



Published in final edited form as:

J Abnorm Child Psychol. 2014 May ; 42(4): 623–634. doi:10.1007/s10802-013-9802-4.

Reliable Ratings or Reading Tea Leaves: Can Parent, Teacher, and Clinician Behavioral Ratings of Preschoolers Predict ADHD at Age Six?

Sarah O'Neill¹, Robyn L. Schneiderman², Khushmand Rajendran¹, David J. Marks^{1,3}, and Jeffrey M. Halperin¹

¹Department of Psychology, Queens College of the City University of New York, USA

²Ferkauf Graduate School of Psychology, Yeshiva University, New York, USA

³Department of Child and Adolescent Psychiatry, NYU Langone Medical Center, New York, USA

Abstract

Objective—To assess the relative ability of parent, teacher, and clinician behavioral ratings of preschoolers to predict ADHD severity and diagnosis at 6 years of age.

Method—Hyperactive/inattentive preschoolers [$N=104$, 75% boys, Mean (SD) age = 4.37 (.47) years] were followed over two years (mean=26.44 months, SD=5.66). At baseline (BL), parents and teachers completed the ADHD-RS-IV and clinicians completed the Behavioral Rating Inventory for Children following a psychological testing session. At age 6, [Mean (SD) age = 6.62 (.35) years], parents were interviewed with the K-SADS-PL; teachers completed the ADHD-RS-IV; and laboratory measures of hyperactivity, impulsivity, and inattention were obtained from children. Hierarchical logistic and linear regression analyses examined which combination of BL ratings best predicted 6-year-old ADHD diagnosis and severity, respectively.

Results—At age 6, 56 (53.8%) children met DSM-IV criteria for a diagnosis of ADHD. BL ratings from parent/teacher/clinician, parent/teacher and parent/clinician combinations significantly predicted children who had an ADHD diagnosis at age 6. Parent and clinician, but not teacher, behavior ratings were significant independent predictors of ADHD diagnosis and severity at 6-years-old. However, only clinician reports of preschoolers' behaviors predicted laboratory measures of over-activity and inattention at follow-up.

Conclusion—Cross-situationality is important for a diagnosis of ADHD during the preschool years. Among parents, teachers and clinicians, positive endorsements from all three informants, parent/teacher or parent/clinician appear to have prognostic value. Clinicians' ratings of preschoolers' inattention, impulsivity and hyperactivity are valid sources of information for predicting ADHD diagnosis and severity over time.

Keywords

ADHD; preschoolers; inter-rater agreement; cross-situationality; longitudinal

Attention-deficit/hyperactivity disorder (ADHD) is among the most frequently diagnosed childhood psychiatric disorders. It is characterized by developmentally inappropriate levels of inattention, hyperactivity, and impulsivity. According to DSM-IV criteria, some

impairing symptoms must be present prior to age 7 years and symptomatology must cause impairment in at least two settings (American Psychiatric Association (APA), 2000).

Given that psychiatric disorders are behaviorally defined, there is clear agreement among researchers and clinicians that reports about children's behavior should be gathered from multiple informants (American Academy of Pediatrics, 2011; APA, 2012; Hunsley & Mash, 2007). However, in general, agreement among raters tends to be low to moderate. Correlations between teachers' and parents' reports are generally between .2 and .5 (Achenbach, McConaughy, & Howell, 1987; Grietens et al., 2004; Mitsis, McKay, Schulz, Newcorn, & Halperin, 2000; Murray et al., 2007; Verhulst & Akkerhuis, 1989). Discrepancies among raters may be accounted for by measurement error, systematic differences related to the informant (e.g., cultural expectations, other children's psychopathology) or situational specificity of children's behavior (Dirks, De Los Reyes, Briggs-Gowan, Cella, & Wakschlag, 2012).

Achenbach and colleagues (1987) concluded that differences among reporters are as informative as similarities. Consistent with this idea, despite what appear to be differences in raters' observations, when multiple informants' reports are integrated, we are able to account for a greater proportion of the variance in children's behavioral outcomes (Grietens et al., 2004). Thus, multiple informants contribute unique information regarding the behavior of any given child and a multi-informant approach is essential for understanding the behavior of children in different contexts (e.g., Achenbach et al., 1987; Murray et al., 2007; Verhulst & Akkerhuis, 1989). This is particularly important for ADHD, which requires cross-situational impairment.

With respect to the assessment of ADHD in *school age* children, practice guidelines recommend that as part of a thorough multi-method, multi-informant assessment, information should be gathered from parents and teachers (American Academy of Pediatrics, 2011; APA, 2012). However, minimal research has been carried out to assess concordance among informants' reports of preschoolers' behavior, or the incremental gain provided by each subsequent rater. Most crucially, we do not know the predictive validity of informants' ratings with respect to continued ADHD diagnosis and later impairment. This issue is especially pertinent to preschoolers, many of whom are judged to exhibit developmentally inappropriate levels of inattention and/or hyperactivity/impulsivity, but *do not* continue to show behavioral dysfunction at school age. In fact, Campbell, Ewing, Breaux, and Szumowski (1986) showed that only about 50% of preschoolers judged to be symptomatic by parent ratings continued to show elevated ADHD symptoms at 6 years of age. Lahey and colleagues integrated parent and teacher reports to identify preschoolers who met DSM-IV criteria for ADHD. The percentage of children who continued to demonstrate inattention and/or hyperactivity/impulsivity at diagnostic threshold at assessment year 7 or 8 ranged from 66.7% (Inattentive Type at preschool) to 81.0% (Combined Type at preschool) (Lahey, Pelham, Loney, Lee & Willcutt, 2005). For early intervention and treatment planning, it would be advantageous to be able to identify preschoolers for whom ADHD will endure versus those whose elevated inattention and/or hyperactivity/impulsivity is more transient. Furthermore, given Campbell et al.'s and Lahey et al.'s findings, it seems particularly important that a second rater supplement parent ratings. Just who the second rater should be for preschoolers is not yet clearly indicated in the literature.

In school age children, clinician ratings do not seem helpful above and beyond parent and teacher reports. Clinicians are generally cautioned against making diagnostic decisions based on observations in their offices during which they fail to see the inattention, hyperactivity, and impulsivity reported by parents and teachers (Barkley, 1998; Sleanor & Ullman, 1981). It is generally considered that, because interactions in the office setting are

one-on-one and often highly structured, children's behavior is likely to be better constrained than at home or school. Consequently, not making a diagnosis based on clinician observations may lead to a higher rate of false negatives (Barkley, 1998; Sleator & Ullman, 1981). In their classic study, Sleator and Ullmann (1981) found that at 3-year follow-up there were no differences in grade point average, medication dose, or teacher ratings of behavior between school-age children whom clinicians judged to be and not be hyperactive. However, children entered this study between kindergarten and grade 6, with results collapsed across age. Given striking developmental differences in children's behavior and their ability to acquiesce to constraints imposed by a structured setting across this age range, closer examination, particularly among young children, is warranted.

Concurrent evidence suggests that the clinical setting could be valid for evaluating preschoolers' behavior. Willcutt, Hartung, Lahey, Loney, and Pelham (1999) had non-clinician examiners complete a simple 7-item rating scale (Hillside Behavior Rating Scale; HBRS) describing 3- to 7-year-olds' behavior immediately following an assessment session. HBRS ratings were significantly associated with parent and teacher ratings of ADHD. Further, HBRS ratings were significantly correlated with impairment (e.g., teacher-rated disruption) after control variables and parent and teacher ratings of ADHD were entered in to the models. This study is informative in that it investigates the utility of non-clinician examiners' ratings of preschoolers' behavior. However, the sample comprised preschoolers with an ADHD diagnosis and matched-control participants without an ADHD diagnosis, with analyses carried out across both groups. This may have inflated correlations obtained among parent, teacher, and non-clinician examiners. Also, predictor and outcomes measures were obtained concurrently. Thus, we still have no knowledge about the predictive validity of these different informants' ratings.

