Clinical Child and Family Psychology Review, Vol. 8, No. 3, September 2005 (© 2005)
DOI: 10.1007/s10567-005-6664-5

Defining the "Disruptive" in Preschool Behavior: What Diagnostic Observation Can Teach Us

Lauren S. Wakschlag,[1,6] Bennett L. Leventhal,[1] Margaret J. Briggs-Gowan,[3] Barbara Danis,[1] Kate Keenan,[2] Carri Hill,[1] Helen L. Egger,[4] Domenic Cicchetti,[3] and Alice S. Carter[5]

This paper presents the clinical/developmental framework underlying a new diagnostic observational tool, the Disruptive Behavior Diagnostic Observation Schedule (DB-DOS). The special importance of observation for clinical assessment during the preschool period is delineated. The developmental rationale for a multi-dimensional assessment of disruptive behavior in young children, including problems in modulation of negative affect and low competence is discussed. The ways in which the DB-DOS will elucidate distinctions between normative and atypical behavior during this developmental period via (a) the integration of qualitative and quantitative dimensions of behavior within a clinically-sensitive coding system, (b) the observation of child behavior both within, and outside of, the parent–child context and (c) the use of specially designed tasks to "press" for clinically salient behaviors are addressed. The promise of this new method for yielding a more precise, developmentally based description of the phenotype of early onset disruptive behavior problems and for providing a standardized clinical tool for observational assessment of disruptive behavior in young children is presented. Large-scale validation of the measure is currently underway.

KEY WORDS: preschool disruptive behavior; behavior problems; developmental methods; developmental psychopathology; observational assessment.

INTRODUCTION

Disruptive behavior problems are the most common reason for mental health referral of preschool children and, these problems are often persistent and impairing (Campbell, 2002; Wakschlag & Danis, 2004).[7] However, because this developmental period is marked by developing autonomy and a great deal of behavioral change, most preschoolers exhibit at least some of the behaviors that fall under the rubric of disruptive behavior. This makes clinical assessment of preschool behavior particularly subtle and complex. Consequently, there has been a lack of consensus about how to validly assess clinical problems

[1]Institute for Juvenile Research, University of Illinois at Chicago.
[2]Department of Psychiatry, University of Chicago, Chicago, Illinois.
[3]Epidemiology and Public Health, Yale University, New Haven, Connecticut.
[4]Psychiatry and Behavioral Sciences, Duke University, Durham, North Carolina.
[5]Department of Psychology, University of Massachusetts at Boston, Boston, Massachusetts.
[6]Address all correspondence to Lauren S. Wakschlag, Ph.D., Institute for Juvenile Research, Department of Psychiatry, University of Illinois at Chicago, 1747 W. Roosevelt Rd, Room 155, MC 747, Chicago, Illinois 60608; e-mail: lwakschlag@psych.uic.edu.

[7]Historically attention deficit/hyperactivity disorder (ADHD) has been included within the broader rubric of externalizing problems. However, DSM-IV differentiates Disruptive Behavior Disorders (DBDs, i.e., Oppositional Defiant Disorder (ODD) and Conduct Disorder (CD)) from ADHD, with DBDs being fundamentally social rather than attentional in nature. This distinction is supported by research showing different developmental pathways and correlates for DBDs and ADHD (Hinshaw & Lahey, 1993; Wakschlag & Hans, 1999). We will use the term "disruptive behavior" to refer to oppositional and conduct problems, the primary focus of this paper.

during this period (Campbell, 2002). This is especially problematic because behavior problems are the most common reason for referral of preschoolers to mental health clinics. Thus, the development of standardized methods that provide a metric for developmentally sensitive clinical decision-making with preschool children is crucial. Such methods will substantially improve our ability to appropriately distinguish transient and/or normative behaviors in preschoolers from behavior disorders requiring treatment and to delineate the nature of early emerging pathways to disruptive behavior disorders.

The goal of this paper is to elucidate how diagnostic observation may be used to advance scientific and clinical knowledge in the area of preschool disruptive behavior.

Measurement of Disruptive Behavior in Young Children: Challenges and Opportunities

The development of self-control is a hallmark of the toddler and preschool periods (Kochanska, Coy, & Murray, 2001; Sroufe, 1996). Frustration tolerance, delay of gratification, the use of verbal negotiation strategies, internalization of standards, and behavioral flexibility are skills that emerge and are consolidated during this period. Developmental contexts and demands also shift during this period, including expansion of the social world to include peers, heightened parental expectations, and limit setting. As these many processes converge, there is a normative increase in behavioral disruption (Campbell, 2002; Tremblay, 2000). Distinguishing typical from atypical behavior during this developmental period thus presents significant diagnostic and methodologic challenges.

Despite these complexities, dramatic progress has been made in the conceptualization and measurement of social–emotional problems and psychopathology in young children over the past decade (DelCarmen-Wiggins & Carter, 2004). A large body of work has established that problem behavior patterns begin in the first years of life (Carter, Briggs-Gowan, Jones, & Little, 2003). Most recently, preliminary validation efforts have demonstrated that clinical problems can be reliably identified in preschool children, based on parent report (Task Force, 2003). Though these studies have focused primarily on concurrent validity (Keenan & Wakschlag, 2002), a few studies have examined the predictive validity of DBDs identified in preschoolers (Lavigne et al., 1998; Speltz, McMellan, DeKlyen, & Jones, 1999). These clinical studies, as well as a large body of developmental research on preschool behavior problems assessed with behavior checklists (e.g., Campbell, 2002; Dumas, Serketich, & LaFreniere, 1995; Gardner, Ward, Burton, & Wilson, 2003; Shaw et al., 1998), have provided a strong empirical foundation for the study of early emerging psychopathology.

However, advancing clinical science in this area requires supplementing information from parent report with data derived from clinically informative observational methods. Observational methods are vitally important for elucidating the boundaries between normal and abnormal behavior in young children and for characterizing the phenotype of disruptive behavior in young children in all its subtlety and complexity. Observational methods are important supplements to parent report because although parents are good historians about children's behavior, they are often less able to make judgments about whether behaviors are normative or atypical (Lord, 1997). In addition, parents' ability to report "accurately" is influenced by a variety of factors, including threshold for misbehavior, developmental knowledge, stress, and parental psychopathology (Briggs-Gowan, Carter, & Schwab-Stone, 1996; Hay et al., 1999). Further, distinguishing normative variation from clinical problems in early childhood requires systematic, nuanced observation of a range of behaviors, and their qualitative features. Historically agreement between direct observation and parent report of preschoolers' behavior has been modest at best, suggesting that each provides unique information (Gardner, 2000; Kagan, 1998). In the discussion that follows, we present diagnostic observation as a promising and innovative methodology designed to supplement parent-report methods. We introduce the Disruptive Behavior Diagnostic Observation Schedule (DB-DOS), a method we have recently developed for observational assessment of disruptive behavior in young children, provide preliminary data on its reliability and validity, and highlight its potential for advancing the science of developmental psychopathology.

The Role of Diagnostic Observation in the Assessment of Young Children

Definition and Potential Utility

Observational methodologies are a longstanding tradition in developmental and clinical research,

including structured laboratory-based parent–child interaction and behavioral analogue methods and home-based naturalistic observation methods (e.g., Buckley, Klein, Durbin, Hayden, & Moerk, 2002; Elardo & Bradley, 1981; Haynes, 2001; Hubbard et al., 2002; Kochanska, 1998; Pelligrini, 2001; Shaw, 2003; Zelenko, 2004). Each of these methods has unique strengths and weaknesses, which we compare and contrast to diagnostic observation below.

In developmental research, *laboratory-based observational methods* have generally focused on structured assessments (e.g., free play, family discussions) designed to elicit characteristic styles of parent–child interactions (Kerig & Lindahl, 2001). Such methods allow for assessment of discrete child behaviors as well as parenting. Typically, a range of tasks is used that include both "demand" and free play tasks. Within the context of these tasks, parents are often encouraged to respond as naturally as possible. One of the most widely used preschool paradigms is the Dyadic Parent–Child Interaction Coding System (D-PICS; Robinson & Eyberg, 1981), a 15 minute paradigm that includes three structured situations: a child-directed interaction, a parent-directed interaction, and a cleanup task. Frequency counts of a range of child behaviors are obtained and then summed to create three total scores: child deviance (sum of whining, yelling, crying, physical negative, aggression, and smart talk), noncompliance (percent of noncompliance in response to parental commands), and child negative affect (Webster-Stratton & Lindsay, 1999). Multiple studies have demonstrated that D-PICS conduct problem scores are higher for referred versus nonreferred preschoolers (Eyberg, Boggs, & Algina, 1995; Speltz, DeKlyen, Greenberg, & Dryden, 1995; Webster-Stratton & Lindsay, 1999). However, in these and other studies, parent and child behavior ratings have been treated as correlates of DBDs rather than used to generate diagnoses (e.g., Speltz et al., 1995;Webster-Stratton & Lindsay, 1999; Wakschlag & Keenan, 2001).

Another type of laboratory-based observation is the "*behavioral analogue*" or "*performance-based*" task, designed to tap into very specific clinically relevant processes (Haynes, 2001). For example, in the study of DBDs in older children, analogue methods have been used to study cheating, social information processing, and response to provocation (Criss, Shaw, & Ingoldsby, 2003; Frick & Loney, 2000). Behavioral analogue tasks include simulations (e.g., computer pinball game in competition with an alleged peer), tasks with the child alone, and tasks where the parental behavior is scripted (e.g., parents are instructed to use a standardized set of commands). Although more commonly used with older children, similar methods have been used with preschool children, including the use of delay tasks to examine anger regulation and impulsivity (Campbell, Pierce, March, Ewing, & Szumowski, 1994; Gilliom, Shaw, Beck, Schonberg, & Lukon, 2002).

