Journal of Consulting and Clinical Psychology
2002, Vol. 70, No. 5, 1099–1111

0022-006X/02/$5.00 DOI: 10.1037//0022-006X.70.5.1099

Preadolescent Girls With Attention-Deficit/Hyperactivity Disorder: II. Neuropsychological Performance in Relation to Subtypes and Individual Classification

Stephen P. Hinshaw
University of California, Berkeley

Estol T. Carte
Kaiser Permanente Medical Group

Nilofar Sami and Jennifer J. Treuting
University of California, Berkeley

Brian A. Zupan
Minnesota Department of Health

This study examined executive functions, motor speed, and language processing in a diverse, preadolescent sample of 93 girls with attention-deficit/hyperactivity disorder (ADHD) combined type, 47 ADHD inattentive type, and 88 age- and ethnicity-matched comparison girls. Testing was performed without stimulant medication. All 10 neuropsychological variables showed significant omnibus subgroup differences, with 8 of 10 combined versus comparison contrasts significant (average effect size medium) and 6 of 10 inattentive versus comparison contrasts significant (average effect size small to medium), but only 2 of 10 combined versus inattentive contrasts significant (average effect size small). Results were robust to statistical control of demographic variables, comorbidities, and IQ. Discriminant function analysis revealed relatively high sensitivity but only modest specificity in predicting ADHD from comparison status from test performance; classification of the inattentive type was extremely poor.

Psychobiological and neuropsychological models of attention-deficit/hyperactivity disorder (ADHD) have come into ascendancy in recent years (Barkley, 1997a; Tannock, 1998). The predominant perspectives pertain to (a) deficits in executive functions (EF; i.e., higher order cognitive operations such as organization, planning, decision making, working memory, and set shifting, as described in Tranel, Anderson, & Benton, 1994) and related frontostriatal circuitry (Pennington & Ozonoff, 1996); (b) impairments in inhibitory control (Barkley, 1997a, 1997b); and (c) deficits in motivation, such as deficient "cognitive energetic" resources (Van der Meere, 1996) or delay aversion (Barkley, 1997b; Sonuga-Barke, 1994). Yet nearly all existing data are based on boys meeting criteria for ADHD (Arnold, 1996; Gaub & Carlson, 1997). Our chief goal was to examine thoroughly the neuropsychological performance of a large and well-characterized sample of preadolescent girls with ADHD, featuring both combined and inattentive types, in contrast with a comparison sample of girls equivalent in age and ethnicity, toward the ends of (a) uncovering cognitive and neuropsychological processes related to the clinical manifestations of ADHD, (b) validating the disorder on the basis of objective measurements rather than potentially biased parent or teacher reports (Seidman, Biederman, Faraone, Weber, & Ouelette, 1997), and (c) encouraging investigation of mechanisms related to ADHD in girls.

Systematic research on EF with well-defined samples of youths with ADHD has appeared relatively recently. In their review of evidence through the early 1990s, Barkley, Grodzinsky, and DuPaul (1992) concluded that evidence for ADHD-related deficits in EF was inconsistent because of small and heterogeneous samples and the diversity of the neuropsychological tests used. In the past decade, using well-characterized samples of sufficient size, several investigators have uncovered clearer evidence for deficits in neuropsychological functions and EF in young men and boys with ADHD. For example, Seidman, Biederman, Faraone, Weber, and Ouelette (1997) contrasted 118 Caucasian, upper-middle-class males, ages 9–22 years, and carefully diagnosed with ADHD on the basis of the revised third edition of the *Diagnostic and Statistical Manual of Mental Disorders* (*DSM–III–R*) criteria (American Psychiatric Association, 1987) with 99 age-matched comparison males on a 2.5-hr neuropsychological battery. The ADHD sample was impaired on several tasks of EF and attentional processing but not on simple motor speed or basic perceptual processes. Group differences were maintained with statistical control of psychiatric comorbidity, socioeconomic status, and learning disabilities. Statistical control of medicated versus unmedicated status did not reveal medication effects (two thirds of the sample were medicated, including 48% receiving stimulant medication), yet it would be prudent to test children with ADHD while unmedicated, as

Stephen P. Hinshaw and Jennifer J. Treuting, Department of Psychology, University of California, Berkeley; Estol T. Carte, Department of Pediatrics, Kaiser Permanente Medical Group, San Rafael, California; Nilofar Sami, Institute for Human Development, University of California, Berkeley; Brian A. Zupan, Center for Health Statistics, Minnesota Department of Health, St. Paul.

The research presented herein was supported by National Institute of Mental Health Grant R01 MH45064. We thank the many research assistants who assisted with data collection and data coding, and we are especially grateful to the girls and families who participated in our research summer camp programs.

Correspondence concerning this article should be addressed to Stephen P. Hinshaw, Department of Psychology, Tolman Hall 1650, University of California, Berkeley, California 94720-1650. E-mail: hinshaw@socrates.berkeley.edu

stimulants enhance performance on neuropsychological measures (Aman, Roberts, & Pennington, 1998; Swanson, McBurnett, Wigal, & Pfiffner, 1993) and improve working memory via frontal circuitry (Mehta et al., 2000).

Carte, Nigg, and Hinshaw (1996) and Nigg, Hinshaw, Carte, and Treuting (1998) reported parallel findings with a group of 171 ethnically and socioeconomically diverse preadolescent boys. The ADHD sample ($n = 100$), tested without medication, yielded deficits relative to a comparison sample ($n = 71$) on tasks demanding motor speed for complex movements, sequential processing, and planning. Whereas disruptive comorbidity produced deficits in verbal IQ and whereas reading disorder was specifically linked with deficits in rapid naming, such comorbidities did not explain the ADHD-related executive deficits (see also Purvis & Tannock, 2000). More recently, Clark, Prior, and Kinsella (2000) found clear EF deficits in adolescents with ADHD, using ecologically valid measures of executive processes, with aggressive comorbidity and reading difficulties controlled.

Regarding neuroanatomical specificity, Aman et al. (1998) contrasted carefully matched samples of ADHD and comparison boys, ages 10–14 years, on three tasks tapping frontally mediated EF and three tasks measuring spatial abilities believed to be subsumed by the right parietal lobe. They found evidence for ADHD-related deficits (during off-medication testing sessions) for both types of tasks, with somewhat larger deficits on the frontal–executive compared with the parietal measures. Beyond neuropsychological tests, a growing body of research has shown evidence for small, but significant, neuroanatomic differences between ADHD and comparison samples, with frontal and striatal areas most consistently implicated (Tannock, 1998). Such neuroanatomic differences correlate with deficits in performance on tasks involving response inhibition and set shifting (Casey, Castellanos, Giedd, & March, 1997; Semrud-Clikeman et al., 2000). Furthermore, investigations using positron emission tomography and functional magnetic resonance imaging have implicated metabolic abnormalities in prefrontal, premotor, and striatal regions of individuals with ADHD (e.g., Rubia et al., 1999; Zametkin et al., 1990; for a review, see Tannock, 1998). Despite such evidence, (a) far more work needs to be done on defining core deficits in ADHD (Barkley, 1997a), (b) neuroimaging investigations are still in their infancy, and (c) young women and girls are severely underrepresented in extant research (Arnold, 1996).

The meta-analytic review of Gaub and Carlson (1997) suggested that girls with ADHD may display greater cognitive and verbal impairment than boys with ADHD, but such conclusions were based on extremely small female samples. Seidman, Biederman, Faraone, Weber, Mennin, and Jones (1997) conducted a pilot study of 43 girls with *DSM–III–R*-diagnosed ADHD and 36 comparison girls using the same battery as Seidman, Biederman, Faraone, Weber, and Oulette (1997). They found no significant group differences, suggesting that girls with ADHD do not show the same levels of EF deficits as do boys. The sample size was relatively small, however, limiting statistical power, and 84% of their ADHD sample were receiving stimulant medication at the time of testing.