It may be that the clinician's office is a particularly useful context for making assessments about preschoolers' behavior. The psychological testing session is one-on-one and highly structured; thus, children who remain dysregulated under these conditions may have the most negative trajectories. Mischel's classic delay gratification studies may offer a useful analogy. The longer preschoolers were able to wait to receive the preferred object in a simple laboratory task, the more competent they were rated by their parents as adolescents. In this way, the structured testing environment served as a valid setting for assessing children's impulsivity, which in turn was a highly significant predictor of parent-rated social, academic, and emotion regulation outcomes during adolescence (Mischel, Shoda, & Peake, 1988).

Given the limited research on the predictive utility of different informants' behavioral ratings of preschoolers, the present study investigated the differential ability of parent, teacher, and clinician ratings of 3- and 4-year-old children's inattention and hyperactivity/impulsivity to predict ADHD-related outcomes at age 6 years. At follow-up, ADHD severity and diagnosis were obtained through interviews and ratings scales, and laboratory measures of activity level, impulsivity and inattention were measured. It was hypothesized that 1) parent and teacher ratings would be the most useful predictors of ADHD, accounting for substantial variance in 6-year-old outcomes; and 2) clinician ratings would also be a valid predictor of ADHD-related outcomes, accounting for additional variance above and beyond parent and teacher ratings.

Method

Participants

Three- and 4-year-old hyperactive/impulsive and/or inattentive (H/I) children were recruited into a longitudinal study via two pathways - screenings through local preschools and direct

clinical referrals. For preschool screenings, after obtaining consent from the principal, parents were sent a letter explaining the study, an ADHD-RS-IV (DuPaul, Power, Anastopoulos, & Reid, 1998), and a consent form to obtain teacher ratings using the ADHD-RS-IV. Additional children, for whom the same ratings were obtained, entered the study via referral by school personnel (i.e., guidance counselor, school administrator, school psychologist or social worker) or less frequently a health professional (pediatrician) because of the child's behavioral difficulties. No child was recruited from a psychiatric facility. To be accepted, children needed 6 or more symptoms rated as *Often* or *Very Often* on the Hyperactivity/Impulsivity and/or Inattention subscale(s) by *either* parent *or* teacher. Thus, children did not necessarily have ADHD, but all had high levels of ADHD symptoms as rated by a parent and/or teacher.

Children were excluded if they had a Pervasive Developmental Disorder, Post-Traumatic Stress Disorder, a neurological disorder, a Full-Scale IQ of less than 80 as measured by the Wechsler Preschool and Primary Scale of Intelligence, 3rd Edition (WPPSI-III; Wechsler, 2004), if they were taking systemic medication for a chronic medical/psychiatric disorder, including ADHD, if they did not attend a preschool/daycare, or if they were not fluent in English.

Baseline (BL) and 6-year-old follow-up assessments were completed for 104 children [M (SD) age at BL = 4.37 (.47) years], the majority of whom were boys [n (%) = 78 (75)]. Participants were ethnically and racially diverse, reflective of the urban area from which they were recruited: 65 (62.5%) were White, 16 (15.4%) Black, 6 (5.8%) Asian, and 17 (16.3%) were of 'other/mixed' ancestry. Just over a third [n (%) = 39 (37.5)] had at least one parent of Hispanic descent. Socioeconomic status (SES) was measured using the Nakao-Treas Socioeconomic Prestige Index (Nakao & Treas, 1994). The mean (SD) SES of children's families at baseline was 62 (18.29) consistent with a moderate-level SES. Mean (SD) Full-Scale IQ fell within the Average range 103.58 (13.01), with scores ranging from 80 to 138.

The mean (SD) time between BL assessment and follow-up was 26.44 (5.66) months; children's mean (SD) age = 6.62 (.35) years. A further 36 preschoolers were assessed at BL who did not return for their 6-year-old evaluation. Children who did and did not return did not differ in age, race, ethnicity, or on any parent, teacher or clinician rating of behavior (all $p > .20$). However, compared to those who completed the 6-year-old assessment, those lost were of lower SES [mean (SD) = 54.72 (17.14)] and FSIQ [mean (SD) = 97.31 (11.29)]; both $p < .05$. These indices are still well within the middle class and average ranges, respectively; thus, the sample for this study is likely to be representative of the population from which it was drawn.

Procedure

At BL and 6-year-old follow-up, parents and teachers completed the ADHD-RS-IV (DuPaul et al., 1998) and at follow-up parents were interviewed with the K-SADS-PL. At age 6 years children were administered a computerized A-X CPT and wore actigraphs around their waist and non-dominant ankle. Following the BL and follow-up assessments, child evaluators (who were blind to all information from parents and teachers) completed the Behavioral Rating Inventory for Children (BRIC; Gopin, Healey, Castelli, Marks, & Halperin, 2010).

Although all children were medication naïve at baseline, at the 6-year-old evaluation 23 (22.1%) children were receiving either stimulant ($n = 20$) or non-stimulant ($n = 3$) medication for ADHD. Five of the 23 children were prescribed more than one medication to manage their behavioral dysregulation. Parents were asked to withhold stimulant and non-stimulant medications for ADHD on the day of the evaluation, but not other medications.

This study was approved by the university's Institutional Review Board (IRB). Following a full description of the study, all parents signed IRB-approved informed consent forms. At the completion of the evaluation parents were compensated for their time and children received a small prize to thank them for taking part.

Measures

ADHD Behavioral Ratings: Baseline—Parent and teacher reports of children's hyperactivity/impulsivity and inattention were measured using the K-12 version of the ADHD-RS-IV (DuPaul et al., 1998). We used this version with a younger sample because at the time at which participants were recruited, the preschool version of the scale (McGoey, DuPaul, Haley & Shelton, 2007) had not been published. Informants indicate the frequency with which children engage in each behavior on a 4-point Likert scale (0 = Never or Rarely; 3 = Very Often). Item scores were summed to give a possible range from 0 to 54, with higher scores indicating greater symptom severity. The K-12 version has been shown to be reliable and valid for 5–7 year olds by DuPaul et al. (1998). The slightly modified preschool version of the scale has been shown to be reliable and valid among 3- and 4-year-olds by McGoey et al. (2007). In our sample, coefficient alpha for the parent and teacher scales were .95 and .97, respectively.

Ratings of children's behavior during psychological testing were obtained from well-trained graduate students or Ph.D.-level psychologists using the BRIC (Gopin et al., 2010). Raters were trained in the use of the BRIC along with their training in the psychometric assessment. After learning to administer the tests, they observed trained testers and completed the BRIC independently of the already trained examiner. All discrepancies were discussed. This simple 5-item rating scale assesses child behavior and affect during a psychological testing session. Two items measure mood and sociability, which are areas of associated impairment in children with ADHD. The remaining three items, activity level, inattention, and impulsivity, assess core symptoms of ADHD. Gopin et al. (2010) have shown these latter three items load on a single factor, which they labeled the ADHD Symptom Triad, and that it had strong test – retest reliability ($r = .78, p < .001$) and inter-rater reliability (estimated $r = .81$). Evaluators completed the BRIC following an approximately 2.5-hour psychological testing session. Scores (ranging from 1 to 5) from the ADHD Symptom triad were summed to give a possible range from 3 to 15. Higher scores indicate greater severity. In this sample, coefficient alpha for the ADHD Triad was .82. Those conducting the psychological testing were blind to the child's parent and teacher ratings, and did not know that the children were from a group characterized as H/I.

ADHD Severity and Diagnosis at 6-year-old follow-up—The Kiddie Schedule for Affective Disorders and Schizophrenia: Present and Lifetime Version (K-SADS-PL; Kaufman, Birmaher, Brent, Rao, & Ryan, 1996) is a semi-structured psychiatric interview used to assess psychiatric disorders in children according to DSM-IV criteria. Parents of children taking medication for ADHD were asked to evaluate their child's behavior during times when s/he was not taking medication or when its effects had worn off. Interviewers were either well-trained graduate students with a background in psychopathology or Ph.D.-level psychologists. These interviewers were different individuals from those who had carried out psychological testing with the children at BL. Interviewers were blind to children's baseline data.