In contrast, *home-based observations* are designed to more naturalistically capture child behavior within the context of routine activities of daily life. Such observations generally involve multiple visits to a home to obtain a varied sampling of behaviors. Families are observed engaging in "typical daily activities." Although virtually all home-based observations include an unstructured component, some also include structured tasks in the home (e.g., cleanup, family meals) (Bradley, Mundfrom, Whiteside, Casey, & Barrett, 1994; Buckley et al., 2002; Gardner, 2000; Webster-Stratton & Lindsay, 1999). Coding methods have varied, ranging from event-based coding to q-sort methodology (Buckley et al., 2002; Gardner, 2000). Rates of noncompliance observed in the home distinguish parent-rated problematic preschoolers from normal controls (Gardner, Sonuga-Barke, & Sayal, 1999).

Standardized *diagnostic observation* provides a direct examiner-based method to systematically elicit the full range of behaviors relevant to the diagnosis of a specific disorder. Diagnostic observation is intentionally structured to "press" for salient behaviors, thereby increasing the likelihood that a range of clinically relevant behaviors will be observed (Lord et al., 2000). The strength of diagnostic observation methods is that they inherently combine research and clinical utility. That is, they are designed to be administered in a standardized fashion, clinically sensitive and relatively brief, and thus feasible for clinicians to use (Mash & Foster, 2001).

By providing a semi-structured standardized method of direct assessment, diagnostic observations also allows for the use of clinical judgment. *Clinical observation* goes beyond observation of discrete behaviors *per se*, offering an integrated examination of multiple facets of the child's behavior and functioning (Wakschlag & Danis, 2004). Building on such observations, *clinical judgment* is an overall assessment of the atypicality of the child's behavior, which weights the salience of particular behaviors based on age-appropriateness and context. The clinician's experience of, and with, the child is a vitally important

dimension of this process (Benham, 2000; Wakschlag & Danis, 2004; Zelenko, 2004).

Standardized diagnostic observation has been used very effectively in research and clinical practice with autism-spectrum disorders. The Autism Diagnostic Observation Schedule (ADOS; Lord et al., 2000) is an interactive schedule that uses examiner "presses" to elicit behaviors fundamental to the diagnosis of autism (i.e., communication, reciprocal social interaction and stereotyped behavior). In conjunction with a standardized diagnostic interview, the ADOS has become an essential part of a "gold standard" battery for diagnostic assessment of autism and has contributed to broad-based consensus about diagnostic validity (Lord, Rutter, & Le Couteur, 1994). In particular, it has been crucial for resolving the controversy about whether autism can be reliably identified in preschool children (Lord, 1995). Further, by providing a method with which to characterize phenotypic variations, it has been a cornerstone of genetic research and studies of diagnosis, treatment and developmental course (Kim et al., 2002; Lord, Leventhal, & Cook, 2001; Owley et al., 2001).

Diagnostic Observation in Context

Diagnostic observations have the capacity to provide very important, unique information but they also have limitations. They assess current, not lifetime symptomatology. The information yielded will necessarily be "de-contextualized" relative to naturalistic or multiple, repeated observations. Further, since low incidence behaviors may not occur in a relatively brief observation, their presence is important but their absence cannot be interpreted as the absence of a symptom (Lord, 1997). Thus, diagnostic observations should not be used in isolation. Rather, they are designed to be used as companion methods to parent interviews, which place these observations in *historical context*. Assessment of historical context includes interviewing parents about the developmental course of the behavior, life events that may have precipitated the onset of the behavior, and events and stressors within the broader social context. Parent-interview data also provide a window into the subjective meaning of the child's behavior to the parent; data which may be informative above and beyond "objective" ratings (Campione-Barr & Smetana, 2004). Information from other key informants in the child's life (e.g., teachers) is also crucial for interpreting whether behavior observed represents enduring patterns (i.e., is "typical" for the child) or is a perturbation due to normative developmental transitions or recent disruptions in the child's social environment.

Distinguishing typical from atypical behavior also requires interpretation of behaviors observed within *developmental context* (Wakschlag & Danis, 2004). As such, standardized evaluation of the child's developmental functioning is also a critical component of a developmentally sensitive assessment battery (for a more extensive discussion, see Carter, Briggs-Gowan, & Ornstein Davis, 2004; Wakschlag & Danis, 2004). This enables control for developmental level when examining the discriminative and predictive utility of observational data as well as determining atypicality based on both chronological and developmental age. For example, noncompliance and negative affect may be interpreted differently if language delays are present.

Varying Observational Methods: Tradeoffs and Opportunities

As we have noted above, each of the various types of observational methodology have unique advantages and disadvantages and such trade-offs must be acknowledged and thoughtfully examined. Using a range of methodologies, studies have correlated observed disruptive behavior with parent-reported behavior problems in preschool children. Such methods have also highlighted dimensions of parenting behavior central in pathways to psychopathology. Their clinical utility is limited, however, because they: (a) measure these behaviors in terms of presence/absence or frequency and do not take qualitative dimensions into account; (b) do not comprehensively assess the constellation of behaviors relevant to a diagnosis of a DBD; (c) are not clinically sensitive or specific, and (d) do not include a method of direct examiner-based clinical assessment. Further, since most methods have relied exclusively on observation of child behavior within the context of parent–child interaction, the extent to which the child's behavior is pervasive or specific to that relationship cannot be systematically observed. With the growing focus on elucidating the nature and pattern of early emerging psychopathology (Task Force, 2003), diagnostically informative observational methodology is increasingly imperative (Carter et al., 2004).

One of the primary advantages of diagnostic observation (i.e., that it is structured to increase the

likelihood of eliciting behaviors of interest) is also its disadvantage (i.e., it is not naturalistic). As Gardner (2000) has noted, the validity of observational methods may be affected by both type of task (e.g., structured versus naturalistic) and location of observation (e.g., home versus lab). As a result, choice of an observational method must carefully consider the unique advantages and disadvantages of each method and the appropriateness of a method within a particular research study must be driven by the particular questions the study is designed to answer. Thus, if the goal is to enhance the accuracy of diagnostic decision-making and characterize the phenotype of disruptive behavior disorders in young children, structured diagnostic observation has significant advantages because the full range of behaviors are unlikely to be observed in an unstructured naturalistic context, unless repeated observations are conducted. Second, although there are clearly significant research advantages to naturalistic observations, they are not practical as clinical tools.

The Disruptive Behavior Diagnostic Observation Schedule (DB-DOS)

The Disruptive Behavior Diagnostic Observation Schedule (DB-DOS; Wakschlag et al., 2002) is an observational clinical research tool for assessing disruptive behavior in young children. The aim of the DB-DOS is to provide a developmentally informed, examiner-based method for the clinical assessment of disruptive behavior in young children that will yield information essential for characterizing the phenotype of early emerging disruptive behavior and serve as a companion to parent-interview methods. In the present paper, we focus primarily on the conceptual framework and methodology of the DB-DOS and highlight the ways in which it has been designed to advance clinical science in the area of preschool disruptive behavior. We have recently been funded by the National Institute of Mental Health to conduct a large-scale validation of the DB-DOS measure (L. Wakschlag, PI). This study, which is designed to establish the reliability and validity of the DB-DOS, will include extensive psychometric analysis; thus a detailed discussion of reliability, validity, and psychometric properties of the measure is premature. Here we provide preliminary evidence of the DB-DOS' reliability and validity, including evidence from a pilot sample of referred and nonreferred preschoolers ($N = 35$) (Wakschlag & Danis, 2004).

DB-DOS Method

Using the ADOS paradigm as a model, the DB-DOS (a) uses tasks developed to systematically "press" for behaviors relevant to DSM diagnostic criteria, (b) combines structured guidelines for examiner behavior (designed to standardize the context and allow salient behaviors to unfold) and the use of clinical judgment (to allow for flexibility and responsiveness to individual differences amongst children), (c) is relatively brief to administer (approximately 50 min), and (d) rates behaviors along a continuum of atypicality in order to assess behavior in a clinically informative manner. However, by definition, diagnostic observations must have components that are unique to the disorder in question because they are specifically designed to press for behaviors salient to this constellation of problems. As a result, the DB-DOS paradigm and coding system also *differ* from those of the ADOS in several fundamental ways that reflect methodologic issues unique to the assessment of disruptive behavior problems.

DB-DOS Structure

The DB-DOS is structured into three mini-contexts, called *modules* (see Fig. 1). Autism spectrum behaviors tend to be pervasively impairing; as such the ADOS diagnostic assessment is with the examiner alone. In contrast, disruptive behaviors may vary substantially across context, with particular salience of the parent–child context. Assessment of the child in interaction with the parent is most reflective of day-to-day interactions, but also makes it difficult to systematically assess the child's capacity for self-regulation since parents vary substantially in their skillfulness and, in fact, may exhibit problematic behaviors that amplify child behavior problems (Patterson, Capaldi, & Bank, 1991). In contrast, interactions between a child and the examiner lack this "history" and can be more easily standardized across children. As a result, the DB-DOS is designed to press for contextual differences in behavior and includes three modules: two with the examiner (Modules 1 and 2) and one with the parent (Module 3).

