THE JOURNAL OF
CHILD PSYCHOLOGY AND PSYCHIATRY

Journal of Child Psychology and Psychiatry 49:1 (2008), pp 59–69 doi:10.1111/j.1469-7610.2007.01809.x

Quantifying ADHD classroom inattentiveness, its moderators, and variability: a meta-analytic review

Michael J. Kofler, Mark D. Rapport and R. Matt Alderson

Department of Psychology, University of Central Florida, USA

Background: Most classroom observation studies have documented significant deficiencies in the classroom attention of children with attention-deficit/hyperactivity disorder (ADHD) compared to their typically developing peers. The magnitude of these differences, however, varies considerably and may be influenced by contextual, sampling, diagnostic, and observational differences. **Methods:** Meta-analysis of 23 between-group classroom observation studies using weighted regression, publication bias, goodness of fit, best case, and original metric analyses. **Results:** Across studies, a large effect size (ES = .73) was found prior to consideration of potential moderators. Weighted regression, best case, and original metric estimation indicate that this effect may be an underestimation of the classroom visual attention deficits of children with ADHD. Several methodological factors–classroom environment, sample characteristics, diagnostic procedures, and observational coding schema–differentially affect observed rates of classroom attentive behavior for children with ADHD and typically developing children. After accounting for these factors, children with ADHD were on-task approximately 75% of the time compared to 88% for their classroom peers (ES = 1.40). Children with ADHD were also more variable in their attentive behavior across studies. **Conclusions:** The present study confirmed that children with ADHD exhibit deficient and more variable visual attending to required stimuli in classroom settings and provided an aggregate estimation of the magnitude of these deficits at the group level. It also demonstrated the impact of situational, sampling, diagnostic, and observational variables on observed rates of on-task behavior. **Keywords:** ADHD, classroom observation, attention, on-task, meta-analysis. **Abbreviations:** ADHD, attention-deficit/hyperactivity disorder; ES, effect size; MTA, Multimodal Treatment Study of ADHD.

Attention-deficit/hyperactivity disorder (ADHD) is a chronic and disabling condition that affects an estimated 3–7% of school-age children (APA, 2000). Inattention–one of three hypothesized core deficits of ADHD–has been studied extensively over the past several decades following the seminal work by Douglas and colleagues (Douglas, 1972; Sykes, Douglas, Weiss, & Minde, 1971).

Controlled field investigations are highly consistent in documenting that children with ADHD are more inattentive (operationalized as *off-task* in educational research) relative to typically developing (TD) children. This finding holds true for observations as brief as 10 minutes (Skansgaard & Burns, 1998; Zentall, 1980) and as long as 90 minutes (Klein & Young, 1979), when using simple (DuPaul & Rapport, 1993; Roberts, 1990) and highly complex classroom coding schemas (Abikoff et al., 2002; Atkins, Pelham, & Licht, 1985; Cunningham & Siegel, 1987; Klein & Young, 1979; Skansgaard & Burns, 1998), and for younger (Schachar, Sandberg, & Rutter, 1986; Zentall, 1980) and older elementary school children (Jacob, O'Leary, & Rosenblad, 1978). Differences in rates of directly observed attentive behavior between children with ADHD and TD controls vary between 4.29% (Campbell, Schleifer, & Weiss, 1978) and 26.6% (Roberts, 1990) based on published reports, and may reflect subject and methodological variables that moderate observations of attention.

Identifying potential moderating variables and quantifying their influence on children's attention may yield findings of both applied and heuristic value. For example, nearly all teacher rating scales used by clinicians conducting a comprehensive diagnostic evaluation of ADHD contain an attention factor (Rapport, Kofler, Alderson, & Raiker, 2007). Scale scores reflect teacher endorsements over widely disparate time intervals, ranging from the past week (Conners, Sitarenios, Parker, & Epstein, 1998) to the preceding six months (Achenbach & Rescorla, 2001), contain no standardized mechanism for recognizing the myriad factors that influence observations of children's attention over these time intervals, and are subject to potential errors associated with retrospective recall, halo effects, and rater expectation bias (Harris & Lahey, 1982; Kent, O'Leary, Diament, & Dietz, 1974). Moreover, scale scores have limited interpretive value. The degree of inattentiveness is inferred based on standard deviation units derived from a standardization sample, but has little relevance for understanding how inattentive children are in their current setting other than to relate back to the scale's measurement

Conflict of interest statement: No conflicts declared.

Published by Blackwell Publishing, 9600 Garsington Road, Oxford OX4 2DQ, UK and 350 Main Street, Malden, MA 02148, USA

metric (e.g., *often inattentive*). In contrast, direct observations of attention derived from extant literature traditionally refer to preferential selection and processing of sensory information, and are operationalized in classroom settings as visual orientation to a required stimulus (i.e., on-task behavior). Its converse, inattention or off-task behavior is inferred by frequent shifts in activity and behavior that is not task-related (Bear, Connors, & Paradiso, 2007; Platzman et al., 1992). Operational definitions employed in both cases generate a quantitative metric that can be reliably estimated and compared across children and settings, and examined to determine the extent to which other factors influence it.

Child-oriented mental health professionals have long recognized the value of obtaining direct observations of children in classroom settings. Classroom observations of student behavior are the most frequently used assessment method by school psychologists (Wilson & Reschly, 1996), and considered a necessary component when conducting a functional analysis of inattention, aggression, and other forms of externalizing behavior problems (Volpe, DiPerna, Hintze, & Shapiro, 2005). Their value is also recognized by the ICD-10, which requires a direct observation of on-task behavior during academic or other tasks to confirm a diagnosis of hyperkinesis/ADHD due to their recognition of the insufficient positive and negative predictive power of available rating scales (World Health Organization, 1993). Several standardized observation coding schemes are currently available for these purposes (for a review, see Volpe et al., 2005); however, none contain information regarding the various factors that potentially influence classroom observations (e.g., impact of length and duration of observations on observed behavior rates).

Quantifying the degree to which attention-deficits adversely influence classroom attention, and factors that moderate these findings, is particularly relevant given that classroom inattention is often the catalyst for clinical referrals (APA, 2000; Pelham, Fabiano, & Massetti, 2005), correlates highly with multiple pejorative outcomes, and portends near, immediate, and long-term functioning both within and external to the educational setting. Inattentive children complete less schoolwork correctly (Rapport, Denney, DuPaul, & Gardner, 1994), and are more likely to receive lower grades and standardized test scores (Zentall, 1993), be placed in special education classrooms, have comorbid learning disabilities, and repeat a grade compared to their TD peers (Faraone et al., 1993). More than half of all children with ADHD fail at least one grade by adolescence (Barkley, DuPaul, & McMurray, 1990), and 23% to 32% of children with ADHD fail to finish high school (Barkley, Fischer, Smallish, & Fletcher, 2006; Mannuzza, Klein, Bessler, Malloy, & LaPadula, 1993). Classroom attentional problems identified in young children predict scholastic underachievement during adolescence (Rapport, Scanlan, & Denney, 1999) and early adulthood (Fergusson & Horwood, 1995), and their negative impact continues into adulthood. For example, longitudinal studies of children with attention problems reveal that they are less likely to attend college, and more likely to have conduct problems, unstable marriages, and lower mean SES scores compared to peers of similar intelligence (Barkley et al., 2006; Mannuzza et al., 1993).