The more recent investigation of Klorman et al. (1999) included both boys and girls, ages 7–13 years (overall $N = 359$). Children with ADHD combined type ($n = 207$) showed clear EF deficits, whereas (a) children with ADHD inattentive type ($n = 102$) did not and (b) comorbid status with oppositional–defiant disorder (ODD) or with reading disorder did not qualify the core findings. Covarying for sex did not alter the basic pattern of results, suggesting EF deficits in the female sample. Furthermore, Castellanos et al. (2000) investigated performance on oculomotor tasks in a sample of unmedicated girls with ADHD, finding large ADHD-related deficits in EF on delayed response and go/no-go tasks but not on the nonexecutive task of smooth eye pursuit. Thus, recent evidence suggests EF deficits in girls with ADHD.

Several key issues pertain to investigations of neuropsychological functioning in child psychopathology. First, what are the relevant domains and tests? We emphasized the EF of set shifting and set maintenance, visual planning, response organization, working memory, and response inhibition (Nigg et al., 1998; Seidman, Biederman, Monuteaux, Weber, & Faraone, 2000; Tranel et al., 1994) but also included tests of speed of word retrieval, rule-based verbal processing, and speeded motor performance to sample other relevant neuropsychological domains.

Second, because of the formal recognition of subtypes of ADHD in the fourth edition of the *Diagnostic and Statistical Manual of Mental Disorders* (*DSM–IV*; American Psychiatric Association, 1994) and the likelihood that girls with ADHD may be relatively more likely to display the inattentive type (Lahey et al., 1994), we explicitly sampled both ADHD combined type and ADHD predominantly inattentive type. Although Klorman et al. (1999) found evidence for greater levels of executive dysfunction in the former (see Hynd et al., 1991; Marshall, Hynd, Handwerk, & Hall, 1997), many other investigations have yielded null findings (see Milich, Balentine, & Lynam, 2001). Given potential problems with the *DSM–IV* definition of the inattentive type (McBurnett, Pfiffner, & Frick, 2001), we also categorized this subtype on the basis of sluggish cognitive tempo, which may correspond more closely to a “purely” inattentive type.

Third, any group differences in neuropsychological test performance could actually reflect the influences of confounding variables, such as age, socioeconomic status, or comorbidities. We therefore included such covariates following our primary analyses. Statistical control of IQ is controversial in research on ADHD. Some contend that it constitutes “overcontrol” because part of any ADHD comparison group differences may be inherently related to IQ discrepancies (Barkley, 1998; see also Seidman, Biederman, Faraone, Weber, & Oulette, 1997); yet it would be important to see whether any EF deficits were independent of general cognitive ability. Thus, we conducted our primary analyses both without and with IQ as a covariate.

Fourth, given the association between statistical significance and both sample size (Cohen, 1988) and multiple statistical testing, we delimited the number of neurospsychological variables analyzed herein to 10, and we feature effect sizes in our interpretation of the importance of subgroup differences. Finally, past research has indicated the presence of high rates of imprecision when group or subgroup differences are supplemented by individual prediction of diagnostic status from cognitive–neuropsychological test performance (Doyle, Biederman, Seidman, Weber, & Faraone, 2000; Perugini, Harvey, Lovejoy, Sandstrom, & Webb, 2000). We therefore performed discriminant function analysis using the neuropsychological measures as predictors of group (ADHD vs. comparison) and subgroup status. A key rationale for these classification analyses is the call for the diagnosis of ADHD to be based, in the future, on objective cognitive and neuropsychological measures rather than on fallible behavior ratings from adult informants (e.g., Conners, 2000).

Despite the sparse data on neuropsychological performance in girls with ADHD, we predicted that our ADHD sample would show deficits relative to comparison girls in EF, complex motor performance, and language-related skills, even with key covariates controlled (see Gaub & Carlson, 1997). Regarding subtype differences, we made few specific predictions, particularly because of our lowered statistical power for such comparisons; but we did hypothesize that motor slowness and difficulty in rapid naming would particularly characterize the inattentive type (Hynd et al., 1991). Finally, on the basis of past research (e.g., Doyle et al., 2000), we hypothesized that diagnostic classification on the basis of test scores would yield only modest accuracy.

Method

Overview of Procedure

We conducted three consecutive summer research programs (1997, 1998, 1999) for both girls with ADHD and comparison girls (see Hinshaw, 2002, for rationale and details). Following recruitment from multiple sources, we performed a multiple-gating procedure of screening and diagnostic assessment each winter and early spring, followed by thorough evaluation of comorbidities, impairments, and cognitive functioning (including the 1st hr of neuropsychological testing) in the late spring. During the summer programs, which were located on the campus of a local school, we featured a structured series of classroom, art, drama, and outdoor activities that afforded observation of naturalistic social interactions and collection of peer sociometric data. Two additional hours of neuropsychological testing were performed during individual "pull outs" at the summer program. We performed the neuropsychological assessments in three separate sessions to minimize fatigue and to assess test–retest reliability for some of the measures. For girls with ADHD receiving stimulant medication, all testing was performed following a minimum 24-hr medication washout period. Staff were unaware of the diagnostic status of participants.

Participants

Girls with ADHD and comparison girls were recruited in parallel formats. For the ADHD sample, we sent mailings to medical settings (including health maintenance organizations), mental health centers, pediatric practices, and local school districts; we gave talks at self-help groups; and we placed advertisements in local newspapers. For the comparison girls, we sent similar mailings to school districts and community centers in the San Francisco Bay Area; we also placed identical or parallel advertisements in the same local newspapers (in the parallel ads, wording emphasized "summer enrichment programs" rather than "summer enrichment programs for girls with attentional problems"). The wide range of referral sources and the heterogeneity of the San Francisco Bay Area ensured that the sample was diverse ethnically (53% Caucasian, 27% African American, 11% Latina, 9% Asian American) and socioeconomically (families from the highest strata to those receiving public assistance). For the ADHD and comparison girls, the age range was 6–12 years. Overall exclusionary criteria included an IQ lower than 70; overt neurological damage, psychosis, or pervasive developmental disorder; and medical conditions precluding participation in a summer camp.

At the first gate, interested families contacted the project representative by phone. Approximately a third of roughly 1,200 initial callers across the 3 years of recruitment made inquiries about programs for children outside our age range or for boys. We mailed program descriptions to the remaining families. Following the mailing, still interested families called to schedule phone intakes ($n = 709$). For the second gate, we mailed packets of questionnaires to eligible families, one set for parents and another for the child's primary teacher. We received completed forms back from 63% (450/709). Participants with ADHD had to show (a) Swanson, Nolan, and Pelham (SNAP) Parent Inattention *and* Teacher Inattention Scales (Swanson, 1992) endorsed with at least 5 of 9 items positive—that is, at a level of 2 (*pretty much*) or 3 (*very much*) on the 4-point scale—with Parent and Teacher Hyperactivity/Impulsivity Scales ranging from 0 to 9 items positive, and (b) Child Behavior Checklist (CBCL) *and* Teacher Report Form Attention Problem scores (Achenbach, 1991) of at least $T = 60$, a cutoff validated by Chen, Faraone, Biederman, and Tsuang (1994). For girls receiving medication, parents and teachers were asked to rate unmedicated behavior patterns. All parents did so, but 35% of teachers had seen girls exclusively on medication, as determined by a question on the cover sheet of the scales; for these cases, we used parental information only. (Chen et al., 1994, report high validity for parent ratings in terms of diagnosing ADHD.) These cutoffs (five SNAP symptoms; CBCL > 60) were intentionally set low to not eliminate potentially eligible girls on the basis of initial rating scales. For comparison girls, scores had to be below these cutoffs. Of the completed packets, 62% (278/450) met screening criteria for either ADHD or comparison status. These families were invited for diagnostic assessments; ineligible families were referred out.