In addition, as part of the follow-up evaluation, children's teachers were asked to complete the ADHD-RS-IV. To determine the presence of individual symptoms for the K-SADS-PL summary score, a combined “and/or” rule was employed that was similar to that used in the MTA Study (MTA Cooperative Group, 1999). If either the parent indicated that the

symptom was present with associated impairment on the K-SADS-PL *or* the teacher indicated that the symptom occurred “very often” on the ADHD-RS-IV, it was considered present. Further, if the parent indicated that the symptom was present but not sufficiently impairing to get scored at threshold *and* the teacher indicated that the symptom occurred “often,” the symptom was coded as present. In addition to using ADHD diagnosis as an outcome measure, a dimensional measure of ADHD severity was calculated using scores for the 18 individual ADHD items on the K-SADS-PL. Scores were re-coded from their original scale (1= not present; 2 = sub-threshold; 3 = threshold) to a 0 to 2 scale, and then summed to give a dimensionalized ADHD severity score ranging from 0 to 36. Higher scores indicate greater severity.

Activity Level at Follow-Up—Children’s activity level was measured using solid state actigraphs (model # AM7164; ActiGraph, LLC, Fort Walton Beach, FL), worn around the waist and non-dominant ankle for the duration of the 6-year-old evaluation, which was 2.5–3 hours long. Activity level, measured using an actigraph, has been significantly associated with teacher ratings of inattention and over-activity in non-referred children (Reichenbach, Halperin, Sharma, & Newcorn, 1992) and with clinician ratings of hyperactivity, impulsivity, and inattention in both typically developing preschoolers and those at risk for ADHD (Gopin et al., 2010).

Impulsivity and Inattention at 6-year-Old Follow-Up—Children were administered an A-X Continuous Performance Test (A-X CPT). In this task, based on Halperin, Sharma, Greenblatt, and Schwartz (1991), children responded by pressing a mouse button only when the letter ‘A’ was immediately followed by the letter ‘X’. Each letter was presented for 200 ms with an interstimulus interval of 1.5 s. The entire task consisted of 400 stimuli and lasted approximately 12 minutes. Separate Impulsivity and Inattention scores were generated based on the errors that children made during the task. The Impulsivity score was calculated by summing the short latency ‘A-not-X’ errors and long-latency ‘A-only’ errors. The Inattention score was calculated by summing the number of misses, very late correct responses, and long-latency ‘X-only’ false alarms across the four blocks (for more details, see Halperin et al., 1991).

Data Analysis

We conducted a series of chi square analyses to determine which combination of informants’ BL ratings best predicted an ADHD diagnosis at age 6. Categorical variables for parent-, teacher-, and clinician-rated ADHD symptoms were created by dichotomizing them into scores below or 90th percentile. Based on McGoey et al.’s (2007) ADHD-RS-IV norms for preschoolers (collapsed across gender), parent scores 25 and teacher scores 29 were 90th percentile. Although the preschool version provides more developmentally appropriate item descriptors, we chose to use McGoey et al.’s (2007) norms because the K-12 version is not normed for children in our age group. For the BRIC, based on Gopin et al.’s (2010) procedure, scores 11 on the ADHD Symptom Triad were 90th percentile. In defining the categorical outcome variable, ADHD diagnosis, we included children who met DSM-IV-TR criteria for any of the three major subtypes of ADHD (APA, 2000). We first examined the number of children who did and did not go on to receive a diagnosis of ADHD at 6 years of age as a function of the number of informants whose BL ratings were 90th percentile. To this end, participants were grouped as being rated 90th percentile by none or 1, any 2, or all 3 informants. Note that we combined the 0 and 1 informant(s) together because only 9 (8.8%) children were classified < 90th percentile by all three informants and there were no significant differences in the likelihood of an ADHD diagnosis at age 6 between children rated 90th percentile by none versus one informant (Fisher’s exact test $p = .24$). To examine which combination of informants is optimal, participants were grouped

as being 90th percentile by specific combinations of informants (e.g., parent and clinician; teacher and clinician and so on). To correct for multiple testing, the Bonferroni Step-Down correction was used (Holm, 1979).

We carried out hierarchical logistic regression to determine the odds of a preschooler receiving an ADHD diagnosis at age 6 depending on different informants' BL ratings. SES and FSIQ were entered as control variables in Block 1 (enter method). Dichotomized parent- and teacher-rated ADHD-RS-IV scores were entered in Block 2 (enter method) and dichotomized clinician-rated BRIC scores were entered in Block 3.

Despite the categorical nature of the DSM, there is substantial literature supporting the conceptualization of ADHD behaviors occurring along a dimension (e.g., Marcus & Barry, 2011). Given this, we carried out hierarchical linear regression analyses to examine the extent to which individual informants' BL ratings best predicted 6-year-old continuous outcome measures and the incremental gain in outcome variance accounted for by each significant rater. Continuous outcome measures examined include: ADHD severity (from the K-SADS-PL), waist and ankle movement (from actigraphs), and impulsivity and inattention (from the A-X CPT). For all of these analyses BL SES and FSIQ were entered as control variables in Block 1 (enter method). Parent- and teacher-rated ADHD-RS-IV total scores were entered in Block 2 (enter method), and clinician-rated BRIC score was entered in Block 3.

Results

Descriptive Statistics

Descriptive statistics of predictor and outcome variables are presented in Table 1. Skewness and kurtosis absolute values exceeding ± 1.0 were used as indicators of normality (Tabachnick & Fidell, 2001). The continuous variables that did not fall within this threshold – ankle and waist actigraph scores, and Impulsivity score - did so after square root transformation.

Baseline variables showed that 66 (63.5%) and 59 (56.7%) participants fell 90th percentile on the ADHD-RS according to parent and teacher ratings, respectively. Nearly half ($n=47$, 45.2%) fell 90th percentile on the BRIC. For these categorical groups, agreement of teachers with parents ($\kappa = -.16$, $p = .096$, 95% CI = $-.35$, $.02$) and clinicians ($\kappa = .13$, $p = .17$, 95% CI = $-.06$, $.32$) was poor, while agreement between parents and clinicians was fair ($\kappa = .23$, $p = .012$, 95% CI = $.06$, $.40$) (Landis & Koch, 1977).

The majority of children met criteria for ADHD at BL ($n = 79$; 76%) as assessed by the K-SADS-PL. Of these, 7 met criteria for Predominantly Inattentive Type; 37 for Predominantly Hyperactive/Impulsive Type; and 35 for Combined Type. There was high correspondence between those whose ADHD-RS-IV scores fell 90th percentile and those with an ADHD diagnosis. Among those 90th percentile according to parent, teacher and clinician ratings, 86.4%, 91.5% and 89.4% received an ADHD diagnosis at BL, respectively. At BL, no child met criteria for conduct disorder, but 39 (37.5%) met criteria for ODD. Among those meeting criteria for ADHD at BL, a substantial portion also met criteria for ODD ($n = 34$, 43%).

Among children who met criteria for ADHD at BL, 50 (63%) continued to do so at age 6 years. In addition, six children who did not meet criteria for ADHD at BL did so at 6-years-old. Therefore, at follow-up, 56 (53.8%) participants received a diagnosis of ADHD. Eight, 10 and 38 children met criteria for the Inattentive, Hyperactive/Impulsive and Combined types, respectively. Clearly, children who met criteria for ADHD at BL were more likely to

have a diagnosis at follow-up than those who did not meet criteria at BL, $\chi^2(1) = 11.80, p = .001$.