Modules 1 and 2, the examiner modules, are designed to systematically assess the way in which external structure may affect the child's capacity to adaptively manage behavior and emotions via systematic variation in the level of support provided

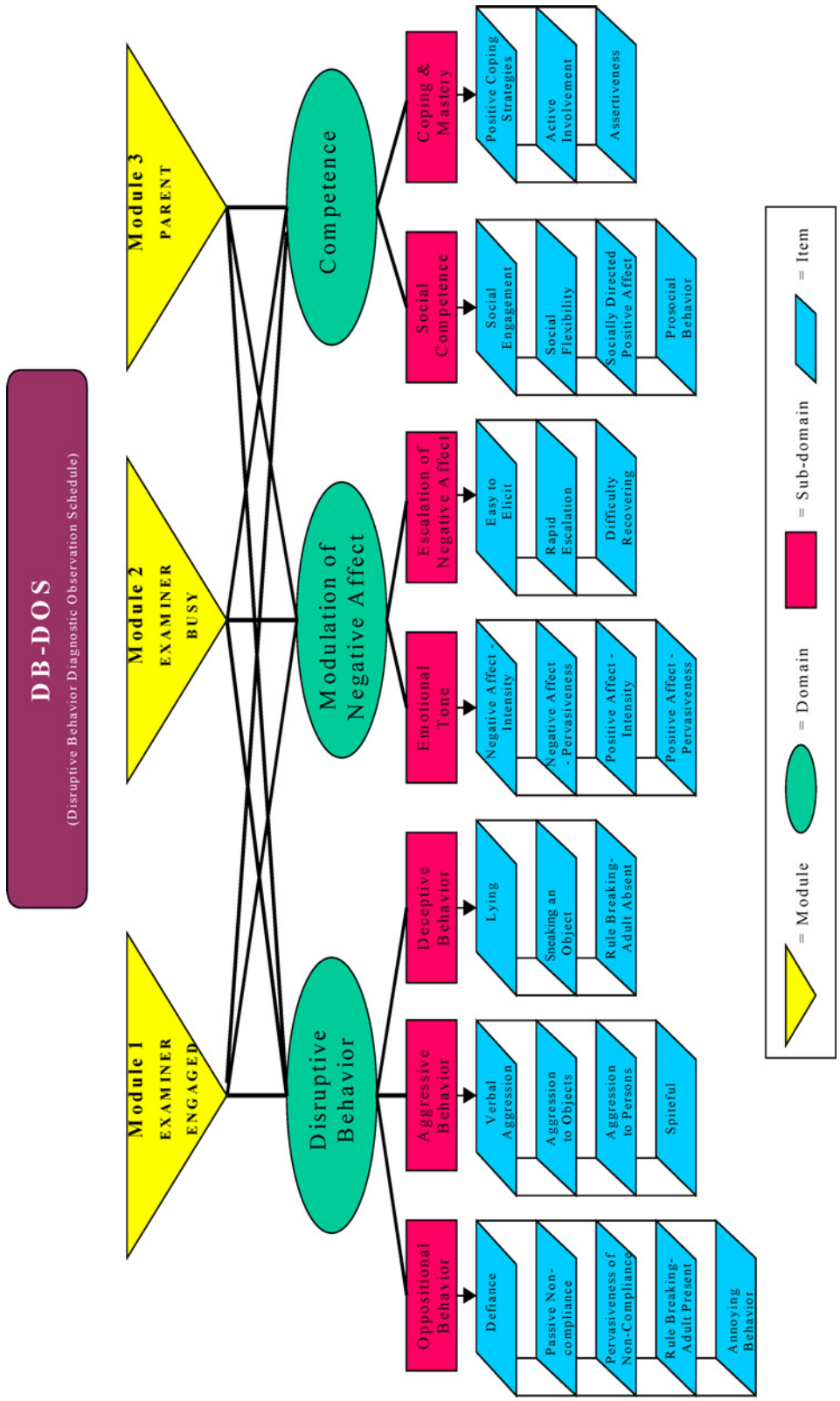

Fig. 1. DB-DOS schematic.

by the examiner. Across all tasks, the examiner responds to the child in a low-key, warm, and natural way. The goal is to put the child at ease in order to allow the child's natural behavior to unfold. Thus, the examiner seeks to strike a balance by making the child feel comfortable without altering or obscuring the child's typical presentation. Within these guidelines, it is emphasized that the examiner's behavior not be rigidly scripted. Rather, clinical judgment is used in determining the nature, pacing, and quality of examiner responses.

The key difference across examiner modules is the extent to which the examiner seeks opportunities to interact with the child. The goal of Module 1 ("*Examiner Engaged*") is to allow the child's behavior to unfold with a responsive examiner. In this module, the examiner is sitting at the child's side and is involved and attentive to the child during all tasks to create "active" opportunities to respond to the child (e.g., smiling, commenting). During this module, the examiner uses normal social responses to child's *positive social- and task-related* behaviors. Within these positive social interchanges, the examiner's responses should be matched to or mirror those presented by the child in terms of modality and intensity, rather than attempting to amplify or modify child behavior.

Module 2 ("*Examiner Busy*") is designed to provide a "withdrawal of attention" context and to press for covert behaviors (e.g., lying). During this module, the examiner does not actively seek opportunities to respond to the child and only engages minimally with the child in response to child initiations. In order to naturally create this structure within this module, the child is given tasks to do independently while the examiner is "busy" during this module with his/her own "work." The examiner may respond to active initiations by the child in a warm and natural manner but with the goal of encouraging the child to resume independent activity.

In contrast, examiner response to child *disruptive behaviors* occurs in essentially the same manner in both modules. Within the DB-DOS context, disruptive behaviors are defined as oppositional, aggressive, and destructive behaviors and/or moderate-to-high levels of irritable/angry affect. Responses to disruptive behavior are guided by the fundamental goal of the DB-DOS to allow the full range of the child's disruptive behavior to emerge, while at the same time providing scaffolded support as a means of testing "what it takes" to help a child exhibiting disruptive behavior to modulate his/her response. Since scoring is based, in part, on the child's responsiveness to examiner support, a hierarchy of well-paced prompts is used to guide examiner responses.

A Level 1 prompt serves as a reminder; thus, Level 1 prompts are designed to redirect the child to the task at hand, and typically utilize restatement of the directions (e.g., "remember you have to do your job before you can do the next thing"). Level 2 prompts provide the child with encouragement. (e.g., "let's see how fast you can do it" "you can play with those later but you need to wait"). Level 3 prompts increase the level of examiner involvement by providing a contingency (e.g., showing child the next toy to be played with once the current task is completed), giving a warning (e.g., "if you can't do it, I can help you"), or trying to join with the child around the tasks (e.g., "let's try it together"). A Level 4 prompt is designed to de-escalate very intense behaviors and may include physical or verbal support for task completion (e.g., helping the child to complete the task, stating directions in a firm tone, and/or moving on to the next task).

Examiners move through the hierarchy of prompts in response to escalating, persistent or serious disruptive behavior. Examiners are trained not to step in too quickly; prompts should be paced to allow the child sufficient time to modulate his/her own behavior. In the event of unsafe or very rapidly escalating behavior (e.g., throwing things, trying to leave the room, aggression towards the examiner), examiners use their judgment to move through the prompts more quickly or begin at a higher level; thus, the examiner may skip earlier levels and provide a Level 3 or 4 prompt as his/her first response.

In contrast to the examiner modules, structuring in the parent module occurs via the variation of task demands rather than by scripting the parent's behavior. This allows the parent's behavior to unfold as naturally as possible. The parent module also provides a standard context with which to assess parental behavior in a clinically informative manner.

DB-DOS *tasks* are activities within the modules that are designed as "presses" to elicit salient behaviors. In developing these tasks, we drew on a number of widely used developmental paradigms to create a series of presses that would be likely to elicit the full range of clinically relevant behavior. Such tasks included compliance "do" and "don't" tasks, cleanup, withdrawal of attention, social play, and frustration tasks; tasks that have been validated across a range of samples of varying ethnicity and socioeconomic status (Campbell, Szumowski, Ewing,

Table I. Overview of DB-DOS Tasks by Module

Behavioral press	Task description
Module 1: Examiner Engaged	
Compliance "Do"	Three consecutive sorting tasks
Frustration	Bubble toy demonstrated but does not work for child
Social play	Examiner and child play with a marble construction toy
Module 2: Examiner Busy	
Compliance "Don't"	Boring crayon task with prohibited toys nearby[a]
Frustration	Trying to complete puzzle to win prize with rigged task[a]
Internalization of rules	Child prohibited from touching novel toy when examiner briefly leaves the room
Module 3: Parent	
Compliance "Do"	Coloring + Clean-up
Frustration	Multi-step puzzle task
Compliance "Don't"	Parent completes questionnaire, child may read book but prohibited from touching toys on shelf
Social play	Parent and child play together with shelf toys

[a]Child is permitted to play with these toys at the close of the session and is always awarded a prize for making "good effort."

Gluck, & Breaux, 1982; Gardner et al., 1999; Garner & Power, 1996; Gilliom et al., 2002; Kochanska & Aksan, 1995; Robinson & Eyberg, 1981; Wakschlag & Keenan, 2001; Webster-Stratton & Lindsay, 1999). Thus, the innovation of the DB-DOS does not lie in the originality of the tasks *per se*. Rather, its unique contribution is in the compilation of a battery of tasks that collectively press for the broad range of behaviors clinically salient in the assessment of DBDs and, in the DB-DOS rating system designed to capture these behaviors in a clinically informative fashion.

There are three tasks in each of the examiner modules and four tasks in the parent module, each approximately 5 min in length. Thus, the entire DB-DOS paradigm takes approximately 50 min. Tasks were designed to be parallel across examiner and parent modules (Table I summarizes DB-DOS tasks by module).

Rating Child Behavior on the DB-DOS

In the development of the DB-DOS, we were guided by two fundamental principles. One was the importance of clinical judgment for distinguishing between typical and atypical behavior in young children (Gilliam & Mayes, 2004; Wakschlag & Danis, 2004). As discussed above, clinical judgment is used in DB-DOS *administration* in terms of determining the timing, pacing, and nature of examiner responses to child behavior. For DB-DOS *scoring* on the other hand, clinical judgment is built into the codes rather than via independent inference by the examiner. That is, the identification of clinically relevant items and clinical inference (i.e., judgment about the atypicality of a behavior based on an integrated view of its quantitative and qualitative features) were used to construct the codes. Thus, for example, clinical expertise drove the selection of "ease of elicitation," "rapid escalation," and "difficulty recovering" as clinically relevant dimensions of problems in modulation of negative affect. Consistent with the framework of clinical observation (Gilliam & Mayes, 2004), the rating system goes beyond discrete behaviors by taking their qualitative features into account. For example, for the item "easy to elicit negative affect" (see Table II), negative affect elicited in response to positive social stimuli is

Table II. Illustrate DB-DOS Item: Easy to Elicit Negative Affect

None (0)	Low (1)	Moderate (2)	High (3)
Negative affect not easily elicited. This may include children who take a *lot* before they get frustrated and then only display low-level negative affect on 1–2 occasions.	Negative affect elicited a few times in response to a build up of frustration, limits or demands.	Negative affect elicited several times in response to even low level frustrations, limits or demands, but not in response to social interactions	Negative affect elicited frequently, and is in response to both low level frustrations, limits, demands and to social interactions