At the diagnostic evaluations, after completing consent and assent forms, parents were administered the Diagnostic Interview Schedule for Children (4th ed.; DISC–IV; Shaffer, Fisher, Lucas, Dulcan, & Schwab-Stone, 2000), and girls were assessed for general intelligence with the Wechsler Intelligence Scale for Children (3rd ed.; WISC–III; Wechsler, 1991). The DISC–IV and WISC–III were administered by highly trained graduate students at the bachelor's or master's level, who were unaware of putative diagnostic status. As noted above, previously medicated girls were unmedicated for these sessions, and parents responded to DISC–IV questions with regard to their daughter's unmedicated behavior. We note that asking parents about unmedicated behavior patterns for ascertaining diagnostic status is necessary for evaluating families of children who are medicated, given the strong influence of stimulant medications on the core behaviors of ADHD (Greenhill & Osman, 2000); this procedure was performed in the Multimodal Treatment Study of Children with ADHD as well (Hinshaw et al., 1997). For final eligibility in the ADHD group, the girl had to meet full criteria for ADHD (either combined or inattentive type) on the DISC–IV and Parent SNAP ratings (note that the DISC–IV algorithm includes duration and impairment criteria in its algorithms; Shaffer et al., 2000). Eligibility for the comparison group required a negative score on the DISC–IV diagnosis of ADHD. Of families attending this gate (6 of the 278 did not show up), 90% (245/272) met full criteria and were accepted into the program. Seventeen of these 245 declined the program or encountered space limitations; the final sample included 140 girls with ADHD and 88 comparison girls.

To promote generalizability of the ADHD sample, we did not exclude common comorbidities (disruptive behavior disorders, anxiety disorders, depression). Although comparison girls could not meet criteria for ADHD, because we did not wish to constitute a "supernormal" comparison sample, we did not exclude ODD or internalizing disorders (Kendler, 1990). In addition, comparison girls may have displayed some of the constituent behaviors of ADHD but could not show either above-threshold symptomatology or above-threshold impairment.

In terms of ADHD subtypes, because the hyperactive–impulsive (HI) subtype may be particularly relevant for preschool-age children (Lahey et al., 1994) and because of our desire to have maximum statistical power for inattentive versus combined type contrasts, we excluded girls meeting criteria for the HI type a priori. Second, to diagnose combined versus inattentive subtypes, we initially used the following formula: If the DISC–IV HI symptom count was greater than or equal to 6 (American Psychiatric Association, 1994) *and* the SNAP Parent HI symptom count was also greater than or equal to 6, the girl was designated as combined;

otherwise, she was designated as inattentive. Yet because the *DSM–IV* criteria present some difficulties for subtype designation (McBurnett et al., 2001)—that is, a child with at least 6 Inattentive items plus 5 HI symptoms is designated as inattentive, the same as an inattentive child with 0 HI symptoms—we reexamined girls ($n = 28$) for whom DISC–IV HI symptoms counts were at 4, 5, or 6 and parent SNAP symptoms counts were also at 4, 5, or 6. For these cases, four senior staff examined diagnostic materials following the summer program and made a clinical judgment as to combined versus inattentive types. In 46% of these cases ($n = 13$), the judgments were identical to those from the *DSM–IV* algorithm; for the remaining 54% ($n = 15$), the consensus judgment reversed the algorithm. Overall, given the consensus judgment, 93 of the 140 girls with ADHD were designated as combined and 47 as inattentive. Note, however, that when we reconducted all primary analyses reported herein using the strict *DSM–IV* algorithm, results from all omnibus analyses were identical, and only 1 of 30 subgroup contrasts changed in significance level (i.e., a combined vs. inattentive contrast was reduced to marginal significance), with effect sizes virtually identical. Thus, we present data for the consensus subgroup designations.

Table 1 presents descriptive information about the ADHD subtypes and the comparison sample regarding demographic variables, behavioral ratings, and comorbidity-related and cognitive covariates (see also Hinshaw, 2002). Groups were not statistically distinguishable with respect to demographic variables, but as expected, the girls with ADHD had elevated rates of comorbid diagnoses and significantly lower IQ scores than did the comparison girls. Whereas Parent SNAP Inattention and HI scores were 1 to 1.5 symptoms higher than the comparable Teacher SNAP scores, both informants clearly distinguished the combined versus inattentive types with respect to HI scores but not with respect to Inattention scores (indeed, effect sizes for subgroup differentiation were nearly identical for parent vs. teacher informants; see Hinshaw, 2002, Table 1). Thus, the ADHD subgroups were matched in terms of inattention and disorganization, but HI and Externalizing scores were far higher in the combined than in the inattentive type.

As an alternative means of identifying inattentive participants, we used the Sluggish Cognitive Tempo factor of McBurnett et al. (2001), composed of “forgetful” from the *DSM–IV* inattentive list plus “daydreams” and “sluggish/drowsy,” included as questions on the SNAP–IV (Swanson, 1992). We defined this new subgroup ($n = 14$) with the following parental criteria: (a) presence of at least two of the three sluggish items, with presence defined as ratings of 2 (*pretty much*) or 3 (*very much*) on the 4-point scale, and (b) four or fewer SNAP–IV HI items. Without this latter criterion, over half of the 38 girls with two or three sluggish items endorsed would have come from the combined type. The 14 girls in the stringently defined sluggish tempo subgroup are therefore a small subset of the 47 girls from the *DSM–IV*-defined inattentive subgroup.

Cognitive and Neuropsychological Measures

We selected well-established and well-validated neuropsychological tests. To delimit multiple statistical testing, we prespecified only one key outcome measure from each test, except for the Rey–Osterrieth Complex Figure Design (ROCF; Osterrieth, 1944) and the Conners’ Continuous Performance Test (CPT; Conners, 1995), for which we selected two

Table 1
Demographic, Behavioral, Comorbidity-Related, and Intellectual Variables by Diagnostic Subgroup

	Combined ($n = 93$)		Inattentive ($n = 47$)		Comparison ($n = 88$)		
Variable	*M*	*SD*	*M*	*SD*	*M*	*SD*	p^a
Age (months)[b]	114.4	20.2	118.0	20.2	113.2	19.8	*ns*
Annual family income[b]	6.1	2.6	6.3	2.8	6.7	2.5	*ns*
White (%)	55.9		57.4		46.6		*ns*
Public assistance (%)	15.1		14.9		11.4		*ns*
Two-parent household (%)	64.5		68.1		77.3		*ns*
Reading disorder (%)[b]	11.1		14.9		4.5		*ns*
DISC–IV Disruptive (%)[b]	75.0_a		48.9_b		6.8_c		$<.01$
DISC–IV Internalizing (%)	35.6_a		23.4_a		3.4_b		$<.01$
Mother							
CBCL Attention Problems T	74.8_a	8.7	74.0_a	9.1	52.2_b	4.6	$<.01$
Externalizing T	68.7_a	8.2	58.6_b	11.0	45.9_c	8.2	$<.01$
CBCL Internalizing T	60.8_a	10.3	60.2_a	10.2	47.4_b	11.2	$<.01$
SNAP Inattention (0–9)	7.5_a	2.1	7.7_a	1.5	0.4_b	0.9	$<.01$
SNAP HI (0–9)	6.7_a	2.3	3.0_b	2.2	0.2_c	0.6	$<.01$
Teacher							
SNAP Inattention (0–9)[c]	6.8_a	2.0	6.3_a	2.4	0.3_b	1.0	$<.01$
SNAP HI (0–9)[c]	5.1_a	2.9	2.0_b	2.2	0.1_c	0.5	$<.01$
WISC–III Full Scale IQ[b]	99.6_a	13.2	99.8_a	14.3	112.0_b	12.7	$<.01$

Note. For annual family income, $1 \leq \$10{,}000$; $9 \geq \$75{,}000$. Reading disorder = Wechsler Individual Achievement Test Reading score < 85. Diagnostic Interview Schedule for Children—Version IV (DISC–IV) Disruptive = oppositional–defiant disorder or conduct disorder diagnoses; DISC–IV Internalizing = anxiety disorder (over and above specific phobias), depression, and/or dysthymic disorder diagnoses; CBCL = Child Behavior Checklist; SNAP = Swanson, Nolan, and Pelham Rating Scale; HI = hyperactive–impulsive; WISC–III = Wechsler Intelligence Scale for Children—3rd Edition.
[a] Significance: one-way analysis of variance for continuous variables; Pearson’s chi-square statistic for categorical variables. In rows with significant omnibus tests, entries with different subscripts differ significantly on the basis of Tukey’s post hoc comparisons or 2×2 chi-square tests. [b] Variables included as covariates in main analyses. [c] $n = 54$ for combined and 37 for inattentive type, as 35% of teacher scales were invalid because teacher could rate only medicated behavior.

measures each. Our total number of neuropsychological variables was 10, over and above the covariates of IQ (dimensional) and reading achievement (used categorically to designate reading disorder).