The need for a meta-analytic approach

Two previous review articles found that most, but not all, direct observation studies report significant differences between children with ADHD and TD peers (Luk, 1985; Platzman et al., 1992). Both reviews noted significant variability among behavioral coding schemes, participant age, and study setting. The reviews, however, were primarily descriptive in nature and failed to quantify between-study differences or analyze the potential effects of moderator variables on observed differences between children with ADHD and TD controls. Box score and descriptive reviews such as these do not consider study power, and results may therefore inaccurately reflect the data (see Howard, Maxwell, & Fleming, 2000, for details and specific examples of this phenomenon).

Although it is generally accepted that children with ADHD are off-task more often than their unaffected peers, the findings from studies directly observing visual attention in the classroom are not unequivocal. Across studies, the magnitude of difference between ADHD and non-ADHD children ranges from zero, indicating no differences, to more than two standard deviations. It may be that problems with attention are highly dependent on situational and contextual factors, rather than a ubiquitous deficit (Porrino et al., 1983). A meta-analysis not only tests empirically the existence of this phenomenon across studies, but provides an aggregate quantification of the magnitude of visual attention difficulties compared to typically developing children. It also provides increased power to detect moderator effects and allows estimation of the goodness of fit of moderator regression models (Hedges & Pigott, 2004).

The present meta-analysis of published and unpublished studies also examines variability of on-task differences across studies. Despite the potentially ubiquitous nature of higher intra- and inter-subject variability in ADHD, past studies and reviews have focused exclusively on estimating the magnitude of differences between children with ADHD and controls. Examining variability and factors that influence it–in addition to off-task rates–may provide a more comprehensive understanding of the nature of inattentive classroom behavior associated with ADHD (Castellanos et al., 2005). The moderating relationship of sample characteristics, diagnostic methods, classroom variables, and observational schema on the magnitude of observed classroom attentional differences are analyzed to determine whether the variability between effect size estimates exceeds levels expected based on study-level sampling error. For example, the ratio of males to females across studies was analyzed due to extant research indicating gender differences in ADHD prevalence rates and symptom manifestation, which may affect the magnitude of observed attentional deficits (Abikoff et al., 2002; Biederman & Faraone, 2004; Yang, Jong, & Chung, 2004). Previous empirical studies support the influence of methodological variables; however, the relative impact of these variables for understanding classroom on-task differences in children with ADHD and TD children remains unknown (Luk, 1985). These factors warrant scrutiny because of their potential to change the nature of dependent-independent variable relationships, with implications for theory development, refinement, and refutation (Holmbeck, 1997).

Method

Literature searches

A three-tier literature search was conducted using PsycInfo, PsycArticles, ERIC, Dissertation Abstracts International, and Social Science Citation Index. Search terms included permutations of the ADHD diagnostic label, class*, observ*, behav*, school, direct, attention, and on/off-task, where asterisks serve as wildcards (e.g., observ* will return studies with the word observation, observations, observer, etc.). Searches were conducted with and without an ADHD search term included. Searches were conducted independently by two researchers (MJK and RMA), and repeated until no new studies were located. To further expand the initial study base, table of contents searches of the following journals likely to publish classroom observation studies of children with ADHD were undertaken: *Journal of Applied Behavior Analysis*, *Behavior Modification*, *Behavior Research and Therapy*, *Behavior Therapy*, *Journal of School Psychology*, *School Psychology Review*, and *Journal of Attention Disorders*. After the initial searches, studies cited by articles observing ADHD children were examined (Tier II backward search), and a forward search (Tier III) was conducted using the Social Science Citation Index to locate studies citing those that observed children with ADHD. These procedures generated 510 peer-reviewed studies, dissertations, and unpublished manuscripts written since 1962.

Inclusion and exclusion criteria

Inclusion and exclusion criteria are described below, with the number of studies omitted for each criterion in parentheses. The following served as inclusion criteria for the review: (a) an independent direct observation of children exhibiting inattentive, hyperactive, and/or impulsive behavior in an elementary classroom setting (204); (b) between 6 and 12 years of age (24); (c) on- or off-task frequency or duration data reported, or statistics reported in between-group studies from which effect size can be estimated (94); (d) a typically developing comparison group (53); and (e) low average or higher estimated intelligence (16). Exclusion criteria included: (a) comorbidity with other mental health disorders beyond learning disabilities (LD), oppositional defiant disorder (ODD), or conduct disorder (CD) (11); (b) repeat data (e.g., study published in journal and as book chapter; follow-up longitudinal study (4); and (c) stimulant or psychotropic medication taken during observation, or no pretreatment baseline condition (74). Studies reporting only placebo (i.e., no medication-free baseline) conditions were excluded based on research demonstrating significant differences in functioning between baseline and placebo conditions in children with ADHD (e.g., Rapport et al., 1994). Twenty-three studies published from 1969 to 2006 met search criteria. An additional six studies met the above criteria but were excluded because they did not report the sample size for their comparison group.

Coding of moderators

All potential moderator variables were coded according to the characteristics reported in Tables 1 and 2. Categorical variables were coded chronologically, where higher values are associated with an addition to the variable in question (e.g., adding matched controls, diagnostic tools, observation time). Observation method was coded dichotomously as (0) alternating or (1) continuous. The proportion of male subjects was dichotomized based on whether study samples contained 10% or more females, based on a previous meta-analytic review of predominately male samples (Mezulis, Abramson, Hyde, & Hankin, 2004). Observation duration, intervals, and days were each divided into three groups of approximately equal cell size. ADD-H and ADHD groups were combined into a single code for diagnostic moniker due to research suggesting that children diagnosed as ADD-H based on the DSM-III typically meet DSM-IV ADHD Combined Type criteria (APA, 1987/2000; August & Garfinkel, 1993; Garfinkel & Amrami, 1992). The specific classroom activity (e.g., structured vs. unstructured academic tasks; math vs. language arts) was reported in only three studies, precluding inclusion as a moderator. Definition of off-task behavior refers to the minimum duration a child must be off-task to be coded as such. This category was coded as (0) partial interval, where time equal to less than an observational interval (as defined above) must pass before a child is coded on-/off-task (e.g., DuPaul & Rapport, 1993, define 15-s coding intervals, during which a child must be off-task for two or more consecutive seconds to be coded off-task for that interval); (1) whole interval, where the subject must be off-task the entire interval; or (2) per incident, for studies coding each behavior change as it occurs. Additional details are available from the authors.