Note. This variable assesses the ease with which negative affect is elicited. "Ease" is based on the frequency but also takes into account the nature of the precipitant. Thus, for example, children whose negative affect is elicited at the beginning of demanding tasks *before* frustration builds and children who often exhibit negative affect during fun tasks would be rated in the moderate to high range for this code. This variable assesses elicitation of *any* negative affect, regardless of intensity.

weighted more heavily along the atypicality scale (coded as a 3) than negative affect elicited to frustration or demands (codes of 1 or 2), even if these occur at the same frequency. As such, while clinical knowledge is an important prerequisite for the clinical use of the DB-DOS, the actual rating of behaviors observed is specified along the continuum of each code, rather than via clinician inference. This structure was designed to get the "best of both worlds," i.e., to make use of clinical expertise while at the same time obtaining equivalent data from one case to the next (Westen & Weinberger, 2004).[8]

The second guiding principle was the importance of a developmentally sensitive and -informed methodology. This included creating a paradigm: (a) informed by the central developmental tasks of the preschool period; (b) with developmentally appropriate tasks, and; (c) testing and validating the measure with preschool populations so that individual differences across this age period could be captured and atypical behaviors could be identified relative to normative behaviors in this age group. It was this principle that led us to design the DB-DOS to examine deficits within the context of developmental domains (rather than as isolated symptoms). We used DSM-IV DBD nosology as a starting point for delineating the core developmental domains impaired in children with disruptive behavior problems. In delineating these core domains, we also went beyond DSM nosology to include a domain that has been conceptualized by developmental theorists as playing a fundamental role in disruptive behavior trajectories (i.e., problems in competence) (Dodge, 1993; Flanagan, Bierman, & Kam, 2003; Webster-Stratton & Lindsay, 1999). This enabled us to take advantage of the benefits of both a "top-down" approach (i.e., applying the well-validated constructs of DSM DBDs for school-age children to preschoolers) and a "bottom-up" approach (i.e., defining problems in relation to developmental domains). We believed that this combined approach would enable us to best identify the *defining features of disruptive behavior in preschoolers, including unique developmental features, and identification of clinically relevant items.*

Drawing on clinical and developmental research (Calkins & Dedmon, 2000; Campbell, 2002; Cole, Michel, & Teti, 1994; Dodge & Crick, 1990; Flanagan et al., 2003; Keenan, 2000; Kochanska et al., 2001; Luthar, Burack, Cicchetti, & Weisz, 1997; Shaw, Gilliom, & Giovannelli, 2000), we identified three core areas of behavioral and socio-emotional functioning of particular salience for understanding the development of disruptive behavior disorders in young children: *behavioral control*, *emotion modulation,* and *social orientation.* In each area, we conceptualized behavior along a continuum from problematic to competent.

Behavioral control reflects the child's capacity to regulate his/her behavior in response to social demands and emotional experiences. For children with disruptive behavior, *problems in behavioral control* may include a tendency to respond aggressively when angry and intransigence. In contrast, *competencies in behavioral control* may include the presence of behavioral coping strategies and assertiveness. *Emotion modulation* reflects the child's capacity to modulate the intensity, duration, and appropriateness of his/her response to emotionally arousing situations and stimuli. Children with disruptive behavior often have *problems in modulation of negative affect*, including responding intensely to frustration, difficulty recovering when distressed, rapid escalation of upset and chronic negative mood. *Competencies in emotion modulation* include positive affectivity and the ability to maintain emotional equilibrium in the face of frustration. *Social orientation* reflects the child's responsiveness to, interest in, concern for, and active engagement with his/her social environment. Children with disruptive behavior often have *problems in social orientation* such as lack of empathy, with concomitant poor internalization of social norms (e.g., lying, deliberate rule breaking), a tendency to respond in a hostile, irritable manner to social bids, and antisocial behaviors (e.g., spiteful and deliberately annoying behavior). In contrast, *competencies in social orientation* include a high level of responsiveness to, interest in, and initiative with the social environment, use of social strategies for coping (e.g., asking for help), and prosocial behaviors (e.g., kind, thoughtful and empathic behaviors).

Using this theoretical framework as a guide, we then "deconstructed" DBD symptoms and other core behaviors into their salient qualitative and quantitative dimensions to develop the items for the DB-DOS rating system. This "deconstruction" and subsequent development of integrative codes is fundamental to the goal of the DB-DOS, i.e., capturing clinically salient behaviors in an integrated manner. Coding is done globally via review of videotapes, with items rated separately for each module.

[8]Clearly, clinicians must then use their clinical judgment to "weight" information derived from this rating system with other information gathered in the diagnostic decision-making process.

Preliminary evidence indicates good inter-rater reliability, with an average intraclass correlation of .84 (Wakschlag & Danis, 2004).

Items on the DB-DOS are organized within three primary *domains,* each of which have several *sub-domains* (see Fig. 1). The *Disruptive Behavior* domain is comprised of three sub-domains: *oppositional*, *aggressive*, and *deceptive* behavior. The *Modulation of Negative Affect* domain has two sub-domains: *negative emotional tone* and *escalation of negative affect*. The third domain assesses child *Competence* and is comprised of two sub-domains: *social competence* and *coping/mastery.* Preliminary evidence suggests that the DB-DOS domains are internally consistent (Cronbach's alphas ranging from .76 to .85) and distinguish between referred and nonreferred children, with referred children exhibiting higher Disruptive Behavior and problems in Modulation of Negative Affect and lower levels of Competence ($F_{(1,34)} = 12.94$, $p < .001$) (Wakschlag & Danis, 2004).

Each DB-DOS item is scored along a continuum of either atypicality or competence. As a result, in assessing problematic behaviors, a broad range of typical behaviors is collapsed within the zero category. Thus, in the Disruptive Behavior and Modulation of Negative Affect problem domains, a score of "0" indicates no evidence of deficit or atypicality, "1" indicates a mild form of the behavior, which may or may not be atypical, and scores of "2–3" are clearly atypical (2: clearly atypical, 3: markedly atypical). Conversely, in the Competence domain, a broad range of problem behaviors are included in the "0" category, a "1" indicates a mild form of the behavior, which may or may not be competent, and scores of "2–3" reflect clear evidence of competence. Notably, although frequency of behavior is generally taken into account, its relative weight varies across different items. For example, serious behaviors such as hitting an adult, may be coded as problematic even if they only occur once, whereas noncompliance in response to task directives is more normatively expectable and must occur more frequently in order to be coded as problematic. DB-DOS items are listed in Fig. 1, Table II provides a more detailed illustration of a DB-DOS item.

DB-DOS Scoring

The DB-DOS was developed as a versatile measure, with the potential for use as both a diagnostic tool and as a research method for characterizing the phenotype of disruptive behavior in young children. Thus, our goal is for it to generate a range of scores including: (a) a clinical algorithm that will establish cut-points to distinguish cases, subclinical cases and non-cases; (b) continuous domain scores, and(c) profile scores (i.e., cross-domain combinations).

Validation of the DB-DOS

Validation of the DB-DOS is centered on establishing its clinical utility at multiple levels, including diagnostic utility (i.e., enhancing precision of diagnosis), predictive utility (i.e., enhancing prediction over time, including the incremental utility of the measure over and above existing methods), and conceptual utility (i.e., advancing understanding of a particular type of psychopathology (Vasey & Lonigan, 2000) as well as ecological validity and validity across cultural contexts.

Diagnostic Utility

By comprehensively assessing the constellation of behaviors that comprise DBD symptoms, the DB-DOS is designed to generate a clinical algorithm that can distinguish cases from noncases with high sensitivity and specificity. Data from our pilot sample indicate very good to excellent sensitivity and specificity (92.9 and 85.7% respectively) in distinguishing between referred and nonreferred children (Wakschlag & Danis, 2004). However, extensive psychometric work will be necessary to fully validate the DB-DOS. This will include establishing clinical cut-points and generating a diagnostic algorithm that weights the clinical significance of specific behaviors, incorporates information about pervasiveness across DB-DOS modules, and takes age and sex differences into account. Validation will require an iterative process in which psychometric findings and clinical considerations inform each other.

Another central dimension of this process is determining the manner in which behaviors will be combined across modules. There are many questions to be resolved in this process, such as: (a) Should a behavior be "weighted" equally if it occurs in one versus multiple modules; and, (b) are behaviors observed with examiner or parent differentially salient? Preliminary analyses have indicated substantial heterogeneity in children's pattern of problems across

the DB-DOS contexts. For example, of the children exhibiting problems during the DB-DOS, approximately 40% exhibit problems across both parent and examiner modules, 40% exhibit problems in the parent module only, and 20% in the examiner modules only (Wakschlag, Leventhal, Hill, Danis, & Keenan, 2004). Analyses that examine the significance of these patterns for determining caseness and for predictive validity will critically inform the development of the clinical algorithm. Since the DB-DOS is specifically designed as a companion to interview-based methods, developing mechanisms for weighting and integrating information across informants and contexts (i.e., school, home and clinic) is also fundamental (Carter et al., 2004; Gardner, 2000).

Conceptual Utility

Although there is an increasing support for the identification of clinical problems in young children, the phenotype of disruptive behavior has not yet been well characterized; symptom counts alone are relatively uninformative for predicting individual differences in presentation or persistence (Greenberg, Speltz, DeKlyen, & Jones, 2001). For example, little is known about sex differences in young children with disruptive behavior. By taking a theoretically driven approach with emphasis on *qualitative patterns* of behavior, the DB-DOS is designed to have a high level of conceptual utility. For example, preliminary analyses suggest that disruptive girls are more likely to exhibit problems in Modulation of Negative Affect and disruptive boys are more likely to display problems with Oppositionality and Aggression (Wakschlag et al., 2004). Further, the DB-DOS was designed to enhance understanding of the role that parenting plays in early emerging disruptive behavior trajectories by providing a standard method for assessing disruptive behavior problems both within and outside of the parent–child relationship context.