WISC–III. To appraise overall cognitive functioning and to exclude participants in the mentally retarded range, we administered the complete WISC–III (except for the supplemental Mazes subtest). The WISC–III is the most widely used measure of general cognitive performance, with extensive psychometric data attesting to its reliability and validity (Wechsler, 1991). We analyzed the Full Scale IQ score (FSIQ), which has internal consistency and test–retest reliability statistics $> .9$ (Wechsler, 1991).

Reading decoding. We administered the Wechsler Individual Achievement Test (WIAT; Wechsler, 1992) Reading Recognition subtest, which measures the taker's sight-reading ability of known words. Its reliability is excellent, as discussed in Hinshaw (2002). The WIAT Reading score was used to subgroup the sample on the basis of reading disorder, such that girls with a standard score < 85 were classified as reading disordered (combined: $n = 10$; inattentive: $n = 7$; comparison: $n = 4$). Consensus has emerged that low reading achievement alone is a sufficient basis for determining reading disorder and that IQ Reading discrepancy scores are no more valid than such absolute reading level scores (Lyon, 1996; Purvis & Tannock, 2000; Shaywitz, Escobar, Shaywitz, Fletcher, & Makuch, 1992).

ROCF. This classic test requires the participant to copy a complex abstract design. Although we administered both a copy condition and a delayed recall condition, we analyzed only the copy condition herein. The ROCF has distinguished patients with frontal and parietal lesions from control samples (Lezak, 1995) and boys with ADHD from comparison boys (Carte et al., 1996; Nigg et al., 1998; Seidman, Biederman, Faraone, Weber, & Oulette, 1997). We analyzed the organization score, scored on a 13-point scale from 1 to 13, representing five developmental levels (see Waber & Holmes, 1985), and an error proportion score. The latter incorporated segments drawn incorrectly (errors) divided by the sum of segments drawn correctly plus incorrectly. This error score was based on operational criteria developed from our modifications of the system of Bernstein and Waber (1996), refined with 50 ROCF drawings from the practice of Estol T. Carte (see Sami, Carte, Hinshaw, & Zupan, 2002, for details). We paid careful attention to the reliability of our scoring procedures, using independent raters who were unaware of diagnostic status to score drawings. For organization, the intraclass correlation between the two primary scorers was .90 (drawings $n = 139$). For errors, the intraclass correlations between pairs of three primary scorers ranged from .91 to .94 (drawings n ranged from 84 to 195 across rater pairs).

Porteus Maze Test (PM; Porteus, 1973). For this task, which requires planning and self-regulation (Lezak, 1995), the child must draw a pencil line carefully on a paper maze, moving forward to arrive at the only exit. Related spatial–planning tasks (e.g., Tower of London) involve prefrontal brain regions (e.g., Levin et al., 1994) and have developmental characteristics similar to those of the PM (Krikorian & Bartok, 1999). The PM has an extensive validation literature regarding ADHD (Barkley, 1997b; Nigg et al., 1998; Seidman, Biederman, Faraone, Weber, & Oulette, 1997). We adhered to administration and termination procedures as described in Porteus (1973) but report a modified test age, in that we did not administer inverted mazes to verify maze success following failure (Carte et al., 1996). Instead, after the specified number of unsuccessful maze attempts, we allowed up to two more attempts of the same maze without inversion. Failure of both supplemental attempts was our failure criterion. Thus, maze failure in our procedure was less stringent than in Porteus (1973). We inverted and readministered the same mazes 4–6 weeks following the main assessment to ascertain test–retest reliability, yielding $r = .72$ ($n = 147$).

Time-to-Do 20 Motor Battery (TTD-20; Denckla, 1974, 1985). We wished to measure relatively pure unrehearsed motor output, free of problem solving, as Sergeant and Scholten (1985) have specified that difficulties with motor preset and motor output characterize children with ADHD. We selected the TTD-20 for its high reliability (Vitiello, Ricciuti, Stoff, Behar, & Denckla, 1988). On the basis of findings from our laboratory (Nigg et al., 1998), from the six individual subtests, we selected complex foot movements (heel–toe rocking, right plus left) and complex hand movements (diadochokinesis, right plus left) and composited these scores for the present analyses (Stokman et al., 1986). Such speeded motor tasks involve prefrontal, premotor, and subcortical brain regions. For a large subset of cases, we had two staff score the times to complete the movements, yielding extremely high intraclass correlations, .98–.99 (ns ranged from 105 to 110). In addition, our 4–6 week test–retest reliabilities ranged from $r = .70–82$ (heel–toe, right and left; ns $= 145$ and 148, respectively), and $r = .64–.65$ (pronation–supination, right and left; ns $= 148$ and 149, respectively).

Grooved Pegboard (GPB; Knights & Norwood, 1979). This is a task that measures distal, complex fine-motor coordination; it is highly dependent on psychomotor speed. Whereas Seidman, Biederman, Faraone, Weber, and Oulette (1997) did not find ADHD–comparison differences on measures of simple finger tapping, the GPB measures speeded performance of a complex fine-motor skill. It is frequently used to assess lateralized cerebral dysfunction (Mitrushina, Boone, & D'Elia, 1999). Participants place 25 pegs into a board, first with the preferred and then with the nonpreferred hand. We analyzed the time to completion for the hand yielding the fastest speed. The right hand was the fastest hand for 86% of the sample; the left hand for 14%. Our test–retest correlations (4–6 weeks) were $r = .74$ (right hand, $n = 133$) and $r = .82$ (left hand, $n = 130$).

CPT. In this visual task of attentional processing, the child is instructed to press the computer spacebar on each presentation of a target letter, defined as any letter except for *X*, to which the child is instructed not to respond. The Conners's version differs from traditional CPTs in that it features frequent, prepotent responses that must occasionally be inhibited, thereby tapping the EF of response inhibition. Trials are presented in six blocks, with interstimulus interval ranging from 1 s, 2 s, and 4 s within each block. Stimulus display time is 250 ms. The total test takes approximately 14 min. From the several measures available from Conners's scoring procedures, we preselected omission errors (percentage of failures to respond to target stimuli out of the total number of targets presented) and commission errors (percentage of key presses for nontargets out of the total number of nontargets presented). We analyze other important measures, including reaction time, in subsequent reports.

Woodcock–Johnson Psychoeducational Battery (W-J): Word Attack subtest (Woodcock & Johnson, 1989/1990). This is a norm-referenced test, with acceptable reliability and validity, which measures the child's ability to decode nonsense words. Because the child has never encountered these pseudowords before, the test is considered to be a measure of "pure" phonologic processing and decoding abilities. As all other neuropsychological measures in the battery were not standardized for age, we analyzed the raw score, which was the number of pseudowords correctly read.