Computation of effect sizes

Multiple effect sizes. Four studies reported data sufficient to calculate multiple effect sizes. Only one effect

Table 1 Demographic, setting, and diagnostic variables

Study	Total *N*	Age mean	Male/Female (% male)	Match ctrl S	A	G	C	Classroom type	Diagnostic method	Diagnostic moniker
Werry & Quay (1969)	21	8.92	21/0 (100)		Y			Simulated	*NR*	Bx Prob
Forness & Esveldt (1975)	48	7.2	48/0 (100)	Y		Y	Y	Regular	PTR	Bx Prob
Shecket & Shecket (1976)	36	NR	*NR*					Regular	PTR	Hyperactive
Abikoff et al. (1977)	120	8.17	112/8 (93.3)	Y	Y			Regular	RS-M	Hyperactive
Campbell et al. (1978)	31	7.67	26/5 (83.9)	Y	Y			Regular	PTR	Hyperactive
Jacob et al. (1978)	16	9.6	14/2 (87.5)	Y				Simulated	RS-T	Hyperactive
Klein & Young (1979)	34	NR	34/0 (100)					Regular	RS-T	Hyperactive
Abikoff et al. (1980)	119	8.4	114/5 (95.8)	Y		Y		Regular	RS-M	Hyperactive
Zentall (1980)	62	7.3	62/0 (100)	Y	Y	Y	Y	Regular	RS-T	Hyperactive
Abikoff & Gittelman (1984)	56	8.5	54/2 (96.4)	Y	Y			Regular	RS-M	ADDH
Abikoff & Gittelman (1985)	56	8.17	54/2 (96.4)	Y		Y		Regular	RS-M	ADDH
Atkins et al. (1985)	47	9.14	34/13 (72.3)	Y		Y	Y	Regular	RS-T	ADD
Book & Skeen (1987)	162	*NR*	*NR*			Y	Y	Regular	PTR	Bx Prob
Cunningham & Siegel (1987)	60	8.71	60/0 (100)	Y	Y			Simulated	RS-P	ADDH
Roberts (1990)	33	8.92	33/0 (100)					Simulated	SSI/RS-P	Hyperactive
DuPaul & Rapport (1993)	56	8.32	46/10 (82.1)					Regular	SSI/RS-M	ADD
Lett & Kamphaus (1997)	55	7.94	41/14 (74.5)					Regular	P&T RS	ADHD
Nolan & Gadow (1997)	68	8.8	62/6 (91.2)					Regular	SSI	ADHD
DuPaul et al. (1998)	26	7.7	21/5 (80.8)	Y		Y	Y	Regular	SSI/RS-M	ADHD
Skansgaard & Burns (1998)	24	*NR*	8/16 (33.3)	Y		Y	Y	Regular	RS-T	ADHD
Solanto et al. (2001)	112	8.45	96/16 (85.7)					Regular	SSI/RS-M	ADHD
Abikoff et al. (2002)	1004	8.4	806/198 (80.3)	Y		Y		Regular	SSI/RS-M	ADHD
Lauth et al. (2006)	106	8.47	76/30 (71.7)	Y	Y	Y	Y	Regular	RS-T	ADHD

Notes: A = Age; Bx Prob = Referred for behavioral problems/hyperactivity; C = Classroom; G = Grade; M = Multiple informants; *NR* = Not reported; P = Parent rating scale only; PTR = Pediatrician or teacher referral only; RS = Rating scale(s); S = Sex; SSI = Structured or semi-structured interview; T = Teacher rating scale only.

Table 2 Observational variables

Study	Obs/ day	Number of codes	Number of observers	Definition of off-task	Observation duration (min)	Observation interval (sec)	Observation method
Werry & Quay (1969)	3	17	6	Partial > 5s	15	20	Continuous
Forness & Esveldt (1975)	6	5	6	*NR*	10	6	Alternating
Shecket & Shecket (1976)	4	19	1	*NR*	20	10	Alternating
Abikoff et al. (1977)	3	13	5	Whole > 15s	16	15	Alternating
Campbell et al. (1978)	2	5	*NR*	*NR*	15	10	*NR*
Jacob et al. (1978)	10	6	*NR*	Partial	30	10	Alternating
Klein & Young (1979)	2	17	2	Whole	90	10	Alternating
Abikoff et al. (1980)	3	12	4	Whole > 15s	16	15	Alternating
Zentall (1980)	5	6	3	Partial > 10s	10	Not divided	Alternating
Abikoff & Gittelman (1984)	3	12	7	Whole > 15s	16	15	Alternating
Abikoff & Gittelman (1985)	3	12	7	Whole > 15s	16	15	Alternating
Atkins et al. (1985)	7	32	2	Partial	*NR*	2	Alternating
Book & Skeen (1987)	1	4	*NR*	Incidence	45	Not divided	Continuous
Cunningham & Siegel (1987)	1	16	2	Partial	15	5	Continuous
Roberts (1990)	1	6	*NR*	Incidence	15	Not divided	Continuous
DuPaul & Rapport (1993)	1	1	*NR*	Partial > 2s	20	15	Continuous
Lett & Kamphaus (1997)	1	13	*NR*	Partial	15	3 (27 sec recording)	Continuous
Nolan & Gadow (1997)	3	5	3	Whole	30	10	Alternating
DuPaul et al. (1998)	3	4	2	*NR*	18	15	Continuous
Skansgaard & Burns (1998)	4	96	2	Whole	10	5	Continuous
Solanto et al. (2001)	1	12	*NR*	Whole > 15s	16	15	Alternating
Abikoff et al. (2002)	3	12	*NR*	Whole > 15s	16	15	Alternating
Lauth et al. (2006)	3	8	1	Partial	*NR*	5	Alternating

Notes: Alternating = *abab* or similar method for subsequently observing multiple children; Continuous = Observing one child exclusively for the entire observation period; Number of codes = Number of behaviors/situations coded simultaneously; *NR* = Not reported; Obs/day = Number of observation days; Observation duration = Minutes of observation per day; Observation interval = Duration of observation (in seconds) before coding child as on- or off-task.

size was used for each study to meet the independence assumption (Lipsey & Wilson, 2001).

Publication bias. Shecket and Shecket (1976) did not provide data sufficient to calculate an effect size, but reported no significant between-group differences. The study was retained in the analysis and assigned an effect size of .00 because omitting it was likely to artificially inflate overall effect size estimates due to publication bias (Rosenthal, 1995).

Results

Tier I: Moderator-independent attentional differences

Publication bias: the file drawer problem. Three tests were conducted to determine the likelihood that missing studies would significantly influence the obtained mean effect sizes (Lipsey & Wilson, 2001; Rosenthal, 1991). The *fail-safe N* indicated that an unlikely 1,945 studies finding no between-group differences would be needed to reduce the confidence interval of the mean effect size to include zero (i.e., result in no significant differences in off-task rates). A rank correlation test (Begg & Mazumdar, 1994) for publication bias was non-significant, Kendall's *tau* $b = .099$, $p = .25$. The Trim and Fill procedure suggested that zero studies were missing from the analysis based on expected symmetry when plotting effect sizes by the inverse of their standard errors (Duval & Tweedie, 2000). Collectively, these analyses suggest that the effect of publication bias is minimal or non-existent in the current meta-analysis.