Predictive Utility

One of the greatest knowledge gaps in the area of preschool disruptive behavior is that of predictive validity. Studies using diagnostic methods with preschoolers are relatively recent and thus, little is known about the predictive validity of DBDs in young children. Existing data suggest moderate stability of disruptive behavior problems in preschool children (Briggs-Gowan & Carter, 1998; Lavigne et al., 1998; Mathiesen & Sanson, 2000; Pierce, Ewing, & Campbell, 1999; Speltz et al., 1999). In order to advance scientific understanding in this area, explorations of predictive validity must go beyond establishing stability *per se* to examining individual differences in these patterns. The DB-DOS was designed to address this issue at several levels. First, by "deconstructing" symptoms to include qualitative features, the DB-DOS lends itself to examination of the predictive value of the nature and pattern of particular types of behaviors in young children. For example, there is evidence that different types of behaviors within the constellation of ODD symptoms differentially predict stability (Speltz et al., 1999; Stormshak, Bierman, & Group, 1998). Thus, the more qualitative and fine-grained assessment of such behaviors on the DB-DOS is likely to be useful predictively. Second, by examining behaviors across multiple domains, the DB-DOS is uniquely suited for person-oriented analyses that can predict individual differences in patterns of persistence and desistance over time.

Ecological Validity

Ecological validity establishes that behavior observed in the laboratory is reflective of "real life" behavior. The DB-DOS is structured to enhance ecological validity by using relatively simple tasks that are reflective of typical experiences of preschoolers (e.g., having to wait, following directions). Data from our pilot sample indicate that parents generally feel that child behavior on the DB-DOS is representative of the child's usual behavior. Nearly two-thirds of parents in our pilot sample rated their child's behavior on the DB-DOS as typical, with the majority of other parents rating their child's behavior as "typically somewhat worse." Interestingly, no parents rated their child's behavior as "typically better," suggesting that observations on the DB-DOS are not over-identifying problems.

Coding Parental Behavior During DB-DOS Module 3

Parallel to the DB-DOS child behavior coding system, the Parenting Clinical Observation Schedule (P-COS; Wakschlag, Hill, Danis, Grace, & Keenan, 2003) was designed to assess parental behavior during the DB-DOS in a clinically sensitive manner. The P-COS draws on parenting coding systems widely

used in developmental research in defining key constructs, including the centrality of both positive and negative dimensions of parenting behavior (e.g., Denham, 1993; Kochanska, 1998; Olson, Bates, & Bayles, 1990; Patterson et al., 1991; Pettit, Bates, & Dodge, 1993; Shaw & Bell, 1993; Wakschlag & Hans, 1999). It uses the clinically oriented structure of the DB-DOS to operationalize these constructs, with problems and competencies coded within each domain. Salient behaviors are also deconstructed so as to "unpack" maternal behavior in a manner that will be informative for treatment. For example, response to misbehavior is coded in terms of *firmness, use of anticipatory versus reactive management strategies,* and the *flexibility with which these strategies are employed.* Similarly, we intentionally code *behavioral responsiveness* and *behavior management* separately because the use of anticipatory strategies may be uniquely important in reducing the risk that behaviors will escalate (Gardner et al., 1999). We also include low incidence but highly pathognomic parenting behaviors in order to capture behaviors that are often "red flags" to the clinician of serious problems (e.g., power struggles). Parenting behaviors rated on the P-COS are illustrated in Table III.

In summary, the DB-DOS, and its companion coding system, the P-COS, were developed in response to both methodologic and substantive gaps in the study of early onset disruptive behavior. We now delineate their potential for advancing clinical science in this domain.

Table III. The Parenting Clinical Observation Schedule (DB-DOS Module 3)

DB-DOS parenting domains
1. Competence
1A. Behavioral
Behavior management
1. Firmness
2. Use of positive behavior strategies
3. Flexibility and modulation of behavior management strategies
Behavioral responsiveness
4. Scaffolding
5. Responsivity to compliance and other positive behaviors
1B. Emotional responsiveness
6. Affectionate behavior
7. Positive engagement
8. Labeling
2. Maternal problem behaviors
9. Spiteful/hostile behavior
10. Verbal aggression toward child—threats
11. Verbal aggression toward child—cursing
12. Physical aggression toward child
13. Engagement in power struggles
14. Emotional misattunement
3. Affective tone
15. Intensity of positive affect
16. Pervasiveness of positive affect
17. Intensity of irritable/angry affect
18. Pervasiveness of irritable/angry affect
19. Anxious/tense
20. Sad/depressed

How the DB-DOS Can Inform Central Questions in the Study of Early Emerging Disruptive Behavior

Question 1: What are the Boundaries Between Typical and Disruptive Behavior in Preschool Children?

The defining feature of DBDs is a pattern of negative behaviors that interfere with social interactions with others (DSM IV, APA, 1994). The hallmarks of ODD are defiance and negative emotionality. The essential features of CD are aggression and rule violation. On the surface, it would appear that these patterns of disordered behavior are virtually identical to the normative behavioral disruption of the preschool years. As a result, in contrast to older children, the presence of such behaviors is *necessary but not sufficient* for determining whether behaviors are problematic at this age. As such, *taking both quantitative and qualitative features into account in an integrated manner is crucial for distinguishing between typical and disruptive behavior in preschool children* (Campbell, 2002; Hay, Castle, & Davies, 2000; Wakschlag & Danis, 2004). While developmentally sensitive parent interview methods can assess each of these features as separate components of behavior, unlike clinical observation they cannot assess the way in which these features "fit together." We suggest below clinical principles that we have developed based on extensive clinical experience with disruptive preschoolers and drawing on developmental research (Wakschlag & Danis, 2004, Campbell, 2002). A primary goal of the DB-DOS is to provide empirical support for these distinctions by generating parameters that identify specific qualitative dimensions of behavior that are discriminative.

Clinical Principle A: Normative Noncompliance in Preschoolers is not Pervasive

Normative or "assertive noncompliance" (Crockenberg & Litman, 1990) is frequent but not

characteristic during the preschool period. Such behavior generally reflects assertions of autonomy and is rarely elicited during interactions that do not involve limit setting. In contrast, clinically significant oppositionality and defiance tend to be pervasive across settings and/or within social relationships and transactions. This is "negativism for its own sake" or a "reflexive no" (Wenar, 1982). On the DB-DOS, we examine pervasiveness of noncompliance in several ways. First, using the code "*pervasiveness of noncompliance"* we rate the extent to which noncompliance is elicited mostly during tasks that involve limit setting or whether it predominates across all types of tasks, including enjoyable ones. The DB-DOS also measures assertive noncompliance via an "*assertiveness* "variable. Second, we can examine the extent to which the child exhibits high levels of noncompliance pervasively across modules or whether noncompliance manifests primarily in one module (e.g., only with parent).

Clinical Principle B: Normative Aggression is Primarily Instrumental Rather than Hostile and Proactive and Largely Reflects Immature Conflict Resolution Skills

Normative aggression during this period tends to reflect an immature strategy for resolving conflict with peers rather than deliberate efforts to hurt another (Hay et al., 2000; Tremblay, 2000) and is generally ameliorated with adult intervention. In contrast, clinically significant aggression is proactive and persistent, has a deliberate, driven quality to it, and may be nasty and/or spiteful. On the DB-DOS, we assess qualitative dimensions (e.g., severity, deliberateness) of a range of aggressive behaviors including *physical* and *verbal aggression, spiteful behavior,* and *aggression towards objects*.

Clinical Principle C: Normative Expressions of Negative Affect are Relatively Well-Modulated

Normative expressions of negative affect are likely to occur in response to frustration or fatigue, are generally of mild-moderate intensity, and, even when high in intensity, have a relatively rapid period of recovery. In contrast, children with disruptive behavior problems tend to have frequent tantrums that are easily elicited, prolonged, and poorly modulated. Thus, on the DB-DOS, we code multiple dimensions of the child's ability to modulate negative affect (e.g., *pervasiveness, ease of elicitation, highest intensity*). For example, the code "*easy to elicit*" incorporates both the frequency of negative affect and the type of task that elicits it (Table II). We also capture qualitative dimensions of emotion modulation in terms of competencies, such as the use of "*positive coping strategies"* in the face of challenge or frustration. Here the emphasis is on both the frequency with which coping strategies are displayed as well as their quality (e.g., the presence of varied coping strategies).

Clinical Principle D: Normative Misbehavior and Negativity are Generally Responsive to Adult Input

In contrast, disruptive behaviors are often entrenched, intransigent, and persistent in the face of adult support. On the DB-DOS we capture this by integrating the intensity of the behavior with its responsiveness to adult input. For example, the codes "*highest level of defiance"* and "*difficulty recovering from negative affect"* rate as most problematic behavior that persists in the face of adult prompts or support.

Question 2: Can DBDs in Young Children be Distinguished from Parent–Child Relationship Problems?

This is a vexing problem, which essentially boils down to the question of whether the constellation of behaviors identified as "disruptive behavior" are *actually* a problem that resides within the child (i.e., a DBD) or whether they reflect behavioral difficulties that are a *reaction to* other problems within the child's social context. We have previously suggested (Keenan & Wakschlag, 2002), that when a *pattern of behavior* is pervasive, persistent, and developmentally impairing it is a clinical problem, *even when problematic parenting or other difficulties in the social context are present.* This reflects our theoretical viewpoint that problem behaviors that impair a child's ability to successfully master normative developmental tasks and hinder participation in developmentally appropriate activities are clinically significant, irrespective of etiology. This viewpoint does not discount the central role that parents play in the development of behavioral regulation in young children, however (e.g., Kochanska et al., 2001). This framework also highlights the fact that there are multiple pathways to the development of early emerging psychopathology,

including the potent contribution of child characteristics (e.g., temperamental vulnerability) to the development of disruptive behavior (Keenan & Shaw, 2003).

The question of preschool psychopathology versus disordered parent–child relationship is hotly debated within the field (Carter et al., 2004; Emde, 2003; Jensen & Hoagwood, 1997; Zeanah, Boris & Scheeringa, 1997). A primary goal of the DB-DOS is to provide data to help answer this question empirically. First, embedding standard assessment of parenting behavior within the diagnostic observation paradigm allows for systematic control of problematic parenting behavior in analyses examining the discriminant, predictive, and clinical utility of child behavior on the DB-DOS. For example, we can examine whether child behavior on the DB-DOS predicts impairment in the school setting, once parenting behavior has been controlled.