At age 6 years, some children were in kindergarten ($n=7$) while others were in grade school (grade 1 $n=89$; grade 2 $n=8$). As such, it is possible that different teacher expectations affected findings. Four (57.1%) kindergarteners and 52 (53.6%) children in grade school met DSM-IV criteria for ADHD at age 6. Grade did not impact on likelihood of diagnosis at school-age, Fisher's exact test $p=1.00$ and no differences in teacher-rated Hyperactivity/Impulsivity, Inattention, and Total ADHD severity at age 6 years were seen as a function of grade (all $p > .10$).

We also examined whether gender accounted for the change in rates of ADHD diagnosis from BL to age 6. At BL, 20 (76.9%) girls met criteria for one of the three subtypes of ADHD. A similar proportion of boys met criteria for ADHD ($n=59, 75.6\%$). Thus, girls were no more likely than boys to meet criteria for ADHD at BL, $\chi^2(1) = .02, p = .90$. At age 6 years, 18 girls (69.2%) and 38 (48.7%) boys met criteria for ADHD. The difference in rates of diagnosis as a function of gender approached significance, $\chi^2(1) = 3.30, p = .07$. Of the 18 girls who met criteria for ADHD at 6 years of age, 15 (83%) of them had met criteria for ADHD during preschool, McNemar test $p = .73$ (note that some cell counts were <5). For boys, 16 (84%) of those who *did not* meet criteria for ADHD at BL continued not to at age 6. In contrast, of the 59 boys who had an ADHD diagnosis during preschool, 24 (40.7%) no longer did at school-age, McNemar test $p < .0001$. One must be cautious not to overstate these findings given the small number of girls in the study, but they provide some evidence to suggest that girls identified as displaying developmentally inappropriate levels of inattention and/or hyperactivity/impulsivity during preschool show reasonable stability of the behaviors over time. In contrast, a significant minority of boys who met criteria for ADHD during preschool lost their diagnosis by age 6.

Finally, we looked at whether comorbidity with ODD may have affected outcomes. Excluding 5 children who had ODD alone, we carried out a 1-way ANOVA to examine whether there were differences in 6-year-old outcomes as a function of BL diagnostic status (No ADHD/ODD; ADHD only; ADHD + ODD). Children with no diagnosis had less severe KSADS scores at age 6 than children with ADHD alone or children with ADHD+ODD, $F(2, 101) = 11.75, p < .0001, \eta_p^2 = .19$. Age 6 severity ratings were no different between children with ADHD alone and children with comorbid ADHD and ODD. No differences in 6 year-old laboratory-measured outcomes were observed as a function of baseline ODD.

Correlational Analyses

Among the baseline variables, the only significant correlation was between teacher and clinician, $r = .33, p < .01$. Parent ratings were not significantly correlated with those of the teacher ($r = -.05, p > .10$) or clinician ($r = .10, p > .10$). Similarly, at age 6 years, agreement between parent and teacher total ADHD-RS score was quite low, $r = .16, p = .14$.

Several significant correlations were observed, however, between predictor and outcome variables. Both SES ($r_s = -.21-.30, p < .05$) and FSIQ ($r_s = -.27 - -.35, p < .01$) were related to one or more of the outcomes at age 6 and were therefore included as covariates in the regression analyses. Parent ($r = .46, p < .01$) and clinician ($r = .22, p < .05$) ratings at BL were significantly positively associated with ADHD severity at 6 years of age. Parent ($r = .35, p < .01$) and teacher ($r = .21, p < .05$) ratings at BL were both significantly positively associated with ADHD diagnosis, while clinician ratings ($r = .19, p < .10$) approached significance. Neither parent nor teacher ratings at BL were significantly related to any of the laboratory measures at age 6 years ($r_s = -.02 - .11, p > .10$). Clinicians' ratings at BL were

significantly positively related to waist movement measured using the actigraph ($r=.29, p<.01$) and the A-X CPT Inattention score ($r=.31, p<.01$).

Chi Square Analyses

A significant linear-by-linear association suggests the percentage of children with an ADHD diagnosis at age 6 years increased as the number of informants whose BL ratings were 90th percentile increased, $\chi^2(1) = 13.67, p = .0002$ (see Figure 1). Cross-situationality of symptoms appears to place preschool children at greater risk for continued difficulties over time.

We next examined which combination of informants is optimal (see Figure 2). A chi-square analysis across the four possible combinations of informants plus those with none or 1 endorsement was significant, $\chi^2(4) = 18.47, p = .001$. To further examine the relations among these groupings, we first looked at whether any combination comprising at least two informants with ratings 90th percentile was more likely to identify children as having ADHD at age 6; no differences were seen, $\chi^2(3) = 5.35, p = .15$ (3 cells had expected count < 5). We then conducted individual pairwise comparisons to see which combination of raters was better able to predict ADHD at age 6 relative to 0 or 1 informants. After adjusting for multiple testing (Holm, 1979), three significant pairwise comparisons emerged. Compared to 0 or 1 informants, children were significantly more likely to be identified as having ADHD when all three informants' BL ratings were 90th percentile, $\chi^2(1)=11.69, p=.004$; when both parent *and* teacher BL ratings were 90th percentile, $\chi^2(1)= 8.62, p=.018$; and when both parent *and* clinician BL ratings were 90th percentile, $\chi^2(1)=5.45, p=.04$. When clinician *and* teacher ratings were 90th percentile, children were no more likely to be identified as having ADHD at age 6 than when 0 or 1 informants' ratings were used, $\chi^2(1)=.05, p=1.00$. Thus parents' ratings seem to be critical to the evaluation of preschoolers, but the second informant could be a clinician in the absence of teacher ratings.

Regression Analyses

Diagnosis at follow-up was regressed on dichotomized parent, teacher, and clinicians' ratings, controlling for SES and FSIQ (see Table 2). Parent and clinician, but not teacher ratings were significant independent predictors of ADHD diagnosis at follow-up. Preschoolers rated 90th percentile in ADHD symptom severity by parents and clinicians were, respectively, 3.6 and 2.8 times as likely to be diagnosed with ADHD at age 6 than children who fell $<90^{\text{th}}$ percentile. The addition of clinician ratings to parent ratings resulted in a significantly better model fit. The same pattern of findings is observed if children are categorized as falling $<$ or 93rd percentile.

Hierarchical linear regression analyses revealed that together, SES, FSIQ, parent and teacher accounted for 24.8% of the variance in ADHD severity at 6 years of age, although only parent report accounted for significant variance. When clinician ratings were included in the model, an additional 3% in outcome variance was explained (see Table 3).

Given the possibility that the association between parent ratings at BL and ADHD severity at age 6 years may have been due to source bias (i.e., parent administered K-SADS-PL at both time points), we investigated whether the same pattern of findings would be observed using laboratory measures of hyperactivity, impulsivity, and inattention (see Table 3). Waist and ankle actigraph scores were regressed on parent, teacher, and clinician ratings, while controlling for SES and FSIQ. None of the informants' ratings at BL was a significant predictor of ankle movement, but clinician ratings were a significant predictor of waist movement, accounting for 6.4% of the variance. Among CPT measures, the Inattention score was predicted by clinicians' BL ratings. Over and above the variables in the model,

clinicians' ratings accounted for 5.0% of the variance. Early ratings did not predict the CPT Impulsivity score.

Finally, post-hoc analyses were carried out to investigate whether the two symptom domains, inattention and hyperactivity/impulsivity, were better predictors of age 6 outcomes than total severity scores. We redid all analyses using parent, teacher and clinician rated inattention as predictors of age 6 Inattention (measured using the K-SADS:PL and AX-CPT), and parent, teacher and clinician rated hyperactivity/impulsivity as predictors of age 6 Hyperactivity and Impulsivity (measured using the K-SADS:PL, AX-CPT, and actigraph). Findings were virtually identical to when total severity score was used as the predictor. The exception was that only parent-rated Inattention at BL predicted Inattention severity at age 6 (based on K-SADS:PL).