Off-task comparisons. Off-task rates are shown in Table 3. Children with ADHD were off-task a weighted average of 25.66% across studies (range = 5.11% to 47.00%), compared with 12.09% (range = .78% to 34.00%) for control children for the 20 studies reporting this data. All but two studies reported significant between-group differences (Cunningham & Siegel, 1987; Shecket & Shecket, 1976). At the group level, children with ADHD were more variable than control children across studies. The mean standard deviation across studies was significantly greater for ADHD groups than control groups, $t(16) = 4.98$, $p < .0005$.

Effect sizes. Hedges' (1982) *g* effect sizes (ES) were computed using the Comprehensive Meta-Analysis software package (see Table 3). Mean weighted effect size for the 23 studies was .73 (95% CI = .65–.82) with a range of .00 to 2.23. This result corresponds to a large effect (Lipsey & Wilson, 2001). The overall test of homogeneity suggests that there is more variance among the effect sizes than would be predicted by study-level error alone, and supports analysis of potential moderator effects ($Q = 138.23$, $p < .0005$).

Tier II: Moderators of obtained effect size

Data screening. Analysis of moderator variables was conducted on 15 studies reporting data for all variables of interest. An ANOVA with Bonferroni corrections was conducted to determine whether there were systematic differences between studies reporting all data and studies with missing data. Results indicate no significant differences for effect size ($F = .66$, $p = .43$), or any moderator variables

Table 3 Mean off-task rates, standard difference scores, and effect sizes in children with ADHD and typically developing children

Study	ADHD % Off-task M (SD)	Control % Off-task M (SD)	Std. diff. scores (%)	Hedges' *g* effect sizes (std. error)
Werry & Quay (1969)	46.3 (12.8)	23 (15.4)	50.3	2.09 (.53)
Forness & Esveldt (1975)	47.0 (16.5)	34 (12.4)	27.7	.88 (.30)
Shecket & Shecket (1976)	*NR*	*NR*	*NR*	.00[4]
Abikoff et al. (1977)	13.1 (10.0)	2.1 (2.6)	84.2	1.50 (.21)
Campbell et al. (1978)	16.73 (15.15)	12.41 (10.88)	25.8	.32 (.35)
Jacob et al. (1978)	15.8 (NR)	10.5 (NR)	33.3	1.41 (.53)[3]
Klein & Young (1979)	39.8 (9.0)	26.6 (5.0)	33.1	1.78 (.40)
Abikoff et al. (1980)	15.1 (23.4)	4.1 (7.8)	72.8	.62 (.19)
Zentall (1980)	15.0 (NR)	7.1 (NR)	52.2	.45 (.25)
Abikoff & Gittelman (1984)	17.4 (12.3)	3.5 (6.6)	79.7	1.39 (.29)
Abikoff & Gittelman (1985)	15.7 (10.4)	2.5 (4.6)	84.1	1.71 (.31)
Atkins et al. (1985)	*NR*	*NR*	*NR*	.59 (.30)[1]
Book & Skeen (1987)	5.11 (4.82)	.78 (1.47)	84.7	1.21 (.17)
Cunningham & Siegel (1987)	33.0 (NR)	26.4 (NR)	19.9	.51 (.26)[2]
Roberts (1990)	39.5 (18.8)	12.9 (20.9)	67.3	1.31 (.39)
DuPaul & Rapport (1993)	44.26 (16.56)	19.72 (11.56)	55.4	1.66 (.31)
Lett & Kamphaus (1997)	18.3 (16.5)	12.7 (12.7)	30.6	.36 (.29)
Nolan & Gadow (1997)	30.5 (15.9)	13.3 (8.3)	56.4	1.34 (.27)
DuPaul et al. (1998)	33.0 (19.2)	9.5 (11.9)	71.2	1.31 (.45)
Skansgaard & Burns (1998)	23.8 (10.3)	4.8 (6.1)	79.8	2.23 (.60)
Solanto et al. (2001)	*NR*	*NR*	*NR*	.58 (.19)[5]
Abikoff et al. (2002)	10.6 (24.0)	3.3 (13.2)	68.8	.38 (.06)
Lauth et al. (2006)	33.3 (13.0)	12.5 (10.4)	62.5	1.76 (.23)
Column M (SD) =	25.66 (13.06)	12.09 (9.45)	61.04 (21.90)	.73 (.04)[6]

Notes: Effect sizes were calculated using means, SDs and sample sizes unless otherwise noted. NR = Not Reported; [1]Effect size calculated using $N = 47$, $t = 2.01$; [2]Effect size calculated using $N = 60$, $p = .052$; [3]Effect size calculated using $N = 16$, $p = .01$; [4]Effect size set at zero–insufficient data to calculate effect size of nonsignificant differences; [5]Effect size calculated using differences in means (.007), common SD (.12), and $N = 112$; [6]SE of effect sizes; Mean is weighted by study sample size (unweighted ES = 1.10, SE = .13).

except diagnostic method ($F = 9.05$, $p = .007$). Studies with missing moderator data include all four studies diagnosing based solely on a referral for behavior problems, and two of the seven (29%) studies using single informant rating scales. When the behavior problem code was eliminated, no significant differences remained ($F = .351$, $p = .56$). Diagnostic method was therefore retained in the model, with the code for referral deleted.

Weighted regression. A fixed effects weighted regression model was adopted to examine the influence of potential moderator variables on the observed variability in effect sizes. Results reveal that the moderator variables described below predicted a large amount of the variance in effect sizes (adjusted $R^2 = .90$; $Q_R = 91.93$, $p < .001$; $Q_E = .58$, $p = ns$).[1] Q_R is analogous to the F-test for the χ^2 distribution–the significant Q_R indicates that the model predicts significant variability in the effect sizes. The nonsignificant Q_E indicates that only random study-level sampling error remains across effect sizes after accounting for variability explained by the model (Lipsey & Wilson, 2001). The influence of each moderator on obtained effect size is shown in Table 4.

Table 4 Inverse variance weighted regression results

Model	χ^2	Df
Q_R	91.93***	13
Q_E	.58	1
$R^2_{Model} = .997$		
Adjusted $R^2_{Model} = .901$		

Variable	Corrected B–weight	SE_B
Age/Grade/Classroom matching	−2.03*	.83
Percent male	.38	.26
Observation duration (min.)	.69	.64
Observation interval (sec.)	1.21**	.45
Classroom type	−1.69**	.40
Definition off-task	.42	.31
Diagnostic moniker[1]	.95**	.30
Observation method	−2.43**	.41
Number of codes	.86**	.29
Days of observations	1.84**	.42
Diagnostic method	−.79*	.31
(Constant)	.22	1.57

Note: *$p < .05$, **$p < .01$, ***$p < .001$; B/SE_B is tested against the *z*-distribution; *B*-weight = unstandardized regression weight; SE_B = standard error of the regression weight.
[1]Diagnostic moniker and publication year were not entered into the model together because the high correlation between these variables ($r = .83$) violates the assumption of heteroscedasticity. Publication year was therefore omitted from the model. Re-running the weighted regression model substituting publication year for diagnostic moniker did not change the significance of any variable or the direction of any variable's *B*-weight. Publication year was not a significant predictor, *B*-weight = .137, $p > .05$, in the alternate model.