Second, standard clinically oriented assessment of parental problems and competencies via the P-COS ratings will enable us to begin to characterize individual differences in parenting patterns amongst the parents of children with disruptive behavior problems. (In contrast, much previous work has compared the parenting received by children with and without behavior problems.) Recently, we conducted preliminary analyses to examine whether highly problematic parenting behavior distinguished children who exhibited problems on the DB-DOS parent-module only. Interestingly, while half of these children had parents displaying highly problematic parenting styles, nearly one-third of this group had parents who exhibited highly competent parenting (Wakschlag et al., 2004). This would suggest that while parent–child relationship problems may be a contributing factor for a sub-group of children with disruptive behavior, there may also be distinct subgroups of children whose behavior problems derive from other pathways, including (a) temperamental vulnerabilities and (b) the interaction of child-specific and parenting factors. Understanding such heterogeneity in causal pathways is a fundamental next step for the field. The opportunity the DB-DOS provides to examine child disruptive behaviors both within and outside of the context of parent–child interaction is likely to be especially informative in this process, because it allows for objective rating of child behavior that is not dependent on parents' subjective experience of the child.

The role of parent behavior in diagnostic conceptualization must also be considered. For our purposes, problematic parenting is considered a risk factor rather than a clinical indicator. Thus, we do not expect to include parenting behavior within the diagnostic algorithm. On the other hand, as noted above, P-COS data on quality of parenting may also be used to delineate sub-types (e.g., disruptive behavior with or without parenting problems). In addition, we will examine whether taking P-COS profile scores into account enhances the predictive utility of the DB-DOS (e.g., examining whether competent parenting predicts desistance).

Question 3: Can the Heterogeneity of Disruptive Behavior in Young Children be Characterized in a Manner that Enhances Prediction and Informs Etiologic Studies?

The *predictive* validity of DBDs in preschoolers is a critical, but relatively untested area of inquiry (but see Speltz et al., 1999). We hypothesize that observational methods will add incremental value in diagnostic classification and prediction over time for several reasons. First, the clinical presentation of young children with DBDs is very heterogeneous ((Wakschlag & Danis, 2004). For example, some children with oppositional-defiant symptoms exhibit chronic negative mood, whereas others present only with emotional dysregulation when upset. It is not known what this heterogeneity "means" in terms of persistence and desistance of disruptive behavior problems. Further, it is not yet clear whether there are distinct disruptive behavior constellations (i.e., separate oppositional and conduct patterns) at this age or whether there is a more nonspecific disruptive behavior disorder in preschoolers. By observing *patterns of behavior*, rather than isolated symptoms, we will be able to systematically characterize individual differences in manifestations (e.g., sex differences, how constellations of behaviors "hang together") in a manner that will enable us to better understand and characterize multiple trajectories over time.

We also contend that the development of a standardized observational method is fundamental to breaking new scientific ground in research on early emerging DBDs, independent of its incremental predictive value. As emphasized in the blueprint for DSM-V, recent advances in developmental science create unparalleled opportunities to broaden the methods and approaches used in refining taxonomies of developmental psychopathology (Pine et al., 2002). A standardized observational method has

unique potential for advancing this line of inquiry in several ways. First, observational methods are uniquely suited to enhancing phenotypic description of disruptive behavior in young children. Advances in genetic research have led to a shift in focus from a search for genetic causes for discrete categorical disorders to one on genetic contributions to dimensional risk factors involved in the multifactorial origins of disorder (Rutter, 1997). Phenotypic description at this level requires "deconstructing" the grosser levels of behavior captured by symptoms into their various components. "Mapping" neuroscientific findings on structural and functional brain deficits onto various types of psychopathology also requires this level of description (Pine et al., 2002). Examining disruptive behavior as a more continuous phenomenon is central to this line of inquiry. All of these approaches necessitate a high level of clarity of description (Pine et al., 2002; Sroufe, 1997; Wakefeld, 1997), which observational methods are especially well-suited to provide.

The DB-DOS may contribute to enhanced phenotypic description in several ways. First, symptoms are "deconstructed" into their various components. As such, *one* DSM symptom such as "touchy/easily annoyed" is measured by three more fine-grained codes on the DB-DOS: *easy to elicit negative affect*, its *rapid escalation,* and *difficulty recovering* from it. Similarly, the symptom "often defies or refuses to comply with adult requests/rules" is captured with four DB-DOS codes assessing intensity, type, and pervasiveness of these behaviors. Second, the inclusion of a set of behaviors that are not part of current DSM nosology, but have frequently been identified as deficient in children with DBDs (i.e., problems in social competence), will enable us to examine whether these behaviors are, in fact, a defining feature of DBDs in young children, thereby expanding beyond the limits of current taxonomies. Further, organizing behavioral assessment in the DB-DOS within developmental domains (i.e., disruptive behavior, modulation of negative affect and competence) rather than in diagnostic categories lends itself to person-oriented analysis. This will enable us to move beyond categorical grouping to typologies of behavior and to assess whether these clusters vary across patterns of disruptive behavior, such as persistence of symptoms, diagnostic status, and level of impairment over time. As has been demonstrated by the significant role the ADOS has played in research on autism, such precise characterization is not only useful for, but also may, in part, *drive* advances in epidemiologic, genetic, developmental neuroscience, and pharmacologic research (Pine et al., 2002; Vitiello, 2001).

Conclusion: Where Do We Go From Here?

We have highlighted the potential of diagnostic observation methods in general, and the DB-DOS in particular, to advance scientific inquiry about the nosology, patterns, and clinical significance of early emerging disruptive behavior. We conclude by outlining fundamental next steps in this process.

Validation of the DB-DOS

As noted above, we are currently conducting a large-scale validation of the DB-DOS measure. The primary goals of this study are: (1) psychometric analysis and refinement of the DB-DOS instrument including (a) developing a clinical algorithm that optimizes sensitivity and specificity for classifying children along a continuum of disruptive behavior problems and (b) testing the incremental value of DB-DOS ratings over and above information derived from parent interview for predicting impairment and persistence of DBDs over time; (2) examining whether DB-DOS scores distinguish children with disruptive behavior problems from children with developmental problems and children who are not pervasively disruptive but are disruptive in the context of problematic parent–child interactions and; (3) examining the predictive value of DB-DOS profiles for disruptive behavior trajectories over time. The validation sample will consist of 360 preschoolers along a disruptive behavior continuum, including children referred for disruptive behavior problems, nonreferred children whose parents have behavioral concerns and nonreferred children without behavioral concerns. Laboratory observations of child behavior on the DB-DOS will be validated in relation to parent and teacher report of child behavior and observations of the child in a real life (i.e., school) setting.

Demonstration of the reliability and validity of the DB-DOS within this study is an important first step towards demonstrating the potential contribution of diagnostic observation to the study of clinical phenomena in young children. This work will lay the foundation for future studies that will examine the validity and reliability of the DB-DOS within social and cultural context. First, examination of behaviors observed on the DB-DOS to child behavior observed

in other real life contexts, such as the home, will be a crucial test of its ecological validity. As has been noted (Gardner, 2000), even assessments with high face validity may not necessarily reflect "real life" behavior. Second, a vitally important next step will be examination of the reliability and validity of the DB-DOS across varying socioeconomic and ethnic contexts. Although the importance of context is widely acknowledged in clinical and developmental research, measurement and model equivalence across contexts is often assumed and rarely tested (Carter et al., 2004; Cicchetti & Aber, 1998; Raver, 2004). As has been noted, a contextual approach is crucial for identifying individual differences in and contextual influences on, risk processes (Cicchetti & Aber, 1998; Raver, 2004). Standardization of the DB-DOS within a representative sample of preschoolers is also an important area for future research.

Clinical Utility

In addition to the scientific issues to be addressed, testing whether the DB-DOS will have substantial "added value" to clinicians is critical to establishing its clinical utility. As noted above, clinicians face the pragmatic issue that there is currently no standardized method for direct clinical assessment of preschool children's behavior, despite high rates of clinical referral (Wakschlag & Danis, 2004). Establishing clinical utility requires demonstrating that: (a) clinical *administration* is feasible; (b) *real-time scoring* is reliable and valid; and (c) that individuals with a range of clinical expertise and experience can be reliably *trained*. Administration of the DB-DOS is relatively brief and is consistent with the length of the ADOS as well as most diagnostic interviews. The DB-DOS was also designed to be a clinically useful measure—such that tasks are as naturalistic as possible, do not feel highly "contrived," and do not require costly or sophisticated equipment or laboratory setups. Testing of the reliability and validity of real-time clinical scoring of the DB-DOS is currently underway. Our ultimate goal is to have an instrument that can be widely used by both research and general clinicians. However, we note that the DB-DOS is not meant to train individuals in clinical observation but rather to provide trained observers with a standard method of clinical observation. Thus, we anticipate that a prerequisite for its use will be training and experience in clinical observation and assessment as well as knowledge of normative development.

With the benefits of diagnostic observation, also come new challenges. By adding an additional source of clinical information, the possibility that these sources will provide discrepant information is increased. Thus, systematic methods for reconciling differences across methods will be needed (Carter et al., 2004). Further, while "going beyond" DSM symptoms may enhance our understanding of individual differences in clinical presentation and trajectories, it also raises the question of whether and how these new dimensions should be incorporated into diagnostic nosology. Whether or not the DB-DOS is sensitive enough to capture treatment gains over time is another question that remains to be tested empirically.

In conclusion, although the DB-DOS holds much promise, the importance of more precise characterization of preschool disruptive behavior extends well beyond a particular diagnostic observation measure. The past decade has taught us much about the emergence and identification of mental health problems in the first few years of life. Delineation of the boundaries between typical and atypical behavior in young children now appears to be possible. While there are many challenges that lie ahead, advancing scientific inquiry in this manner is likely to deepen our understanding of the origins of psychopathology, in all its diversity and complexity, in a manner with direct implications for prevention, assessment, and treatment of clinical problems in young children.