Rapid Automatized Naming (RAN; Denckla & Rudel, 1974). Here, children rapidly name repeating sets of automatically processed digits and pictured stimuli, which require effortful semantic processing and retrieval. This type of task appears to require left-prefrontal brain regions (Gabrieli et al., 1996). Carte et al. (1996) found that, for boys, optimal ADHD versus comparison group differentiation was found for the naming of common objects; we therefore prioritized this task for the present female sample. Nigg et al. (1998) found that RAN scores were related specifically to reading disorder over and above ADHD status; we wished to evaluate the same issue for our female sample. Our score was an accuracy rate, defined as 50 (i.e., the total number of stimuli to be named) minus the sum of incorrect recognitions plus omissions, with that quantity divided by the time to complete the task. From prior research with boys, we obtained an alpha reliability of .79 for naming times from alternate sets of pictures objects (Nigg et al., 1998).

Cancel Underlining (CUL). This modification of the Underlining Test (Rourke & Orr, 1977) measures rapid, accurate visual discrimination. In research with boys, discrimination of a consonant sequence (*fsbm*: Subtest

9) optimally discriminated ADHD from comparison participants (Carte et al., 1996; Nigg et al., 1998), with errors (rather than slow speed) characterizing ADHD (MacLeod & Prior, 1996). Anterior cingulate prefrontal regions are involved in such tasks (Cabeza & Nyberg, 1997). For our girls, we administered two uppercase consonant sequences (*GWRS* and *DHCT*) among distractor sequences of the same letters in different orders, using a fast/accurate instruction set. We eliminated vowels and used uppercase letters to reduce similarity to automatically processed sight words; similarly, we used two different consonant sets to prevent automatization or memorization. We instructed the girls to underline targets and to cancel out nontargets (ratio of 1:5) to establish a set of cancellation. Underlining of the rare targets thus requires inhibition of a prepotent response. We analyzed correct minus incorrect responses (Nigg et al., 1998).

Overall, on the basis of conceptual and empirical formulations (Tranel et al., 1994), our measures of EF include the two variables from the ROCF, the two from the CPT, and one each from PM and CUL. Our two measures of motor speed are the TTD-20 and GPB; our two language-related measures constitute W-J Word Attack and RAN.

Covariates

First, because our dependent measures were raw scores rather than age-standardized scores, we reasoned that age should exert clear effects on these scores, and we used age (in months) as a dimensional covariate. Second, we entered an ordinal variable regarding family income (1–9) as a proxy for socioeconomic status. Third, comorbidity is a potential confound in neuropsychological investigations of children with ADHD (Nigg et al., 1998; Seidman, Biederman, Faraone, Weber, & Oulette, 1997). For our categorical indicator of reading disorder, we used the cutoff of less than 85 on the WIAT Reading Recognition test. For comorbid disruptive disorders, we formed a categorical variable of girls with disruptive disorders from those who had *DSM–IV* diagnoses of ODD or conduct disorder (CD) (coded as 1) versus those who had neither disorder (coded as 0), using parental data from the DISC–IV (Shaffer et al., 2000), scoring diagnoses in terms of current presence (which, for the DISC–IV, includes the time period from 12 months before the interview until the present). For comorbid internalizing disorders, we formed a categorical variable of girls who had one or more anxiety disorders (over and above specific phobias) and/or major depression or dysthymia (coded as 1), versus those without such diagnoses (coded as 0), again using the DISC–IV current diagnoses. Our designations follow directly from the *DSM–IV* in that ODD and CD are listed as the disruptive behavior disorders, and anxiety disorders and depression are the prototypic internalizing categorizations (see Achenbach, 1991). We also entered FSIQ as a covariate, given that it is the most robust estimate of overall intellectual ability.

Procedure

Neuropsychological assessment was performed by Estol T. Carte as well as by extensively trained graduate students and bachelor's-level research assistants, who were supervised closely by Estol T. Carte. Our goal was to have a core, each summer, of major neuropsychological instruments (those featured in the present analyses), as well as more experimental measures (e.g., sequencing tasks in which we varied instructional set across repeated assessments). For the latter, we varied tasks somewhat in successive summer programs, meaning that the overall order of test administration did not stay constant across the three samples each summer.

Examiners were unaware of the diagnostic status of the participants. For girls with ADHD receiving stimulant medication at the time of their referral, neuropsychological tests were administered after a minimum 24-hr medication washout. However, 16 (11.4%) of the ADHD sample were receiving other psychotropic medication (e.g., selective serotonin reuptake inhibitors, clonidine) during the summer program. When we analyzed data excluding these cases, the major findings were preserved almost entirely. We thus present our data analyses below for all 140 girls with ADHD.

Data Analyses

Missing neuropsychological scores were rare; reasons were absences, task refusal, or scoring software difficulties (CPT). From among the total sample of 228 girls, sample sizes for the tests ranged from $n = 219$ (CPT) to $n = 225$ (GPB). All statistical analyses were performed with SPSS for Windows, Version 10 (SPSS, Inc., 1999). After inspecting data carefully for out-of-range values, our initial data analytic step was to perform a multivariate analysis of variance (MANOVA) on the entire set of 10 neuropsychological measures, with the independent variable constituting diagnostic subgroup (combined, inattentive, comparison). Following a significant MANOVA, we used univariate analyses of variance (ANOVAs) with respect to each measure (retaining the traditional alpha level of .05 for each ANOVA), then used Tukey's post hoc contrasts for all three subgroup comparisons (combined vs. inattentive, combined vs. comparison, inattentive vs. comparison). We emphasize effect sizes in our interpretation, calculated as the subgroup mean difference divided by the pooled standard deviation from the relevant ANOVA (see Hinshaw, 2002) and using Cohen's (1988) standards for effect size magnitudes: .2 = small, .5 = medium, and .8 = large. We then repeated this analytic strategy with statistical control of our covariates. To examine the predictability of subgroup status from our battery, we performed discriminant function analysis (DFA), using any neuropsychological variables showing significant subgroup effects in the main analysis. For significant functions, we constructed classification tables to examine the sensitivity (proportion of cases predicted to be in the clinical range) and specificity (proportion of noncases predicted to be under the clinical threshold) of group (ADHD vs. comparison) and subgroup (combined, inattentive, comparison) membership from the optimal linear combination of test scores, using the actual proportions of girls in each group or subgroup as the prior probability.

Results

Main Subgroup Analyses

The overall MANOVA yielded a highly significant effect of subgroup, Pillai's trace = .251, $F(20, 386) = 2.76$, $p < .01$. Follow-up univariate analyses revealed significant omnibus one-way ANOVAs for each of the 10 neuropsychological variables (see Table 2). The typical rank ordering was for the combined type to score worst, the inattentive type somewhat better, and the comparison girls best. Exceptions were the motor speed tests (TTD-20, GPB), on which the inattentive girls were ranked slowest, and CPT commissions, on which the inattentive girls made the fewest number of errors.