Tier III: Best case estimation and original metric

Best case estimation. The large percentage of variance predicted by the Tier II weighted regression model, and evidence for model stability, allows us to empirically estimate the average magnitude of between-group attentional differences expected under optimal observational conditions–after accounting for the influence of moderators (Lipsey & Wilson, 2001). The best case metric is value neutral–it does not imply whether a lower or higher range of values is desirable–and facilitates conclusions concerning overall differences in classroom attentive behavior between children with ADHD and their peers. Best case estimation involves solving the Tier II fixed effects weighted regression equation using values for each significant moderator corresponding to empirically-supported best practice. Decisions for each type of moderator were based on the following reviews: Direct observations (Harris & Lahey, 1982; Heyman et al., 2001; Volpe et al., 2005), diagnostic assessment (Rapport, Kofler, Alderson, & Raiker, 2007[2]), and study design (Cohen, 2003; Tabachnick & Fidell, 2007). Values corresponding to age/grade or classroom matching, regular classroom settings, longer observation intervals, use of the ADHD moniker based on semi-structured clinical interviews and multiple informants, fewer simultaneous codes, continuous observation schemes, and greater total days of observation were selected based on best practice for each significant predictor in the regression equation (Lipsey & Wilson, 2001). Solving the regression equation using these values and corresponding *B*-weights (Table 4) suggests that an effect size of 1.40 would be expected on average for studies employing this combination of observational and diagnostic methodology, and sample and classroom characteristics.

Original metric. Original metric scores translate the effect size onto the control group distribution, and are calculated by adding the product of the overall weighted mean effect size and the control group standard deviation to the overall weighted mean of the control group (Lipsey & Wilson, 2001).

Across studies, typically developing children were off-task an average of 12.09% (SD = 9.45). An effect size of 1.40, based on the above *best case* estimation, corresponds to a 25.32% off-task average for

[1] Hedges and Pigott (2004) goodness of fit: $GOF = 1-H(c_\alpha \mid k-p-1; \lambda_E) = .57$. This value exceeds the recommended power value of .50 for meta-analyses suggested by Muncer, Craigie, and Holmes (2003) and suggests adequate model specificity.

[2] The recommendations by Rapport and colleagues were chosen *in lieu* of the recent recommendations by Pelham and colleagues (2005) because the latter focused on cost effectiveness rather than comprehensiveness, and did not consider disorder-specific patterns of onset, course, and duration necessary to differentiate ADHD from myriad childhood disorders featuring inattention and/or hyperactivity as secondary symptoms.

children with ADHD. In other words, one would expect typically developing children to be on-task an average of 87.91% of the time, compared to 74.68% for children with ADHD.

Discussion

Direct observations of children with ADHD in classroom settings reveal significant deficiencies in their ability to maintain visual attention relative to peers. The large effect size (ES = .73) obtained prior to considering moderator effects indicates that the visual attention of children with ADHD is sufficiently impaired to be detected across a broad range of sampling and situational variables, diagnostic practices, and observational methods. A best case estimate obtained by solving the significant moderator regression equation using values corresponding to empirically supported best practices suggests that, under ideal conditions, the difference between the ability of children with and without ADHD to pay attention in the classroom is approximately 1.40 standard deviation units (i.e., an effect size of 1.40). Translating this finding into its original metric indicates that children with ADHD are able to focus their attention in classroom settings approximately 75% of the time relative to an 88% on-task average for typically developing children. These values reflect expected group means, and may not reflect the classroom attention of individually observed children due to the high within- and between-day variability associated with the disorder (Abikoff et al., 2002; Castellanos et al., 2005).

Regression analysis revealed that several moderators–sampling and situational variables, diagnostic practices, and observational methods–significantly influenced effect size differences in classroom visual attention to an extent that no systematic variability remained among studies after accounting for their effects (shrunken $R^2 = .90$). Some of these moderators likely impacted directly the behavior of observed children (classroom type), some impacted the characteristics of children observed (diagnostic moniker and method, age/grade/classroom matching), and others impacted overall differences in on-task rates by changing aspects of coding schema (observation interval and methods, days of observations, and number of simultaneous codes). The direct impact of classroom factors on attentive behavior illustrates a potential shortcoming of normed-based, direct observation coding systems of attention (Volpe et al., 2005), and indicates that contextual dynamics need to be considered when interpreting information about classroom behavior (Lauth et al., 2006).

The somewhat counterintuitive finding–that increasingly rigorous diagnostic criteria were associated with smaller effect sizes–may have emerged for at least two reasons. Selecting children based exclusively on teacher or parent rating scale cutoff scores (e.g., 2 SDs above the mean on various factors including inattentiveness) is likely to identify a more highly pathological group relative to procedures that include both formal diagnostic interviewing and rating scales. The former method is likely to include nearly all true positive cases in addition to children with comorbid and other disabling clinical disorders, whereas the latter will exclude a higher rate of false positive cases. The finding may also reflect a potential experimental confound, wherein between-group differences become inflated when group assignment is based on the variable of interest (inattention) rather than the presence of a distinct clinical syndrome.

The number of days observing children and the observation interval required before recording inattentive behavior significantly affected observed group differences, whereas total observation minutes per day did not. This latter finding appears to indicate that the inattentive behavior of children with ADHD in classroom settings is so palpable and frequent that it readily distinguishes them from their peers if observed across days for even brief time intervals. The nonsignificant effect of off-task definition may appear counterintuitive unless the ADHD child's classroom inattentiveness is characterized by both a higher rate and different pattern of off-task behavior that deviates significantly from the norm. A recent study revealed that incidences of off-task behavior in children with ADHD last approximately 1-m, 40-s compared to 20-s for TD, same-age peers (Timko, Kofler, & Rapport, 2004). This extended interval of off-task behavior, if typical of ADHD, would be detected equally well by behavioral definitions requiring differing intervals (e.g., 2-s, 10-s, or 15-s) of continuous off-task behavior.

Intergroup variability

Higher behavioral variability is observed across a wide range of standardized tests, neurocognitive tasks, and experimental paradigms, and may be a ubiquitous feature of ADHD not currently explained by neuropsychological or cognitive endophenotypical models (Castellanos et al., 2005; Doyle et al., 2005). Our meta-analytic review findings were consistent with this literature in revealing significantly higher variability in classroom visual attention in ADHD relative to TD children. The summative nature of the analysis precluded examination of possible inter- and intra-ADHD related variability; however, the phenomenon merits investigation. Direct observation studies of children's visual attention in the classroom may be uniquely positioned to examine potential processes underlying the variability of visual attention deficits in ADHD because of the time series data they can provide. This applies particularly to data derived from continuous direct observations that can be exploited by sophisticated wavelet-based analyses and growth mixture modeling (Castellanos et al., 2005).