ACKNOWLEDGMENTS

The development of the DB-DOS and the writing of this paper have been supported by NIMH grant 1R01 MH068455-01 to Dr. Wakschlag and 1 R01 MH62437-01 to Dr. Keenan, support to Drs. Wakschlag, Carter, Briggs-Gowan & Egger from National ZERO TO THREE (via a grant to ZERO TO THREE from the Robert Wood Johnson Foundation) and ongoing support by the Walden and Jean Young Shaw Foundation and the Irving B. Harris Center for Developmental Studies to the Department of Psychiatry at the University of Chicago. Dr. Catherine Lord's seminal work on the ADOS was the inspiration for the DB-DOS, her continuing incisive and generous guidance has been vitally important to its development. We are indebted to Dr. Chaya Roth, whose teachings laid the foundation for this work. Dr. Edwin Cook's ongoing critical feedback is gratefully acknowledged. We are also

indebted to Dr. Adrian Angold for his critical comments. We are very appreciative of the contribution of our research team and clinical students whose feedback has greatly enriched this effort. Finally, the DB-DOS is dedicated to the memory of our beloved student and colleague, Kathleen Kennedy Martin, Psy.D., a consummate clinical observer who embodied the principles upon which the DB-DOS rests.

REFERENCES

Benham, A. (2000). The observation and assessment of young children including use of the Infant–Toddler Mental Status Exam. In C. Zeanah (Ed.), *Handbook of infant mental health* (2nd ed., pp. 249–266). New York, NY: Guilford Press.

Bradley, R. H., Mundfrom, D. J., Whiteside, L., Casey, P. H., & Barrett, K. (1994). A factor analytic study of the infant–toddler and early childhood versions of the HOME inventory administered to White, Black, and Hispanic American parents of children born preterm. *Child Development*, *65,* 880–888.

Briggs-Gowan, M., & Carter, A. (1998). Preliminary acceptability and psychometrics of the Infant–Toddler Social and Emotional Assessment (ITSEA): A new adult report questionnaire. *Infant Mental Health Journal*, *19,* 422–445.

Briggs-Gowan, M. J., Carter, A. S., & Schwab-Stone, M. (1996). Discrepancies among mother, child, and teacher reports: Examining the contributions of maternal depression and anxiety. *Journal of Abnormal Child Psychology*, *24*(6), 749–765.

Buckley, M., Klein, D., Durbin, C., Hayden, E., & Moerk, K. (2002). Development and validation of a Q-sort procedure to assess temperament and behavior in preschool children. *Journal of Clinical Child and Adolescent Psychology*, *31,* 525–539.

Calkins, S. D., & Dedmon, S. E. (2000). Physiological and behavioral regulation in two-year-old children with aggressive/destructive behavior problems. *Journal of Abnormal Child Psychology*, *28,* 103–119.

Campbell, S. (2002). *Behavior problems in preschool children: Clinical and developmental issues* (2nd ed.). New York, NY: Guilford Press.

Campbell, S., Pierce, E., March, C., Ewing, L., & Szumowski, E. (1994). Hard-to-manage preschool boys: Symptomatic behavior across contexts and time. *Child Development*, *65,* 836–851.

Campbell, S. B., Szumowski, E. K., Ewing, L. J., Gluck, D. S., & Breaux, A. M. (1982). A multidimensional assessment of parent-identified behavior problem toddlers. *Journal of Abnormal Child Psychology*, *10,* 569–592.

Campione-Barr, N., & Smetana, J. (2004). In the eye of the beholder: Subjective and observer ratings of middle-class African American mother-adolescent interactions. *Developmental Psychology*, *40*(6), 927–934.

Carter, A. S., Briggs-Gowan, M. J., Jones, S. M., & Little, T. D. (2003). The Infant-Toddler Social and Emotional Assessment (ITSEA): Factor structure, reliability, and validity. *Journal of Abnormal Child Psychology*, *31,* 495–514.

Carter, A., Briggs-Gowan, M., & Ornstein Davis, N. (2004). Assessment of young children's social-emotional development and psychopathology: Recent advances and recommendations for practice. *Journal of Child Psychology, Psychiatry and Allied Disciplines*, *45,* 1–26.

Cicchetti, D., & Aber, J. L. (1998). Contextualism and developmental psychopathology. *Development and Psychopathology*, *10,* 137–142.

Cole, P., Michel, M., & Teti, L. (1994). The development of emotion regulation and dysregulation: A clinical perspective. In N. Fox (ed.), *The development of emotion regulation: Biological and behavioral considerations* (serial no. 240, vol. 59, pp. 73–102). Chicago: University of Chicago Press.

Criss, M. M., Shaw, D., & Ingoldsby, E. M. (2003). Mother-son positive synchrony in middle childhood: Relation to antisocial behavior. *Social Development*, *12,* 379–401.

Crockenberg, S., & Litman, C. (1990). Autonomy as competence in 2 year olds: Maternal correlates of child defiance, compliance, and self assertion. *Developmental Psycholology*, *26,* 961–971.

DelCarmen-Wiggins, R., & Carter, A. (2004). *Handbook of infant, toddler and preschool mental health assessment*. New York: Oxford University Press.

Denham, S. (1993). Maternal emotional responsiveness and toddlers' social-emotional competence. *Journal of Child Psychology and Psychiatry*, *34,* 715–728.

Dodge, K., & Crick, N. (1990). Social information-processing bases of aggressive behavior in children. *Personality and Social Psychology Bulletin*, *16,* 8–22.

Dodge, K. (1993). Social-cognitive mechanisms in the development of conduct disorder and depression. *Annual Reviews of Psychology*, *44,* 559–584.

Dumas, J., Serketich, W., & LaFreniere, P. (1995). "Balance of power": A transactional analysis of control in mother-child dyads involving socially competent, aggressive, and anxious children. *Journal of Abnormal Psychology*, *104,* 104–113.

Elardo, R., & Bradley, R. (1981). The Home Observation for Measurement of the Environment Scale: A review of research. *Developmental Review*, *1,* 113–145.

Emde, R. (2003). Commentary: RDC-PA: A major step forward and some issues. *Journal of the American Academy of Child and Adolescent Psychiatry*, *42*(12), 1513–1516.

Eyberg, S., Boggs, S., & Algina, J. (1995). New developments in psychosocial, pharmacological, and combined treatments of conduct disorders in aggressive children. *Psychopharmacology Bulletin*, *31*, 83–91.

Flanagan, K., Bierman, K., Kam, C., & Conduct Problems Prevention Research Group (2003). Identifying at-risk children at school entry: The usefulness of multibehavioral problem profiles. *Journal of Clinical Child and Adolescent Psychology*, *32,* 396–407.

Frick, P., & Loney, B. (2000). The use of laboratory and performance-based measures in the assessment of children and adolescents with conduct disorders. *Journal of Clinical Child Psychology (Special Section: Laboratory and performance-based measures of childhood disorders)*, *29,* 540–555.

Gardner, F. (2000). Methodological issues in direct observation of parent-child interaction: Do observational findings reflect the natural behavior of participants? *Clinical Child and Family Psychology Review*, *3,* 185–198.

Garner, P., & Power, T. (1996). Preschoolers' emotional control in the disappointment paradigm and its relation to temperament, emotional knowledge, and family knowledge. *Child Development*, *67,* 1406–1419.

Gardner, F., Sonuga-Barke, E. J., & Sayal, K. (1999). Parents anticipating misbehaviour: An observational study of strategies parents use to prevent conflict with behaviour problem children. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *40,* 1185–1196.

Gardner, F., Ward, S., Burton, J., & Wilson, C. (2003). The role of mother-child joint play in the early development of children's conduct problems: A longitudinal observational study. *Social Development*, *12*(3), 361–379.

Gilliam, W., & Mayes, L. (2004). Integrating clinical and psychometric approaches: Developmental assessment and infant mental health evaluation. In R. Del Carmen Wiggins &

A. Carter (Eds.), *Handbook of infant, toddler and preschool mental health assessment* (pp. 185–203). New York: Oxford University Press.

Gilliom, M., Shaw, D. S., Beck, J. E., Schonberg, M. A., & Lukon, J. L. (2002). Anger regulation in disadvantaged preschool boys: Strategies, antecedents, and the development of self-control. *Developmental Psychology*, *38,* 222–235.

Greenberg, M. T., Speltz, M. L., DeKlyen, M., & Jones, K. (2001). Correlates of clinic referral for early conduct problems: Variable- and person-oriented approaches. *Development and Psychopathology*, *13*(2), 255–276.

Hay, D. F., Castle, J., & Davies, L. (2000). Toddlers' use of force against familiar peers: A precursor of serious aggression? *Child Development*, *71,* 457–467.

Hay, D., Pawlby, S., Sharp, D., Schmucker, G., Mills, A., Allen, H., & Kumar, R. (1999). Parents' judgements about young children's problems: Why mothers and fathers disagree yet still predict later outcomes. *Journal of Child Psychology, Psychiatry and Allied Disciplines*, *40,* 1249–1258.

Haynes, S. (2001). Introduction to the special section on clinical applications of analogue behavioral observation. *Psychological Assessment (Special Issue on Clinical Applications of Analogue Behavioral Observation)*, *13,* 3–4.

Hubbard, J., Smithmyer, C., Ramsden, S., Parker, E., Flanagan, K., Dearing, K., Relyea, N., & Simons, R. (2002). Observational, physiological, and self-report measures of children's anger: Relations to reactive versus proactive aggression. *Child Development*, *73,* 1101–1119.

Jensen, P., & Hoagwood, K. (1997). The book of names: DSM-IV in context. *Developmental Psychopathology*, *9,* 231–250.

Kagan, J. (1998). Biology and the child. In W. Damon & N. Eisenberg (Eds.), *Handbook of child psychology: Social, emotional and personality development* (Vol. 3, pp. 177–236). New York: Wiley.

Keenan, K. (2000). Emotion dysregulation as a risk factor for psychopathology. *Clinical Psychology Science and Practice*, *7,* 418–434.

Keenan, K., & Shaw, D. (2003). Starting at the beginning: Exploring the etiology of antisocial behavior in the first years of life. In B. Lahey, T. Moffit, & A. Caspi (Eds.), *Causes of conduct disorder and juvenile delinquency* (pp. 153–181). New York: Guilford.