The subscripts in Table 2 reveal the significance of pairwise contrasts of subgroups; effect sizes for these contrasts are found in Section I of Table 3. First, girls in the combined type scored significantly worse than comparison girls for 8 of the 10 variables, with effect sizes ranging from medium (most variables) to large (ROCF error proportion). The exceptions were TTD-20 and CPT commissions, for which effect sizes were small. Second, the inattentive type scored significantly worse than comparisons for 6 of the 10 variables, with medium effect sizes for all 6 (when all 10 variables are counted, effect sizes range from medium to small). Third, the combined type diverged significantly from the inattentive type for only 2 of the 10 measures, ROCF errors and CPT commissions, with combined type displaying worse performance. These two effect sizes were of medium strength, but all others were small. Overall, girls with ADHD (particularly those in the combined type) differed significantly from comparison girls, with effects of at least medium strength; but the ADHD subtypes were

Table 2
Neuropsychological Performance by Subgroup

	Combined (n = 93)		Inattentive (n = 47)		Comparison (n = 88)		
Variable	*M*	*SD*	*M*	*SD*	*M*	*SD*	p^a
ROCF organization, 1–13	5.36_{a}	3.12	$6.23_{a,b}$	3.15	6.81_{b}	3.33	.05
ROCF error proportion	0.37_{a}	0.19	0.30_{b}	0.17	0.22_{c}	0.15	.01
Porteus Maze: Test age	11.64_{a}	2.90	11.63_{a}	2.95	13.13_{b}	2.17	.01
TTD-20 Complex, time	$18.24_{a,b}$	5.98	19.37_{a}	7.92	16.48_{b}	3.67	.05
Grooved Pegboard, time	80.78_{a}	20.67	83.72_{a}	23.18	72.86_{b}	14.78	.01
Conners's CPT omissions (%)	12.04_{a}	15.76	$9.13_{a,b}$	11.29	5.65_{b}	7.82	.01
Conners's CPT commissions (%)	59.47_{a}	19.93	50.43_{b}	21.92	$53.33_{a,b}$	20.76	.05
W-J Word Attack (raw correct)	12.87_{a}	7.73	14.00_{a}	8.19	17.70_{b}	7.21	.01
RAN accuracy proportion	0.83_{a}	0.30	$0.88_{a,b}$	0.26	0.97_{b}	0.22	.05
CUL, correct minus incorrect	12.78_{a}	6.06	12.43_{a}	6.87	15.92_{b}	5.04	.01

Note. ROCF = Rey–Osterrieth Complex Figure Design; TTD–20 = Time-to-Do 20 Motor Battery; CPT = Continuous Performance Test; W-J = Woodcock–Johnson Psychoeducational Battery; RAN = Rapid Automatized Naming; CUL = Cancel Underlining Test.
[a] Significance based on one-way analyses of variance following significant omnibus multivariate analyses of variance (see text). In a given row, means with different subscripts differ significantly on the basis of Tukey's post hoc comparisons.

themselves rarely distinguishable. The ADHD–C column in Table 3 (final column under Section I) presents effect sizes for the girls with ADHD (across subtypes) versus the comparison group. As can be seen, the ROCF effect was large; the ROCF organization, PM, TTD-20, GPB, CPT omissions, Word Attack, RAN, and CUL effects were medium (or nearly medium); and the CPT commissions effect was small.

Covariates

How did these differences hold up with control of key covariates? Controlling for year of program (cohort) yielded no changes whatsoever in overall results; therefore, we excluded it from further consideration. Our key tests included two omnibus multivariate analyses of covariance (MANCOVAs): The first partialled (a) age, (b) family income, (c) comorbid reading disorder, (d) comorbid disruptive disorder, and (e) comorbid internalizing disorder; we termed this analysis the *demographic plus comorbidity set* of covariates. The second partialled all five of these covariates plus FSIQ; we termed this analysis the *full set* of covariates. The MANCOVAs for both sets yielded highly significant multivariate effects of subgroup: Demographic plus comorbidity set: Pillai's trace = .328, $F(20, 344) = 3.38$, $p < .01$; full set: Pillai's trace = .253, $F(10, 338) = 2.45$, $p = .01$. Follow-up analyses of covari-

Table 3
Neuropsychological Performance Effect Sizes by Subgroup, With Statistical Control of Covariates

	I. No covariates[a]				II. Demographic plus comorbidity set[b]				III. Full set[c]			
Variable	A–B	A–C	B–C	ADHD–C	A–B	A–C	B–C	ADHD–C	A–B	A–C	B–C	ADHD–C
ROCF organization	.27	.45	.18	.35	.10	.53	.43	.45	.16	.38	.22	.29
ROCF error proportion	.40	.84	.44	.71	.24	.90	.66	.75	.30	.73	.43	.55
Porteus Maze: Test age	.00	.54	.55	.54	−.09	.61	.71	.58	−.05	.55	.61	.49
TTD-20 Complex, time	−.20	.30	.50	.37	−.21	.15	.37	.34	−.19	.03	.22	.19
Grooved Pegboard, time	−.15	.40	.55	.45	−.13	.55	.68	.59	−.08	.28	.36	.31
Conners's CPT omissions (%)	.23	.51	.28	.43	−.05	.38	.42	.60	−.04	.31	.35	.49
Conners's CPT commissions (%)	.43	.29	−.14	.15	.60	.50	−.10	.12	.59	.49	−.10	.14
W-J Word Attack (raw correct)	.14	.61	.47	.56	.13	.52	.39	.54	.15	.36	.21	.36
RAN accuracy	.21	.53	.32	.47	.24	.44	.21	.41	.26	.37	.11	.31
CUL, correct minus incorrect	−.06	.52	.58	.54	.01	.57	.56	.53	.04	.46	.42	.52

Note. The effect sizes are Cohen's *d*. Positive value of effect sizes reflects greater deviance in the first subgroup (e.g., lower Rey–Osterrieth Complex Figure Design [ROCF] organization, higher ROCF proportion of errors). For no covariates and demographic plus comorbidity set, 10/10 analyses of variance (ANOVAs) were significant. For full set, 7/10 ANOVAs were significant; the 3 ANOVAs failing to attain significance with statistical control of the full set of covariates were ROCF organization, TTD-20, and Grooved Pegboard. For all groups, multivariate analysis of variance and multivariate analysis of covariance *p*s ≤ .001. A–B = combined group versus inattentive group; A–C = combined group versus comparison group; B–C = inattentive group versus comparison group; ADHD–C = attention-deficit/hyperactivity disorder group versus comparison group; TTD-20 = Time-to-Do 20 Motor Battery; CPT = Continuous Performance Test; W-J = Wood–Johnson Psychoeducational Battery; RAN = Rapid Automatized Naming; CUL = Cancel Underlining Test.
[a] See Table 2. [b] Covariates were age, income, reading disorder, disruptive disorder, and internalizing disorder. [c] Covariates were the same as for the demographic plus comorbidity set plus Wechsler Intelligence Scale for Children—3rd edition, Full Scale IQ.

ance (ANCOVAs) revealed that control of the demographic plus comorbidity set still yielded significant subgroup effects for each of the 10 neuropsychological variables (see Table 3). For the full set of covariates, 7 of the 10 variables maintained significance; ROCF organization, TTD-20, and GPB were rendered nonsignificant. Thus, even with this latter, stringent degree of statistical control, 5 of our 6 measures of EF (ROCF errors, PM, CPT omissions and commissions, CUL) and both measures of word retrieval and word-finding ability (Word Attack, RAN) maintained significant effects of subgroup.

Another means of conveying the ramifications of such statistical control is to denote the patterns of subgroup contrasts and effect sizes in the presence of each set of covariates (see Sections II and III of Table 3). First, regarding the combined versus comparison contrast, controlling for the demographic plus comorbidity set reduced the effect size for TTD-20 from small–medium to very small, yet slightly increased the effect size for CPT commissions and maintained the medium-to-large effect sizes for the other variables. Controlling the full set of covariates again reduced the TTD-20 effect (now to nearly nonexistent), reduced the effect for CPT omissions from .51 to .31, somewhat reduced the medium effect for RAN, maintained the effect for CPT commissions, and left the other medium or large effects roughly equivalent. Second, regarding the inattentive versus comparison contrast, controlling the demographic plus comorbidity set of covariates maintained or slightly increased the effect sizes for the ROCF variables, PM, GPB, and CPT omissions. Controlling for the full set reduced the effect sizes for tests of motor speed (TTD-20, GPB) and verbal ability (Word Attack, RAN); yet key measures of EF (ROCF, CUL) maintained medium effect sizes. Third, regarding the combined versus inattentive contrast, the only two significant effects in the main analyses were ROCF error proportion (effect size = .40) and CPT commissions (effect size = .43). Statistical control (of either set) slightly reduced the effect sizes for ROCF errors but maintained (with slight increases) the medium effect for CPT commissions.