Limitations

The restricted age range in reviewed studies precluded examination of differences in classroom visual attention between younger and older elementary age children, but such differences are likely to emerge based on developmental studies of vigilance (Hart, Lahey, Loeber, Applegate, & Frick, 1995) and cognition (Carlson, Lahey, & Neeper, 1986). Inclusion of younger and older subject groupings in future investigations are warranted to better understand whether the visual attentiveness of children with ADHD in classroom settings is developmentally immature, qualitatively different, or reflects elements of both models.

Questions concerning the potential influence of specific academic tasks and structured/unstructured classroom activities on children's classroom visual attention could not be addressed in this review. Accumulating evidence suggests that several variables appear to influence children's visual attention in classroom settings. These include small, seemingly insignificant factors, such as distractors on the border of an assignment (Radosh & Gittelman, 1981), as well as the volume and linguistic content of noise (Zentall & Shaw, 1980), structured nature of the environment (Whalen et al., 1978; Jacob et al., 1978), and type of academic assignment (Rapport, Murphy, & Bailey, 1982). Quantifying the degree to which these variables affect children's visual attentiveness in the classroom, and determining whether on-task observations can serve as a proxy for attention across different activities merit additional scrutiny.

Girls with ADHD were underrepresented in available studies, and may have influenced our failure to find a relationship between gender and observed differences in classroom attention. Gender differences in behavioral and cognitive symptoms are reported in some (Gaub & Carlson, 1997), but not all (Biederman et al., 2005) studies of children with ADHD, with girls tending to exhibit more inattentive and fewer hyperactive symptoms than males (Abikoff et al., 2002; Biederman & Faraone, 2004).

The presence of comorbidity on classroom inattentiveness was not evaluated in the review because none of the reviewed studies included comorbid diagnoses beyond LD, ODD, or CD. Eleven studies were excluded from the review that contained children with comorbid diagnosis; however, 55% of these studies included children with qualitatively different developmental histories (i.e., comorbid mental retardation); the remaining five studies included bipolar disorder (1), adjustment disorder (1), severe emotional disturbance (2), and depression (1). Higher rates of inattentive behavior are likely in comorbid samples based on the elevated inattention ratings obtained in the large MTA study (Abikoff et al., 2002). The fundamental question concerning whether classroom visual inattention is specific to children with ADHD, rather than a non-specific effect of psychiatric diagnosis in general, could not be addressed by the review due to the limited number of studies with a psychiatric control group.

Finally, potential interaction effects among moderator variables were not examined due to power considerations associated with marginal or empty cells. It is statistically unlikely, however, that the addition of interaction terms or other moderator variables would have provided significant incremental benefit, considering the large amount of variance explained by the current model (i.e., adjusted $R^2 = .90$). Statistical analyses indicated adequate power of the current model; however, the relatively low ratio of moderators to included studies tempered our findings and suggests caution when interpreting the results.

In conclusion, our findings confirm that children with ADHD exhibit deficient and more variable visual attending to required stimuli in classroom settings relative to their peers, and provide an aggregate estimation of the magnitude of these deficits at the group level based on a meta-analysis of 23 studies incorporating direct observation methodology. The findings also illustrate the moderating influence of situational, sampling, diagnostic, and observational variables on classroom attention. After accounting for these factors, children with ADHD were on-task approximately 75% of the time compared to an 88% average for typically developing children. The findings have implications for future classroom investigations of ADHD, and the emergent use of norm-based, direct observation coding systems of classroom attention for diagnostic and treatment monitoring purposes. Investigations of provocation and rarefaction factors contributing to classroom inattention are needed to inform the design of in-class interventions for this population.

Correspondence to

Mark D. Rapport, Dept of Psychology, University of Central Florida, 4000 Central Florida Boulevard, Orlando, FL 38217, USA; Email: mrapport@mail.ucf.edu

References

Studies preceded by an asterisk were included in the between-groups meta-analysis

*Abikoff, H., & Gittelman, R. (1984). Does behavior therapy normalize the classroom behavior of hyperactive children? *Archives of General Psychiatry*, *41*, 449–454.

*Abikoff, H., & Gittelman, R. (1985). The normalizing effects of methylphenidate on the classroom behavior of ADDH children. *Journal of Abnormal Child Psychology*, *13*, 33–44.

*Abikoff, H., Gittelman-Klein, R., & Klein, R. (1977). Validation of a classroom observation code for hyperactive children. *Journal of Counseling and Clinical Psychology*, *45*, 772–783.

*Abikoff, H., Gittelman, R., & Klein, R. (1980). Classroom observation code for hyperactive children: A replication of validity. *Journal of Counseling and Clinical Psychology*, *5*, 555–565.

*Abikoff, H.B., Jensen, P.S., Arnold, L.L., Hoza, B., Hechtmen, L., Pollack, S., et al. (2002). Observed classroom behavior of children with ADHD: Relationship to gender and comorbidity. *Journal of Abnormal Child Psychology*, *30*, 349–359.

Achenbach, T.M., & Rescorla, L.A. (2001). *Manual for the ASEBA School-Age Forms & Profiles*. Burlington, VT: University of Vermont Research Center for Children, Youth, & Families.

American Psychiatric Association (1987). *Diagnostic and statistical manual of mental disorders (3rd edn, revised (DSM-III-R))*. Washington, DC: American Psychiatric Association.

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders (4th edn, text revision (DSM-IV-TR))*. Washington, DC: American Psychiatric Association.

*Atkins, M.S., Pelham, W.E., & Licht, M.H. (1985). A comparison of objective measures and teacher ratings of attention deficit disorder. *Journal of Abnormal Child Psychology*, *13*, 155–167.

August, G.J., & Garfinkel, B.D. (1993). The nosology of attention-deficit hyperactivity disorder. *Journal of the American Academy of Child and Adolescent Psychiatry*, *32*, 155–165.

Barkley, R.A, DuPaul, G.J., & McMurray, M.B. (1990). Comprehensive evaluation of attention deficit disorder with and without hyperactivity as defined by research criteria. *Journal of Consulting and Clinical Psychology*, *58*, 775–789.

Barkley, R.A., Fischer, M., Smallish, L., & Fletcher, K. (2006). Young adult outcome of hyperactive children: Adaptive functioning in major life activities. *Journal of the American Academy of Child and Adolescent Psychiatry*, *45*, 192–202.

Bear, M.F., Connors, B.W., & Paradiso, M.A. (2007). *Neuroscience: Exploring the brain (3rd edn)*. Maryland: Lippincott.

Begg, C.B., & Mazumdar, M. (1994). Operating characteristics of a rank correlation test for publication bias. *Biometrics*, *50*, 1088–1101.