Discussion

Cross-situational impairment is a DSM-IV criterion for ADHD (APA, 2000). Yet, lack of agreement among raters about children's behavior in different settings appears to be the rule rather than the exception (Achenbach et al., 1987; Grietens et al., 2004; Murray et al., 2007; Verhulst & Akkerhuis, 1989). Furthermore, for a number of preschoolers with clinically significant ADHD symptoms, their inattention and/or hyperactivity/impulsivity remit through the school years (Campbell et al., 1986; Lahey et al., 2005). The present study investigated the ability of parent, teacher, and clinician ratings of preschoolers' inattention, hyperactivity, and impulsivity to predict ADHD diagnosis and symptom severity at 6 years of age.

Low agreement was observed among raters, both for categorical groups (children whose level of hyperactivity/inattention fell < vs. 90th percentile) and continuous ratings of symptom severity. Low agreement among raters may reflect the restricted range of the sample, which was limited to children who displayed elevated levels of inattention and/or hyperactivity/impulsivity. With greater variability, agreement may have been stronger. However, it is children with developmentally inappropriate levels of inattention and/or hyperactivity/impulsivity who are referred for assessment. Therefore, the nature of the current sample provides ecological validity.

Low agreement was also obtained between informants' ratings and laboratory measures of inattention, hyperactivity, and impulsivity. This is not surprising given the robust literature showing weak-moderate concurrent associations between behavioral rating scales and laboratory and analogue measures of children's ADHD behavior (Conners, 2004; Reichenbach et al., 1992; see Barkley, 1991, for a review). The fact that such weak associations are observed at the same point in time suggests that we would be even less likely to see significant relations among adult report and laboratory measures over the 2-year delay period in our study.

Preschool children identified as showing symptoms 90th percentile by at least two raters were more likely to have an ADHD diagnosis at 6 years than those rated at or above threshold by one or none of the raters. Yet, the combination of raters was important. Positive ratings from the clinician and teacher combination had limited utility. However, positive BL reports from either of these informants along with positive parent BL ratings were highly predictive of later ADHD. Thus, cross-situationality is important and one informant should be the parent. Yet, preschoolers who are only reported to be symptomatic by parents have reduced likelihood of developing ADHD, suggesting that situational variables or rater characteristics may explain the reported behavior. Notably, however, additional endorsement from either the teacher or the clinician substantially improves predictive utility.

The prognostic value of each informant's BL ratings for determining ADHD diagnosis at age 6 was further evaluated using logistic regression analysis, which accounted for all raters, as well as IQ and SES. These findings indicated that parent and clinician ratings were the most useful. Children rated above threshold by either of these informants were around three times more likely to receive an ADHD diagnosis at school age than their peers who were rated below the 90th percentile. Clinician, but not teacher, ratings improved the model fit, providing further evidence that multiple raters of preschoolers provide greater predictive validity for later ADHD.

Parent ratings also emerged as the best predictor of later ADHD severity, perhaps because the amount of time that parents spend with their children enables them to see the breadth and chronicity of their children's difficulties (Duhig, Renk, Epstein, & Phares, 2000). Yet, the strong association between parent report at BL and 6-year-old outcomes based on interview and rating scales may also reflect source bias. That is, parent report formed the basis of K-SADS-PL interview at follow-up, with teacher ratings providing supplemental information. It is perhaps not surprising, therefore, that parent ratings at BL emerged as the best predictor of these outcomes.

Notably, baseline ratings obtained from clinicians, but again not teachers, explained additional variance in ADHD severity at age 6 years. Several factors might account for the unique "added value" of clinician ratings. Clinicians' experience with a large number of children may provide them with their own set of developmental norms (Hartman, Rhee, Willcutt, & Pennington, 2007), enabling them to identify children whose behavior is clearly outside of that expected for their age. Clinicians may also uniquely detect those highly impaired/symptomatic children who are not more restrained within the context of a structured one-on-one setting. Clinicians' observations of preschoolers' behavior in a structured one-on-one setting not only predicted later clinical status, but also predicted later laboratory-measured outcomes (excessive truncal movements and inattention). As such, the pattern of findings support the conclusion that clinicians' behavioral ratings of preschoolers are a valid way of assessing their elevated inattention, hyperactivity and impulsivity that is likely to persist over time.

That clinician ratings were so useful may seem counterintuitive; after all, compared with home and school, the clinician's office is relatively artificial. It may be that children who display elevated inattention, hyperactivity, and impulsivity in the office environment, just like Mischel et al.'s (1988) preschoolers who could not wait for the preferred object in the lab, are those who display the most dysregulated behaviors and therefore are those with the poorest trajectories.

When teachers' BL ratings were combined with parents' BL ratings, children were more likely to be identified as having ADHD at age 6 than when a single informant's ratings were used. However, teacher ratings did not help to explain additional variance in 6-year-old continuous outcomes, nor were teacher ratings at BL a significant predictor of ADHD diagnosis at age 6, above and beyond SES, IQ, the parent and the clinician. The latter finding contrasts with the literature regarding school-age children, in which teacher reports are considered to be superior to parent ratings (Loeber, Green, & Lahey, 1990; Power et al., 1998). It also contrasts somewhat with Willcutt et al.'s (1999) finding that teacher ratings are significantly associated with impairment in preschoolers. However, Willcutt et al. (1999) assessed teacher ratings and children's impairment concurrently - they did not assess teachers' capacity to predict outcomes over time and their sample included many children with no behavioral difficulties.

The relatively weak predictive validity of preschool teacher reports may be due to situational variables. Preschoolers first adjusting to the structured and novel school environment may have displayed dysregulated behavior that was transient rather than chronic. That is, as children settled into the more structured environment of school and learned to meet the behavioral expectations of that setting, a diminution of their ADHD-like behaviors may have occurred. It is also possible that many preschool teachers allow for a wide range of behavioral disturbance prior to reporting difficulties. Factors affecting teachers' perceptions of children's behavior and abilities have been the focus of study by Mashburn and colleagues. More experience, a higher student:teacher ratio, longer school days, and lower teacher self-efficacy have been related to lower ratings of children's competence and higher ratings of children's behavioral problems (Mashburn, Hamre, Downer, & Pianta, 2006). Key differences between preschool and kindergarten teachers may also influence their assessment of children. Mashburn and Henry (2004) found that the relation between preschool teacher ratings and children's observed school readiness skills was lower than that between kindergarten teacher ratings and children's observed skill level. In part, differences in teacher ratings was related to level of education, such that teachers with higher degrees generated ratings more closely related to children's observed skill level. Classroom type was also a factor, with head start preschool teachers tending to overestimate children's skills. Thus, critical differences among type of classroom, education of the teacher, and classroom environment may affect teachers' perceptions of children's behavior, potentially explaining the poor predictive validity of preschool teachers' behavioral ratings. Nevertheless, this variability in types of schools and teacher credentials is reflective of the real world – our sample of preschoolers is not unique in this respect and as such our findings likely reflect the true nature of preschool teacher ratings in the U.S.

The possibility of source bias influencing findings by using parent report at BL and follow-up was discussed above. Shared method variance may have been a factor in one other way too. Parents rated preschoolers' behavior primarily in the home (although they may have some knowledge of their child's behavior in other settings); teachers rated behavior at school; clinicians rated behavior in the laboratory. This confounding of rater and setting is problematic (Costello, Loeber, & Stouthamer-Loeber, 1991; Dirks et al., 2012), but is difficult to overcome given that studies provide little evidence to support the practice of asking a third party about the child's behavior in a setting that is different from the one in which interactions with the child generally occur (Mitsis et al. 2000). Simple, quick, reliable, and valid observational measures may go some way to address these issues (Dirks et al., 2012), such as has been developed by Wakschlag and colleagues for the assessment of disruptive behavior in preschoolers (Wakschlag et al., 2005), or the BRIC (Gopin et al. 2010), which was used in this study.

Another methodological issue that may have impacted finding is the use of the K-12 version of the ADHD-RS-IV (DuPaul et al., 1998) to assess preschoolers' behavior. Subsequent to the baseline evaluation, a preschool version of the ADHD-RS-IV was published (McGoey et al., 2007). It is unclear if results would have been altered in some way if we had been able to use the preschool scale.