In terms of ADHD (subtypes together) versus comparison group effect sizes (see ADHD–C columns of Sections II and III in Table 3), the demographic plus comorbidity set of covariates still yielded a large effect size for ROCF errors and medium effect sizes for ROCF organization, PM, TTD-20, GPB, CPT omissions, Word Attack, RAN, and CUL. CPT commissions remained small. The full set of covariates yielded medium effects for ROCF errors, PM, CPT omissions, Word Attack, and CUL, with small effects for the remaining variables.

Overall, the results of the main analyses proved quite robust to the statistical control of key covariates. In particular, control of demographic and comorbidity-related variables yielded almost no reduction in significance or magnitude of subgroup effects, and control of FSIQ showed its strongest reductions in significance and effect size for tests of motor speed rather than EF per se.[1]

Discriminant Function Analysis and Individual Classification of Participants

Because of the lack of salient combined versus inattentive subgroup contrasts in our primary analyses, we performed our initial DFA with a two-group classification: ADHD (merging the combined and inattentive types) versus comparison. The sample sizes in our DFAs are lower because of listwise deletion of cases for which missing values occurred for any variable. The analysis yielded a highly significant discriminant function, Wilks's $\Lambda = .831$, $\chi^2(10, N = 204) = 36.54$, $p < .01$. The function showed its highest structure loadings for ROCF errors, PM test age, CUL, Word Attack, GPB, CPT omissions, and RAN scores, all of which had loadings $> .4$. Thus, the function spanned various EF (planning, response organization, set maintenance, response inhibition) as well as linguistic skills and motor speed. When girls were classified by this function into ADHD versus comparison groups, the overall correct classification rate was 70.1% (see Tables 4 and 5). As can be seen, the sensitivity of classification was reasonably good, as 78.0% of the girls with ADHD were so classified. The specificity was weaker, however (58.0%), because of the large number of false-positive classifications on the basis of the DFA (that is, comparison girls who were wrongly predicted to have ADHD). We performed a three-group DFA as well, using combined, inattentive, and comparison subgroups. Only one function was statistically significant, nearly identical in composition to the general dimension above. The overall correct classification rate was poorer, with only 56.9% of the girls correctly classified. Whereas 72.8% of the comparison girls and 62.5% of the combined girls were classified accurately, the comparable rate for the inattentive girls was only 16.3%. Indeed, inattentive girls were far more likely to be placed in the comparison group or the combined type than to be accurately classified. Finally, an exploratory DFA with the sluggish tempo classification yielded one significant function as well, similar to those above. Only 25.0% of the sluggish tempo group was accurately classified, with classification statistics for the ADHD combined and comparison groups comparable to those in Tables 4 and 5. Overall, for both the *DSM–IV*-defined inattentive type and the sluggish tempo group, the neuropsychological battery performed poorly in terms of accurate diagnostic prediction.

Discussion

With a large, diverse, and carefully diagnosed sample of preadolescent girls with ADHD, in contrast to an age- and ethnicity-matched comparison sample of girls, we found deficits in EF, spanning self-regulation, planning, response organization, set maintenance requiring both long- and short-term memory, vigilance, and inhibitory control. Motor speed and measures of lan-

[1] In analyzing the performance of our sluggish tempo classification of girls, we found that this subgroup's scores on most of the 10 neuropsychological measures were quite similar to those of the original inattentive group. The exceptions were TTD-20 and GPB, on which the sluggish tempo group scored a standard deviation slower than the remainder of the inattentive girls, and CPT omissions and Word Attack, on which the sluggish tempo subgroup scored slightly better than the remaining inattentive girls. In a parallel set of analyses to the core analyses, we found that the sluggish tempo girls scored worse than comparisons for only 1 of 10 measures (GPB) and that they differed significantly from the combined type for only one measure, scoring better than the combined girls on Word Attack. In all, although reduced statistical power is clearly pertinent to analyses with such a small subgroup, the sluggish tempo designation did not yield substantially different performance from the *DSM–IV*-defined inattentive type, except for slower motor speed.

Table 4
Classifications of Diagnostic Groups From Discriminant Function Analysis

Predicted	Actual: ADHD	Actual: Comparison	Total
ADHD	96	34	130
Comparison	27	47	74
Total	123	81	204

Note. Sample sizes are lower than those in the core analyses because of listwise deletion of data for discriminant function analyses. Sensitivity (attention-deficit/hyperactivity disorder [ADHD] accurately predicted): 96/123 = 78.0%; specificity (comparison accurately predicted): 47/81 = 58.0%; overall correct classification rate: 143/204 = 70.1%.

guage processing were also impaired in the girls with ADHD. The combined type showed the most consistent pattern of deficits in relation to comparison girls, with contrast effect sizes ranging from medium to large. The inattentive type showed EF, motor, and linguistic deficits as well, but with effect sizes (in relation to the comparison group) that were somewhat smaller. The two ADHD subtypes differed significantly on only 2 of 10 core variables, yielding some (but limited) suggestion of greater EF deficits in the combined group. Crucially, our ADHD versus comparison differences were robust to control of age, family income, and comorbidity (with reading disorder as well as disruptive and internalizing disorders); EF and language deficits (but not motor speed deficits) were also independent of IQ. Importantly, however, despite the clear mean differences between ADHD and comparison girls, individual classification based on DFA of the neuropsychological tests yielded a plethora of false positives (i.e., misclassification of comparison girls as having ADHD), and classification of the inattentive type was extremely poor. In addition, our sluggish tempo subgroup did not yield a qualitatively different pattern of neuropsychological deficits from that of the inattentive type defined by *DSM–IV*. Overall, our findings signal key EF-, motor-, and language-related deficits in girls with ADHD but do not provide endorsement of the use of neuropsychological testing to recapture diagnostic classifications.

Our overall findings contradict the contention of Seidman, Biederman, Faraone, Weber, Mennin, and Jones (1997) that girls with ADHD do not show noteworthy deficits in EF but support the data of Castellanos et al. (2000) and Klorman et al. (1999), who found that girls with ADHD did, in fact, show deficits on various tasks measuring EF. In both of these latter studies, as in the present report, the girls were tested while off of stimulant medications. Consensus is therefore emerging that EF deficits in girls with ADHD do indeed exist.

Our results do not, however, reveal strong evidence for a differentiation of combined versus inattentive types of ADHD in terms of EF, at least in the present sample of girls and at least with respect to the particular composition of our neuropsychological battery. In fact, only 2 of our 10 neuropsychological measures yielded significant subtype differentiation: ROCF errors (our most sensitive measure of EF) and CPT commissions, both of which revealed higher percentages of errors in the combined subgroup. Two comments are salient. First, we had reduced statistical power available for detecting this subgroup contrast, as opposed to the greater power available to test ADHD versus comparison differences. Yet effect sizes (which are not sensitive to sample size) were typically smaller for the subtype contrast than for distinctions between either ADHD subgroup and the comparison group. Second, because we emphasized parental data in making our diagnostic classifications, it is possible that teachers did not rate the girls with ADHD combined type as showing sufficient HI symptoms to warrant their designation as combined, which could explain the lack of subtype differences we found. As shown in Table 1, however, for the reduced subsample of girls with ADHD whose teachers rated them during unmedicated periods, whereas HI scores were l to 1.5 symptoms lower than comparable Parent SNAP ratings, they were still considerably higher in the combined ($M = 5.1$) than they were in the inattentive ($M = 2.0$) subgroup, with comparable effect sizes regarding subgroup differentiation, as were found with the parental data (see Hinshaw, 2002). Overall, our results suggest limited support for the claim of greater executive dysfunction in the combined type.