Biederman, J., Kwon, A., Aleardi, M., Chouinard, B., Marino, T., Cole, H. et al. (2005). Absence of gender effects on attention deficit hyperactivity disorder: Findings in non-referred subjects. *American Journal of Psychiatry*, *162*, 1083–1089.

Biederman, J., & Faraone, S.V. (2004). The Massachusetts General Hospital studies of gender influences on attention-deficit/hyperactivity disorder in youths and relatives. *Psychiatric Clinics of North America*, *27*, 225–232.

*Book, R.M., & Skeen, J.A. (1987). Behavioral observations of disruptive behaviors with referred and non-referred elementary school children. *Psychology in the Schools*, *24*, 399–405.

*Campbell, S.B., Schleifer, M., & Weiss, G. (1978). Continuities in maternal reports and child behaviors over time in hyperactive and comparison groups. *Journal of Abnormal Child Psychology*, *6*, 33–45.

Carlson, C.L., Lahey, B.B., & Neeper, R. (1986). Direct assessment of the cognitive correlates of attention deficit disorders with and without hyperactivity. *Journal of Psychopathology and Behavioral Assessment*, *8*, 69–86.

Castellanos, F.X., Sonuga-Barke, E.J., Scheres, A., Di Martino, A., Hyde, C., & Walters, J.R. (2005). Varieties of attention-deficit/hyperactivity disorder-related intra-individual variability. *Biological Psychiatry*, *57*, 1416–1423.

Cohen, J. (2003). *Methodological issues & strategies in clinical research (3rd edn)*. Washington, DC: American Psychological Association.

Conners, C.K., Sitarenios, G., Parker, J.D.A., & Epstein, J.N. (1998). Revision and restandardization of the Conners Teacher Rating Scale (CTRS-R): Factor structure, reliability, and criterion validity. *Journal of Abnormal Child Psychology*, *26*, 279–291.

*Cunningham, C.E., & Siegel, L.S. (1987). Peer interactions of normal and attention-deficit-disordered boys during free-play, cooperative task, and simulated classroom situations. *Journal of Abnormal Child Psychology*, *15*, 247–268.

Douglas, V.I. (1972). Stop, look and listen: The problem of sustained attention and impulse control in hyperactive and normal children. *Canadian Journal of Behavioural Science*, *4*, 259–282.

Doyle, A.E., Willcutt, E.G., Seidman, L.J., Biederman, J., Chouinard, V., Silva, J., & Faraone, S.V. (2005). Attention-deficit/hyperactivity disorder endophenotypes. *Biological Psychiatry*, *57*, 1324–1335.

*DuPaul, G.J., Ervin, R.A., Hook, C.L., & McGoey, K.E. (1998). Peer tutoring for children with attention deficit hyperactivity disorder: Effects on classroom behavior and academic performance. *Journal of Applied Behavior Analysis*, *31*, 579–595.

*DuPaul, G.J., & Rapport, M.D. (1993). Does methylphenidate normalize the classroom performance of children with attention deficit disorder? *Journal of the American Academy of Child and Adolescent Psychiatry*, *32*, 190–198.

Duval, S., & Tweedie, R. (2000). Trim and Fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics*, *56*, 455–463.

Faraone, S.V., Biederman, J., Lehman, B.K., Spencer, T., Norman, D., Seidman, L.J. et al. (1993). Intellectual performance and school failure in children with attention deficit hyperactivity disorder and in their siblings. *Journal of Abnormal Psychology*, *102*, 616–623.

*Forness, S.R., & Esveldt, K.C. (1975). Classroom observation of children with learning and behavioral problems. *Journal of Learning Disabilities*, *8*, 382–385.

Fergusson, D.M., & Horwood, L.J. (1995). Early disruptive behavior, IQ, and later school achievement and delinquent behavior. *Journal of Abnormal Child Psychology*, *23*, 183–199.

Garfinkel, B.D., & Amrami, K.K. (1992). A perspective on the attention-deficit disorders. *Hospital and Community Psychiatry*, *43*, 445–448.

Gaub, M., & Carlson, C.L (1997). Gender differences in ADHD: A meta-analysis and critical review. *Journal of the American Academy of Child and Adolescent Psychiatry*, *36*, 1036–1045.

Harris, F.C., & Lahey, B.B. (1982). Recording system bias in direct observational methodology: A review and critical analysis of factors causing inaccurate coding behavior. *Clinical Psychology Review*, *2*, 539–556.

Hart, E.L., Lahey, B.B., Loeber, R., Applegate, B., & Frick, P.J. (1995). Developmental change in attention-deficit hyperactivity disorder in boys: A four-year longitudinal study. *Journal of Abnormal Child Psychology*, *23*, 729–749.

Hedges, L. (1982). Estimation of effect sizes from a series of independent experiments. *Psychological Bulletin*, *92*, 490–499.

Hedges, L.V., & Pigott, T.D. (2004). The power of statistical tests for moderators in meta-analysis. *Psychological Methods*, *9*, 426–445.

Heyman, R.E., Chaudhry, B.R., Treboux, D., Crowell, J., Lord, C., Vivian, D. et al. (2001). How much observational data is enough? An empirical test using marital interaction coding. *Behavior Therapy*, *32*, 107–122.

Holmbeck, G.N. (1997). Toward terminological, conceptual, and statistical clarity in the study of mediators and moderators: Examples from the child-clinical and pediatric literatures. *Journal of Consulting and Clinical Psychology*, *65*, 599–610.

Howard, G.S., Maxwell, S.E., & Fleming, K.J. (2000). The proof of the pudding: An illustration of the relative strengths of null hypothesis, meta-analysis, and Bayesian analysis. *Psychological Methods*, *5*, 315–332.

*Jacob, R.G., O'Leary, K.D., & Rosenblad, C. (1978). Formal and informal classroom settings: Effects on hyperactivity. *Journal of Abnormal Child Psychology*, *6*, 47–59.

Kent, R.N., O'Leary, K.D., Diament, C., & Dietz, A. (1974). Expectation biases in observational evaluation of therapeutic change. *Journal of Consulting and Clinical Psychology*, *42*, 774–780.

*Klein, A.R., & Young, R.D. (1979). Hyperactive boys in their classroom: Assessment of teacher and peer perceptions, interactions, and classroom behaviors. *Journal of Abnormal Child Psychology*, *7*, 425–442.

*Lauth, G.W., Heubeck, B.G., & Mackowiak, K. (2006). Observation of children with attention-deficit hyperactivity (ADHD) problems in three natural classroom contexts. *British Journal of Educational Psychology*, *76*, 385–404.

Lett, N.J., & Kamphaus, R.W. (1997). Differential validity of the BASC student observation system and the BASC teacher rating scale. *Canadian Journal of School Psychology*, *13*, 1–14.

Lipsey, M.W., & Wilson, D.B. (2001). *Practical meta-analysis*. California: Sage.

Luk, S. (1985). Direct observation studies of hyperactive behavior. *Journal of the American Academy of Child Psychiatry*, *24*, 338–344.