Another limitation of this study is that clinicians' ratings were obtained over a 2- to 3-hour assessment. While this type of assessment is often carried out by school psychologists or neuropsychologists, such a lengthy assessment may be atypical in many clinical settings. It is important to determine whether clinicians' ratings would be as useful in predicting later outcomes if obtained during a shorter assessment period.

In summary, parents, teachers, and clinicians showed limited agreement in their ratings of preschoolers' ADHD behavior. Despite their discrepancies, in combination, parent/clinician

reports were better predictors of ADHD severity and diagnosis over time than parent reports alone, and clinicians' ratings were the sole predictor of laboratory measures of hyperactivity and inattention. Clinical impressions of preschool children's behavior during a structured assessment should, therefore, not be dismissed as irrelevant data - at least when the assessment takes place over a 2- to 3-hour period. These observations may be used to augment parent reports and appear to be a valid source of information for predicting long-term outcome.

Acknowledgments

This work was supported by grant #R01 MH68286 from the National Institute of Mental Health to Jeffrey Halperin. The content is solely the responsibility of the authors and does not necessarily represent the official view of the National Institute of Mental Health.

References

- Achenbach TM, McConaughy SH, Howell CT. Child/adolescent behavioral and emotional problems: Implications of cross-informant correlations for situational specificity. *Psychological Bulletin*. 1987; 101:213–232. [PubMed: 3562706]
- American Academy of Pediatrics Subcommittee on Attention-Deficit/Hyperactivity Disorder. ADHD: Clinical practice guideline for the diagnosis, evaluation, and treatment. *Pediatrics*. 2011; 128:1–16. [PubMed: 21646265]
- American Psychiatric Association. Diagnostic and statistical manual of mental disorders. 4th ed., text revision. Washington DC: Author; 2000.
- American Psychiatric Association. DSM-V development proposed revisions: attention deficit hyperactivity disorder. 2012. Retrieved from <http://www.dsm5.org/ProposedRevision/Pages/proposedrevision.aspx?rid=383#>.
- Barkley RA. The ecological validity of laboratory and analogue assessment methods of ADHD symptoms. *Journal of Abnormal Child Psychology*. 1991; 19:149–178. [PubMed: 2056161]
- Barkley, RA. Attention-deficit hyperactivity disorder: A handbook for diagnosis and treatment. 2nd ed.. New York: Guilford Press; 1998.
- Campbell S, Ewing L, Breaux A, Szumowski E. Parent-referred problem three-year olds: Follow-up at school entry. *Journal of Child Psychology and Psychiatry*. 1986; 27:473–488. [PubMed: 3745347]
- Conners, CK. Conners' continuous performance test (CPT II) version 5 for windows: Technical guide and software manual. Toronto, Canada: Multi-Health Systems, Inc; 2004.
- Costello EJ, Loeber R, Stouthamer-Loeber M. Pervasive and situational hyperactivity -confounding effect of informant: a research note. *Journal of Child Psychology and Psychiatry*. 1991; 32:367–376. [PubMed: 2033114]
- Dirks MA, De Los Reyes A, Briggs-Gowan M, Cella D, Wakschlag LS. Annual research review: Embracing not erasing contextual variability in children's behavior – theory and utility in the selection and use of methods and informants in developmental psychopathology. *Journal of Child Psychology and Psychiatry*. 2012; 53:558–574. [PubMed: 22360546]
- Duhig AM, Renk K, Epstein MK, Phares V. Interparental agreement on internalizing, externalizing, and total behavior problems: A meta-analysis. *Clinical Psychology: Science and Practice*. 2000; 7:435–453.
- DuPaul, GJ.; Power, TJ.; Anastopoulos, AD.; Reid, R. ADHD Rating Scale IV: Checklists, norms, and clinical interpretation. New York: Guilford; 1998.
- Gopin C, Healey D, Castelli K, Marks D, Halperin JM. Usefulness of a clinician rating scale in identifying preschool children with ADHD. *Journal of Attention Disorders*. 2010; 13:479–488. [PubMed: 19395646]
- Grietens H, Onghena P, Prinzie P, Gadeyne E, Van Assche V, Ghesquiere P, et al. Comparison of mothers', fathers', and teachers' reports on problem behavior in 5- to 6-year-old children. *Journal of Psychopathology and Behavioral Assessment*. 2004; 26:137–146.

- Halperin JM, Sharma V, Greenblatt E, Schwartz ST. Assessment of the continuous performance test: Reliability and validity in a non-referred sample. *Psychological Assessment*. 1991; 3:603–608.
- Hartman CA, Rhee SH, Willcutt EG, Pennington BF. Modeling rater disagreement for ADHD: are parents or teachers biased? *Journal of Abnormal Child Psychology*. 2007; 35:536–542. [PubMed: 17333362]
- Holm S. A simple sequentially rejective Bonferroni test procedure. *Scandinavian Journal of Statistics*. 1979; 6:65–70.
- Hunsley J, Mash EJ. Evidence-based assessment. *Annual Review of Clinical Psychology*. 2007; 3:29–51.
- Kaufman, J.; Birmaher, B.; Brent, D.; Rao, U.; Ryan, N. The schedule for affective disorders and schizophrenia for school-age children present and lifetime version (version 1.0). Pittsburgh, PA: Dept of Psychiatry, University of Pittsburgh School of Medicine; 1996.
- Lahey BB, Pelham WE, Loney J, Lee SS, Willcutt E. Instability of the DSM-IV subtypes of ADHD from preschool through elementary school. *Archives of General Psychiatry*. 2005; 62:896–902. [PubMed: 16061767]
- Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977; 33:159–174. [PubMed: 843571]
- Loeber R, Green SM, Lahey BB. Mental health professionals' perception of the utility of children, mothers, and teachers as informants of childhood psychopathology. *Journal of Clinical Child Psychology*. 1990; 19:136–143.
- Marcus DK, Barry TD. Does attention-deficit/hyperactivity disorder have a dimensional latent structure? A taxometric analysis. *Journal of Abnormal Psychology*. 2011; 120:427–442. [PubMed: 20973595]
- Mashburn AJ, Hamre BK, Downer JT, Pianta RC. Teacher and classroom characteristics associated with teachers' ratings of prekindergartners' relationships and behaviors. *Journal of Psychoeducational Assessment*. 2006; 24:367–380.
- Mashburn AJ, Henry GT. Assessing school readiness: Validity and bias in preschool and kindergarten teachers' ratings. *Educational Measurement: Issues and Practice*. 2004; 23:16–30.
- McGoey KE, DuPaul GJ, Haley E, Shelton TL. Parent and teacher ratings of Attention-Deficit/Hyperactivity Disorder in preschool: the ADHD Rating Scale-IV Preschool Version. *Journal of Psychopathology and Behavioral Assessment*. 2007; 29:269–276.
- Mischel W, Shoda Y, Peake PK. The nature of adolescent competencies predicted by preschool delay of gratification. *Journal of Personality and Social Psychology*. 1988; 54:687–696. [PubMed: 3367285]
- Mitsis EM, McKay KE, Schulz KP, Newcorn JH, Halperin JM. Parent-teacher concordance for *DSM-IV* Attention-Deficit/Hyperactivity Disorder in a clinic-referred sample. *Journal of the American Academy of Child & Adolescent Psychiatry*. 2000; 39:308–313. [PubMed: 10714050]
- MTA Cooperative Group. A 14-month randomized clinical trial of treatment strategies for attention-deficit/hyperactivity disorder. *Archives of General Psychiatry*. 1999; 56:1073–1086. [PubMed: 10591283]
- Murray DW, Kollins S, Hardy KK, Abikoff HB, Swanson JM, Cunningham C, et al. Parent versus teacher ratings of attention-deficit/hyperactivity disorder symptoms in preschoolers with attention-deficit/hyperactivity disorder treatment study (PATS). *Journal of Child and Adolescent Psychopharmacology*. 2007; 17:605–619. [PubMed: 17979581]
- Nakao K, Treas J. Updating occupational prestige and socioeconomic scores: How the new measures measure up. *Sociological Methodology*. 1994; 24:1–72.
- Power TJ, Doherty BJ, Panichelli-Mindel SJ, Karustis JL, Eiraldi RB, Anastopoulos AD, et al. The predictive validity of parent and teacher reports of ADHD symptoms. *Journal of Psychopathology and Behavioral Assessment*. 1998; 20:57–81.
- Reichenbach LC, Halperin JM, Sharma V, Newcorn JH. Children's motor activity: reliability and relationship to attention and behavior. *Developmental Neuropsychology*. 1992; 8:87–97.
- Sleator EK, Ullmann RK. Can the physician diagnose hyperactivity in the office? *Pediatrics*. 1981; 67:13–17. [PubMed: 7243422]
- Tabachnick, BG.; Fidell, LS. Using multivariate statistics. 4th ed.. Boston: Allyn Bacon; 2001.