Keenan, K., & Wakschlag, L. (2002). Can a valid diagnosis of disruptive behavior disorder be made in preschool children? *American Journal of Psychiatry*, *59,* 351–358.

Kerig, P., & Lindahl, K. (2001). *Family observational coding systems: Resources for systemic research*. Mahwah, NJ: Erlbaum.

Kim, S., Young, L., Gonen, D., Veenstra-VanderWeele, J., Courchesne, R., Courchesne, E., Lord, C., Leventhal, B., Cook E. Jr, & Insel, T. (2002). Transmission disequilibrium testing of arginine vasopressin receptor 1A (AVPR1A) polymorphisms in autism. *Molecular Psychiatry*, *7,* 503–507.

Kochanska, G. (1998). Mother-child relationship, child fearfulness, and emerging attachment: A short-term longitudinal study. *Developmental Psychology*, *34,* 490–490.

Kochanska, G., & Aksan, N. (1995). Mother-child mutual positive affect, the quality of child compliance to requests and prohibitions, and maternal control as correlates of early internalization. *Child Development*, *66,* 236–254.

Kochanska, G., Coy, K., & Murray, K. (2001). The development of self-regulation in the first four years of life. *Child Development*, *72,* 1091–1111.

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.

Lord, C. (1995). Follow-up of two-year-olds referred for possible autism. *Journal of Child Psychology and Psychiatry*, *36,* 1365–1382.

Lord, C. (1997). Diagnostic instruments in autism spectrum disorders. In D. Cohen & F. Volkmar (Eds.), *Handbook of autism and pervasive developmental disorders* (2nd ed., pp. 460–483). New York: Wiley.

Lord, C., Leventhal, B., & Cook, E. (2001). Quantifying the phenotype in autism spectrum disorders. *American Journal of Medical Genetics: Neuropsychiatric Genetics*, *105,* 36–38.

Lord, C., Risi, S., Lambrecht, L., Cook, E., Leventhal, B., DiLavore, P., Pickles, A., & Rutter, M. (2000). The Autism Diagnostic Observation Schedule-Generic: A standard measure of social and communication deficits associated with the spectrum of autism. *Journal of Autism and Developmental Disorders*, *30,* 205–223.

Lord, C., Rutter, M., & Le Couteur, A. (1994). Autism Diagnostic Interview-Revised: A revised version of the a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *Journal of Autism and Developmental Disorders*, *24,* 659–685.

Luthar, S., Burack, J., Cicchetti, D., & Weisz, J. (Eds.). (1997). *Developmental psychopathology: Perspectives on adjustment, risk and disorder*. New York, NY: Cambridge University Press.

Mash, E., & Foster, S. (2001). Exporting analogue behavioral observation from research to clinical practice: Useful or cost-defective? *Psychological Assessment*, *13,* 86–98.

Mathiesen, K., & Sanson, A. (2000). Dimenisons of early childhood behavior problems: Stability and predictors of change from 18 to 30 months. *Journal of Abnormal Child Psychology*, *28,* 15–31.

Olson, S., Bates, J., & Bayles, K. (1990). Early antecedents of childhood impulsivity: The role of parent-child interaction, cognitive competence and temperament. *Journal of Abnormal Child Psychology*, *18,* 317–334.

Owley, T., McMahon, W., Cook, E., Laulhere, T., South, M., Mays, L., Shernoff, E., Lainhart, J., Modahl, C., Corsello, C., Ozonoff, S., Risi, S., Lord, C., Leventhal, B., & Filipek, P. (2001). Multi-site double-blind, placebo-controlled trial of procine secretin in autism. *Journal of the American Academy of Child and Adolescent Psychiatry*, *40,* 1293–1299.

Patterson, G., Capaldi, D., & Bank, L. (1991). An early starter model for predicting delinquency. In D. Pepler & K. Rubin (Eds.), *The development and treatment of aggression in childhood* (pp. 139–168). Hillsdale, NJ: Erlbaum.

Pelligrini, A. (2001). The role of direct observation in the assessment of young children. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *42,* 861–870.

Pettit, G., Bates, J., & Dodge, K. (1993). Family interaction patterns and children's conduct problems at home and school: A longitudinal perspective. *School Psychology Review*, *22,* 403–420.

Pierce, E., Ewing, L., & Campbell, S. (1999). Diagnostic status and symptomatic behavior of hard-to-manage preschool children in middle childhood and early adolescence. *Journal of Clinical Child Psychology*, *28,* 44–57.

Pine, D., Alegria, M., Cook, E., Jr, Costello, E., Dahl, R., Koretz, D., Merikangas, K., Reiss, A., & Vitiello, B. (2002). Advances in developmental science and DSM-IV. In D. Kupfer, M. First, & D. Regier (Eds.), *A research agenda for DSM-V* (pp. 85–122). Washington, DC: American Psychiatric Association.

Raver, C. (2004). Placing emotional self-regulation in sociocultural and socioeconomic contexts. *Child Development*, *75,* 346–354.

Robinson, E. A., & Eyberg, S. M. (1981). The dyadic parent-child interaction coding system: Standardization and validation. *Journal of Consulting and Clinical Psychology, 49*, 245–250.

Rutter, M. (1997). Nature-nurture integration: The example of antisocial behavior. *American Psychologist, 52*, 390–398.

Shaw, D. (2003). Innovative approaches and methods to study of children's conduct problems. *Social Development*, *12,* 309–314.

Shaw, D., & Bell, R. (1993). Developmental theories of parental contributors to antisocial behavior. *Journal of Abnormal Child Psychology*, *21,* 493–518.

Shaw, D., Gilliom, M., & Giovannelli, J. (2000). Aggressive behavior disorders. In C. Zeanah (Ed.), *Handbook of infant mental health* (2nd ed., pp. 397–411). New York: Guilford Press.

Shaw, D., Winslow, E., Owens, E., Vondra, J., Cohn, J., & Bell, R. (1998). The development of early externalizing problems among children from low-income families: A transformational perspective. *Journal of Abnormal Child Psychology*, *26,* 95–108.

Speltz, M., McMellan, J., DeKlyen, M., & Jones, K. (1999). Preschool boys with oppositional defiant disorder: Clinical presentation and diagnostic change. *Journal of the American Academy of Child and Adolescent Psychiatry*, *38*(7), 838–845.

Speltz, M. L., DeKlyen, M., Greenberg, M. T., & Dryden, M. (1995). Clinic referral for oppositional defiant disorder: Relative significance of attachment and behavioral variables. *Journal of Abnormal Child Psychology*, *23,* 487–507.

Sroufe, L.A. (1997). Psychopathology as outcome of development. *Development and Psychopathology*, *9,* 251–268.

Sroufe, L. A. (1996). *Emotional development: The organization of emotional life in the early years*. New York: Cambridge University Press.

Stormshak, E., Bierman, K., & Group, C. P. P. R. (1998). The implications of different developmental patterns of disruptive behavior problems for school adjustment. *Development and Psychopathology*, *10,* 451–468.

Task Force on Research Diagnostic Criteria: Infancy and Preschool (2003). Research diagnostic criteria for infants and preschool children: The process and empirical support. *Journal of the American Academy of Child and Adolescent Psychiatry*, *42,* 1504–1512.

Tremblay, R. (2000). The development of aggressive behavior during childhood: What have we learned in the past century? *International Journal of Behavioral Development*, *24,* 129–134.

Vasey, M., & Lonigan, C. (2000). Considering the clinical utility of performance-based measures of childhood anxiety. *Journal of Clinical Child Psychology (Special Section: Laboratory and performance-based measures of childhood disorders)*, *29,* 493–508.

Vitiello, B. (2001). Psychopharmacology for young children: Clinical needs and research opportunities. *Pediatrics*, *108,* 983–989.

Wakefeld, J. (1997). When is development disordered? Developmental psychopathology and the harmful dysfunction analysis of mental disorder. *Development and Psychopathology*, *9,* 269–290.

Wakschlag, L., & Danis, B. (2004). Assessment of disruptive behavior in young children: A clinical-developmental framework. In R. DelCarmen-Wiggins & A. S. Carter (Eds.), *Handbook of infant, toddler and preschool mental health assessment* (pp. 421–440). New York: Oxford University Press.

Wakschlag, L., & Hans, S. (1999). Relation of maternal responsiveness during infancy to the development of behavior problems in high risk youths. *Developmental Psychology*, *37,* 569–579.

Wakschlag, L., Hill, C., Danis, B., Grace, D., & Keenan, K. (2003). Parenting Clinical Observation Schedule (P-COS) for the DB-DOS. (Version 1.1). Unpublished manual, Chicago: University of Chicago.

Wakschlag, L., & Keenan, K. (2001). Clinical significance and correlates of disruptive behavior symptoms in environmentally at-risk preschoolers. *Journal of Clinical Child Psychology*, *30,* 262–275.

Wakschlag, L., Leventhal, B., Danis, B., Keenan, K., Hill, C., Egger, H., & Carter, A. (2002). *The Disruptive Behavior Diagnostic Observation Schedule (DB-DOS).* (Version 1.2). Unpublished manual, Chicago: University of Chicago.

Wakschlag, L., Leventhal, B., Hill, C., Danis, B., & Keenan, K. (2004, October). *Observed disruptive behavior in the parent-child context and impairment*. Paper presented at the Meetings of the American Academy of Child and Adolescent Psychiatry, Washington, DC.

Webster-Stratton, C., & Lindsay, D. (1999). Social competence and conduct problems in young children: Issues in assessment. *Journal of Clinical Child Psychology*, *28,* 25–43.

Wenar, C. (1982). On negativism. *Human Development*, *25,* 1–23.

Westen, D., & Weinberger, J. (2004). When clinical description becomes statistical prediction. *American Psychologist*, *59,* 595–614.

Zeanah, C., Boris, N., & Scheeringa, M. (1997). Psychopathology in infancy. *Journal of Child Psychology, Psychiatry and Allied Disciplines*, *38,* 81–99.

Zelenko, M. (2004). Observation in infant-toddler mental health assessment. In R. DelCarmen-Wiggins & A. Carter (Eds.), *Handbook of Infant, Toddler and Preschool Mental Health Assessment* (pp. 205–222). New York: Oxford University Press.