Mannuzza, S., Klein, R.G., Bessler, A., Malloy, P., & LaPadula, M. (1993). Adult outcome of hyperactive boys: Educational achievement, occupational rank, and psychiatric status. *Archives of General Psychiatry*, *50*, 565–576.

Mezulis, A.H., Abramson, L.Y., Hyde, J.S., & Hankin, B.L. (2004). Is there a universal positivity bias in attributions? A meta-analytic review of individual, developmental, and cultural differences in the self-serving attributional bias. *Psychological Bulletin*, *130*, 711–747.

Muncer, S.J., Craigie, M., & Holmes, J. (2003). Meta-analysis and power: Some suggestions for the use of power in research synthesis. *Understanding Statistics*, *2*, 1–12.

*Nolan, E.E., & Gadow, K.D. (1997). Children with ADHD and tic disorder and their classmates: Behavioral normalization with methylphenidate. *Journal of the American Academy of Child and Adolescent Psychiatry*, *36*, 597–604.

Pelham, W.E., Fabiano, G.A., & Massetti, G.M. (2005). Evidence-based assessment of attention deficit hyperactivity disorder in children and adolescents. *Journal of Clinical Child and Adolescent Psychology*, *34*, 449–476.

Platzman, K.A., Stoy, M.R., Brown, R.T., Coles, C.D., Smith, I.E., & Falek, A. (1992). A review of observational methods in attention deficit hyperactivity disorder (ADHD): Implications for diagnosis. *Psychology Quarterly*, *7*, 155–177.

Porrino, L.J., Rapoport, J.L., Behar, D., Sceery, W., Ismond, D.R., & Bunney, W.E. Jr. (1983). A naturalistic assessment of the motor activity of hyperactive boys: I. Comparison with normal controls. *Archives of General Psychiatry*, *40*, 681–687.

Radosh, A., & Gittelman, R. (1981). The effect of appealing distractors on the performance of hyperactive children. *Archives of General Psychiatry*, *40*, 681–687.

Rapport, M.D., Denney, C.B., DuPaul, G.J., & Gardner, M.J. (1994). Attention deficit disorder and methylphenidate: Normalization rates, clinical effectiveness, and response prediction in 76 children. *Journal of the American Academy of Child and Adolescent Psychiatry*, *33*, 882–893.

Rapport, M.D., Kofler, M., Alderson, M., & Raiker, J. (2007). Attention-deficit/hyperactivity disorder. In M. Hersen & D. Reitman (Eds.), *Handbook of psychological assessment, case conceptualization and treatment, volume 2: Children and adolescents*. NJ: Wiley & Sons.

Rapport, M.D., Murphy, A., & Bailey, J.S. (1982). Ritalin vs. response cost in the control of hyperactive children: A within-subject comparison. *Journal of Applied Behavior Analysis*, *15*, 205–216.

Rapport, M.D., Scanlan, S.W., & Denney, C.B. (1999). Attention-deficit/hyperactivity disorder and scholastic achievement: A model of dual developmental pathways. *Journal of Child Psychology and Psychiatry*, *40*, 1169–1183.

*Roberts, M.A. (1990). A behavioral observation method for differentiating hyperactive and aggressive boys. *Journal of Abnormal Child Psychology*, *18*, 131–142.

Rosenthal, R. (1995). Writing meta-analytic reviews. *Psychological Bulletin*, *118*, 183–192.

Rosenthal, R. (1991). *Meta-analytic procedures for social research*. Thousand Oaks, CA: Sage.

Schachar, R., Sandberg, S., & Rutter, M. (1986). Agreement between teachers' ratings and observations of hyperactivity, inattentiveness, and defiance. *Journal of Abnormal Child Psychology*, *14*, 331–345.

*Shecket, S.M., & Shecket, W.C. (1976). *Behaviors of children referred by classroom teachers as hyperactive*. Washington State University, (ERIC Document Reproduction Service No. ED 129–440).

*Skansgaard, E.P., & Burns, G.L. (1998). Comparison of DSM-IV ADHD combined and predominantly inattention types: Correspondence between teacher ratings and direct observations of inattentive, hyperactivity/impulsivity, slow cognitive tempo, oppositional defiant, and overt conduct disorder symptoms. *Child and Family Behavior Therapy*, *20*, 1–14.

*Solanto, M.V., Abikoff, H., Sonuga-Barke, E., Schachar, R., Logan, G.D., Wigal, T., et al. (2001). The ecological validity of delay aversion and response inhibition as measures of impulsivity in AD/HD: A supplement to the NIMH multimodal treatment study of AD/HD. *Journal of Abnormal Child Psychology*, *29*, 215–228.

Sykes, D.H., Douglas, V.I., Weiss, G., & Minde, K.K. (1971). Attention in hyperactive children and the effect of methylphenidate (Ritalin). *Journal of Child Psychology and Psychiatry*, *12*, 129–139.

Tabachnick, B.G., & Fidell, L.S. (2007). *Using multivariate statistics (5th edn)*. Needham Heights, MA: Allyn & Bacon.

Timko, T.M., Kofler, M.J., & Rapport, M.D. (2004). *Attention-deficit/hyperactivity disorder: Competing models of attention*. Paper presented at the 16th annual conference of the American Psychological Society.

Volpe, R.J., DiPerna, J.C., Hintze, J.M., & Shapiro, E.S. (2005). Observing students in classroom settings: A review of seven coding schemes. *School Psychology Review*, *34*, 454–474.

*Werry, J.S., & Quay, H.C. (1969). Observing the classroom behavior of elementary school children. *Exceptional Children*, *35*, 461–470.

Whalen, C.K., Collins, B.E., Henker, B., Alkus, S.R., Adams, D., & Stapp, J. (1978). Behavior observations of hyperactive children and methylphenidate (Ritalin) effects in systematically structured classroom environments: Now you see them, now you don't. *Journal of Pediatric Psychology*, *3*, 177–187.

Wilson, M.S., & Reschly, D.J. (1996). Assessment in school psychology training and practice. *School Psychology review*, *25*, 9–23.

World Health Organization. (1993). *The ICD-10 classification of mental and behavioural disorders: Diagnostic criteria for research*. Geneva: World Health Organization.

Yang, P., Jong, Y., & Chung, L. (2004). Gender differences in a clinic referred sample of Taiwanese attention-deficit/hyperactivity disorder children. *Psychiatry and Clinical Neurosciences*, *58*, 619–623.

*Zentall, S.S. (1980). Behavioral comparisons of hyperactive and normally active children in natural settings. *Journal of Abnormal Child Psychology*, *8*, 93–109.

Zentall, S.S. (1993). Research on the educational implications of attention deficit hyperactivity disorder. *Exceptional Children*, *60*, 143–153.

Zentall, S.S., & Shaw, J.H. (1980). Effects of classroom noise on performance and activity of second-grade hyperactive and control children. *Journal of Educational Psychology*, *72*, 830–840.

Manuscript accepted 6 June 2007