- Verhulst FC, Akkerhuis GW. Agreement between parents' and teachers' ratings of behavioral/emotional problems of children aged 4–12. *Journal of Child Psychology and Psychiatry*. 1989; 30:123–136. [PubMed: 2925818]
- Wakschlag LS, Leventhal BL, Briggs-Gowan MJ, Danis B, Keenan K, Hill C, et al. Defining the “disruptive” in preschooler behavior: what diagnostic observation can teach us. *Clinical Child and Family Psychology Review*. 2005; 8:183–201. [PubMed: 16151617]
- Wechsler, D. *The Wechsler Preschool and Primary Scale of Intelligence, 3rd edition, technical and interpretive manual*. San Antonio, TX: The Psychological Corporation; 2004.
- Willcutt EG, Hartung CM, Lahey BB, Loney J, Pelham WE. Utility of behavior ratings by examiners during assessments of preschool children with attention/deficit/hyperactivity disorder. *Journal of Abnormal Child Psychology*. 1999; 27:463–472. [PubMed: 10821628]

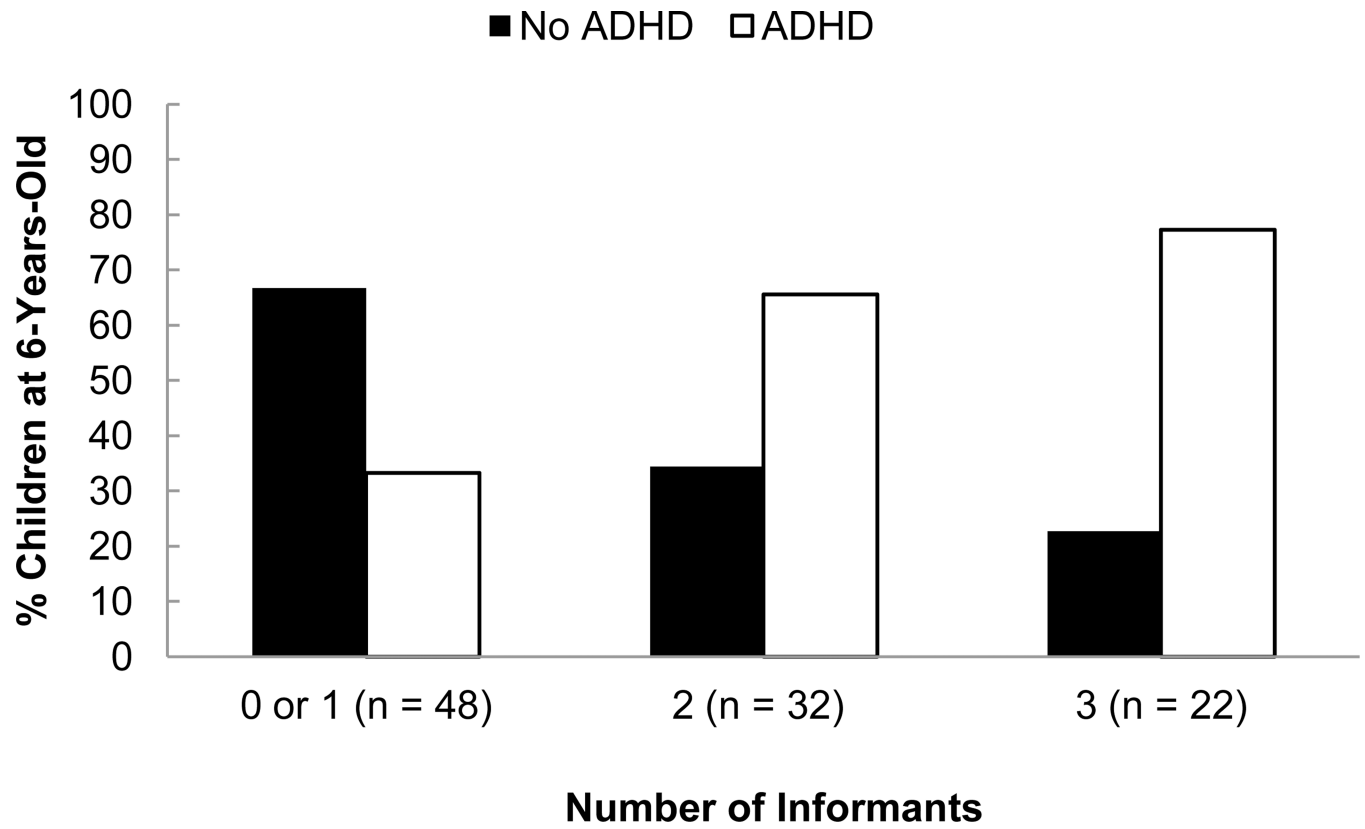


Fig. 1. Percentage of children who do and do not meet DSM-IV criteria for ADHD at Age 6 years as a function of number of informants with BL ratings 90th %ile, $\chi^2(1) = 13.67, p = .0002$

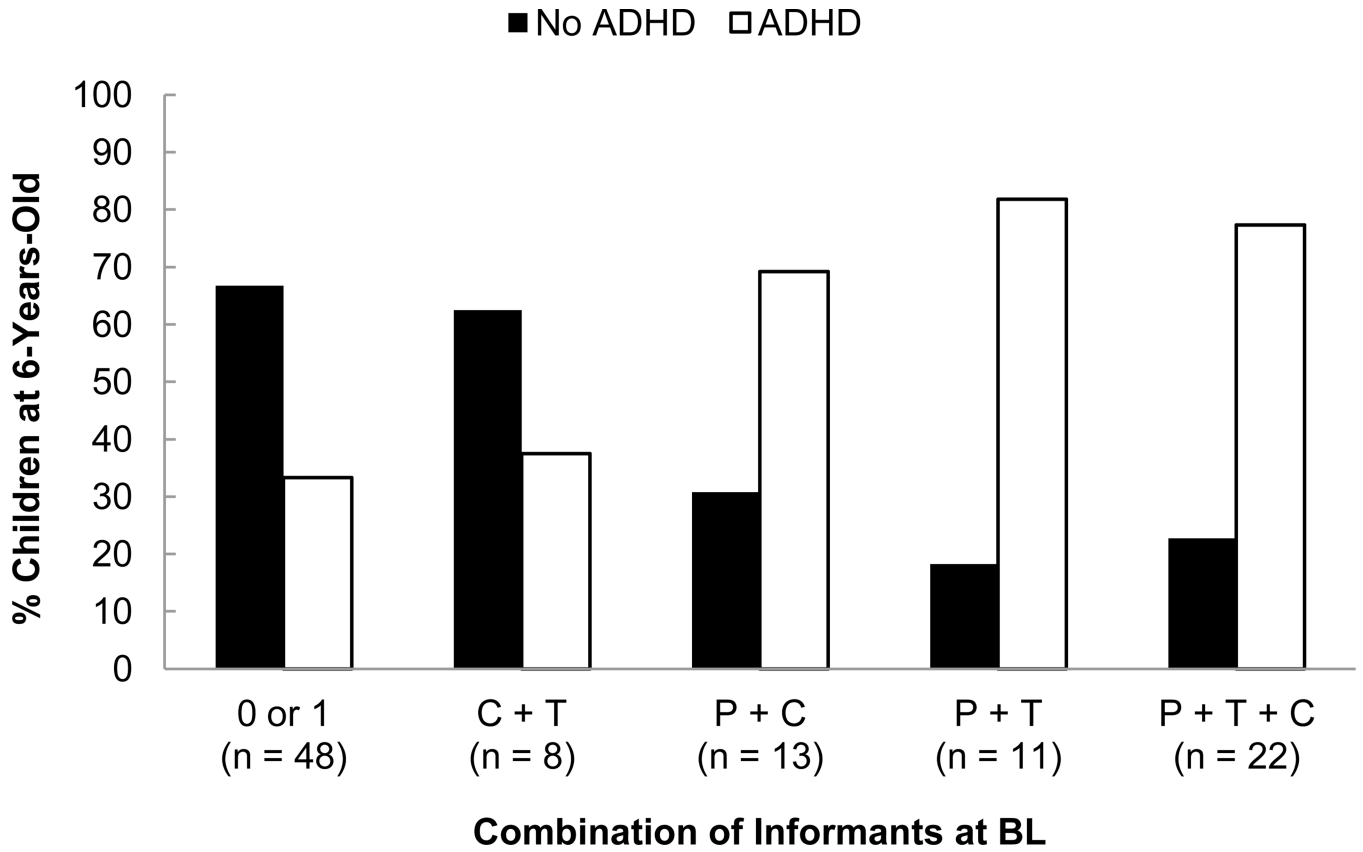


Fig. 2. Percentage of children who do and do not meet DSM-IV criteria for ADHD at age 6 years as a function of combination of informants with BL ratings 90th %ile: P+T+C > 0 or 1, $\chi^2(1) = 11.69, p = .004$; P+T > 0 or 1, $\chi^2(1) = 8.62, p = .018$; P+C > 0 or 1, $\chi^2(1) = 5.45, p = .04$; C+T = 0 or 1, $\chi^2(1) = .05, p = 1.00$ [all *p* values adjusted for multiple testing using Bonferroni Step-Down correction (Holm, 1979)]
 Note: C = clinician rating; P = parent rating; T = teacher rating

Table 1Descriptive Statistics of Predictor and Outcome Variables for Hyperactive/Inattentive Preschoolers (N = 104)^a

Variable	Mean	SD	Range
Parent ADHD-RS Score BL	27.71	10.60	0.00–47.00
Teacher ADHD-RS Score BL	29.50	13.42	0.00–54.00
Clinician BRIC Score BL	9.84	3.24	3.00–15.00
K-SADS-PL Score Age 6	22.30	10.21	0.00–36.00
Ankle Actigraph (sqrt) Age 6	22.46	8.85	3.99–51.95
Waist Actigraph (sqrt) Age 6	15.81	6.64	2.78–34.26
A-X CPT Inattention Age 6	17.28	10.08	1.00–40.00
A-X CPT Impulsivity (sqrt) Age 6	2.48	1.50	0.00–5.74
N (%)			
Parent ADHD-RS 90th percentile BL	66 (63.5)		
Teacher ADHD-RS 90th percentile BL	59 (56.7)		
Clinician BRIC 90th percentile BL	47 (45.2)		
ADHD Diagnosis BL ^b	79 (76.0)		
ADHD Diagnosis Age 6 ^b	56 (53.8)		

^aNs may differ for some variables due to missing data^bMeet DSM-IV criteria for ADHD Predominantly Inattentive, Predominantly Hyperactive/Impulsive or Combined Types.

Table 2

Summary of Hierarchical Logistic Regression Analysis for All Informants' Baseline Ratings Predicting ADHD Diagnosis at 6-Years-Old (N = 102)

Block	Variable	OR	95% CI	p value
1	SES	0.98	0.96–1.00	.11
	FSIQ	1.01	0.98–1.04	.71
2	SES	0.99	0.96 – 1.01	.30
	FSIQ	1.00	0.97–1.03	.95
	Parent ADHD-RS 90 th Percentile	4.74	1.88–11.95	.001
	Teacher ADHD-RS 90 th Percentile	2.15	0.88–5.22	.09
3	SES	0.99	0.96–1.01	.37
	FSIQ	1.01	0.98–1.05	.48
	Parent ADHD-RS 90 th Percentile	3.64	1.39–9.50	.008
	Teacher ADHD-RS 90 th Percentile	1.87	0.75–4.64	.18
	Clinician BRIC 90 th Percentile	2.81	1.07–7.35	.036

OR unadjusted odds ratio, CI Confidence Interval

^aNagelkerke R Square = .03, Model χ^2 (2) = 2.64, $p = .27$

^bNagelkerke R Square = .19, χ^2 (2) = 13.06, $p = .001$; Model χ^2 (4) = 15.69, $p = .003$

^cNagelkerke R Square = .24, χ^2 (1) = 4.52, $p = .034$; Model χ^2 (5) = 20.21, $p = .001$

Table 3

Summary of Hierarchical Linear Regression Analyses

Variable	B	SE B	β	Model Summary		
				Adj R ²	ΔR ²	F change, p
ADHD Severity at Age 6 (n = 102)						
Step 1				.03		
SES BL	-.12	.06	-.21*			
FSIQ BL	.06	.08	.08			
Step 2				.23	.21	13.71, p < .0001
SES BL	-.08	.05	-.14			
FSIQ BL	.05	.07	.07			
Parent ADHD-RS Score BL	.43	.09	.45***			
Teacher ADHD-RS Score BL	.11	.07	.14			
Step 3				.25	.03	4.00, p = .048
SES BL	-.08	.05	-.14			
FSIQ BL	1.00	.07	.12			
Parent ADHD-RS Score BL	.41	.09	.42***			
Teacher ADHD-RS Score BL	.06	.07	.08			
Clinician BRIC Score BL	.63	.32	.20*			
Waist Movement at Age 6 (n = 96)						
Step 1				-.02		
SES BL	-.02	.04	-.06			
FSIQ BL	.03	.05	.06			
Step 2				-.01	.02	1.07, p = .35
SES BL	-.01	.04	-.04			
FSIQ BL	.03	.05	.06			
Parent ADHD-RS Score BL	.06	.07	1.00			
Teacher ADHD-RS Score BL	.06	.05	.13			
Step 3				.04	.06	6.35, p = .01
SES BL	-.01	.04	-.04			

Variable	B	SE B	β	Model Summary		
				Adj R ²	ΔR^2	F change, p
FSIQ BL	.08	.05	.15			
Parent ADHD-RS Score BL	.03	.07	.05			
Teacher ADHD-RS Score BL	.01	.05	.03			
Clinician BRIC Score BL	.59	.24	.29*			
Inattention (A-X CPT) at Age 6 (n = 96)						
Step 1				.08		
SES BL	-.09	.06	-.17			
FSIQ BL	-.18	.08	-.23*			
Step 2				.06	.00	.02, p = .98
SES BL	-.09	.06	-.16			
FSIQ BL	-.17	.08	-.23*			
Parent ADHD-RS Score BL	.01	.10	.01			
Teacher ADHD-RS Score BL	.01	.08	.02			
Step 3				.10	.05	5.27, p = .02
SES BL	-.09	.06	-.17†			
FSIQ BL	-.11	.08	-.14			
Parent ADHD-RS Score BL	-.03	.10	-.03			
Teacher ADHD-RS Score BL	-.05	.08	-.07			
Clinician BRIC Score BL	.84	.37	.26*			

*** p < .01

* p < .05

† p < .10