A growing area of debate, in fact, centers around the contention that the inattentive type of ADHD is qualitatively distinct from types marked by significant levels of hyperactivity and impulsivity (Milich et al., 2001). A key criterion variable in this regard would be neuropsychological measures, which are not subject to the potential biases found in informant reports. Although we did find some evidence for higher rates of errors on the ROCF and higher rates of comissions on the CPT in the combined than in the inattentive type, we did not administer such potentially sensitive measures as the Tower of Hanoi, a task with clear EF components that yielded strong subgroup effects in Klorman et al. (1999). The search for pertinent measures may be difficult: For example, whereas Nigg, Blaskey, Huang-Pollock, and Rappley (2002) found that different tests differentiated the combined versus the inattentive type from comparison children, the types themselves were usually not statistically distinguishable. It is still possible, however, that girls with predominantly inattentive symptomatology will show neuropsychological deficits on different tests with appropriate sensitivity.

To address the potential argument that the *DSM–IV* diagnostic criteria do not accurately capture a truly inattentive subtype of

Table 5
Classifications of Diagnostic From Discriminant Function Analysis

Predicted	Actual: ADHD–C	Actual: ADHD–I	Actual: Comparison	Total
ADHD–c	50	20	20	90
ADHD–i	6	7	2	15
Comparison	24	16	59	99
Total	80	43	81	204

Note. Sample sizes are lower than those in the core analyses because of listwise deletion of data for discriminant function analyses. Sensitivity (attention–deficit/hyperactivity disorder–combined group [ADHD–C] accurately predicted): 50/80 = 62.5%; sensitivity (ADHD–inattentive [ADHD–I] accurately predicted): 7/43 = 16.3%; specificity (comparison accurately predicted): 59/81 = 72.8%; overall correct classification rate: 116/204 = 56.9%.

ADHD (Hinshaw, 2001; McBurnett et al., 2001), we made a preliminary attempt to define a subgroup of our inattentive participants who display sluggish tempo. Only 14 girls, however, showed high scores on the pertinent items along with below-threshold scores for HI symptoms. Although their chief neuropsychological characteristic was markedly slow performance on the motor speed tests (TTD-20, GPB), their performance on the other tests was not distinct from the rest of the inattentive type, and DFA could not readily identify them on the basis of test performance. Considerable work remains to be done to (a) define an inattentive variant of ADHD that poses a distinct alternative to traditional diagnostic conceptions and (b) use cognitive–neuropsychological measures that might optimally indicate such a group.

Our core findings were quite robust to statistical control of age, family income, common comorbid conditions (reading disorder, disruptive and internalizing disorders), and IQ. Independence of ADHD-related EF deficits from comorbid conditions is consistent with much current research (Clark et al., 2000; Klorman et al., 1999; Nigg et al., 1998; Seidman, Biederman, Faraone, Weber, & Oulette, 1997). Unlike Purvis and Tannock (2000), we did not find an effect of reading disorder on our measures of inhibitory control; rather, this effect was specific to ADHD status. Furthermore, when we partialled IQ in the full set of covariates, we lost significance for three measures, one pertinent to EF (ROCF organization) but two related to motor speed (TTD-20, GPB). In all, robustness of our EF deficits to statistical control provides strong evidence for the independence of ADHD-related neuropsychological dysfunction, even from general cognitive ability.

What do our particular measures tell us about the neuropsychological deficits of girls with ADHD? We note, at the outset, that nearly all neuropsychological measures are complex, making it difficult to isolate specific skills or deficits on the basis of task performance; indeed, EF are, by definition, skills tapping complex cognitive functions. Nonetheless, regarding EF, the ROCF features planning, response organization, motor control, and attention to detail. It is noteworthy that our largest ADHD–comparison effects were found on the error scoring system of this measure (see Sami et al., 2002). The PM, as noted in the *Measures* section, requires careful planning, organizational abilities, and self-regulation, classic features of executive control. Moderate-sized effects were found for this task, as they were for CUL, another measure clearly tapping components of EF. Not surprisingly, the traditional omission and commission variables from the CPT showed significant effects as well.

We included two tasks of pure motor abilities: GPB, which requires fine motor control and speed, and the TTD-20 battery, which requires (for the variables we preselected) gross motor speed. It is provocative that statistical control of FSIQ reduced ADHD and comparison effect sizes for both of these measures but not for EF per se. We plan on exploring the complete set of motor functions tapped by the TTD-20 as well as the role of controlling Verbal versus Performance IQ in subsequent publications. Word Attack measures phonologic processing ability, and RAN taps quick linguistic retrieval, both of which yielded ADHD and comparison discrepancies that, surprisingly, were not accounted for by comorbid reading disorder (Purvis & Tannock, 2000).

What additional measures could profitably be used? Castellanos et al. (2000) found ADHD and comparison differences on oculomotor executive tests, yet these had limited utility because of strong practice effects. As for response inhibition, we did not use the Stop Task, but this test may characterize aggressive children as well as those with ADHD (Oosterlaan, Logan, & Sergeant, 1998). In this regard, sound theory is needed to derive appropriate and specific tasks of inhibitory control (Nigg, 2000). Also, the ecologically valid executive measures used by Clark et al. (2000)—the Six Elements Test and the Hayling Sentence Completion Test—are less structured than traditional neuropsychological measures but tap the kinds of real-world set shifting and response organization that are thought to underlie EF. The Cognitive Assessment System of Naglieri and Das (1997) incorporates EF elements of planning, attention, and simultaneous and sequential processing in its composition. Preliminary data attest to its diagnostic accuracy in distinguishing ADHD from comparison children (Wasserman & Becker, 1999). Finally, the Behavior Rating Inventory of Executive Function shows promising psychometric properties (Gioia, Isquith, Guy, & Kenworthy, 2001), but as a rating instrument, it is subject to the biases of adult informant reports.

The results of our DFA, which showed only the modest ability of our neuropsychological battery to predict individual diagnostic classifications, were disappointing but not unexpected (Doyle et al., 2000; Perugini et al., 2000). Despite significant findings and medium (or even large) effect sizes in terms of group differences, the variability of response was sufficiently large to yield a distressingly large number of false-negative and, particularly, false-positive classifications. Problems in this regard plague not only cognitive–neuropsychological performance but also the very diagnostic algorithms used to define ADHD (Mota & Schachar, 2000). Neuropsychological specificity regarding ADHD and other child disorders remains elusive.

Several limitations of the present study are salient. First, the designation of combined versus inattentive types of ADHD is potentially problematic with *DSM–IV* criteria (McBurnett et al., 2001); reduced sample sizes lowered statistical power for this subgroup contrast. Yet deriving an alternative means of classifying a purely inattentive group has eluded clinical and research efforts. Second, as noted earlier in this section, use of a different battery of neuropsychological instruments might have yielded even stronger group and subgroup differences. Third, despite our attempts to equate comparison girls with ADHD girls on the basis of age and ethnicity, such matching was not perfect. This point is salient because of the lack of national norms for most neuropsychological instruments. Fourth, we can make no claim as to the representativeness of the ADHD sample to the population of girls with ADHD across the United States in general. Yet we recruited from multiple sources, obtained a wide socioeconomic and ethnic diversity, and included previously treated as well as untreated participants. Fifth, we did not include the hyperactive–impulsive type of ADHD. Sixth, our division of neuropsychological testing into three, hour-long segments may have prevented fatigue from enhancing ADHD and comparison differences but may not be typical of such test administration.

Nonetheless, our findings of clear EF deficits in girls with ADHD (both combined and inattentive types) and our limited findings of stronger degrees of executive dysfunction in the combined subgroup are bolstered by our large sample size, psychometrically sound tests, careful ascertainment of diagnostic status, and testing of participants while unmedicated. They are tempered by the far-from-perfect individual diagnostic classification of participants on the basis of test performance, which bespeaks the strong variability of neuropsychological functioning in girls with ADHD. Our hope is that these efforts will spur the field towards theoretically rigorous attempts to understand the underlying processes and mechanisms responsible for ADHD in both boys and girls and to provide a sound scientific basis for efforts towards better classification, prediction, and intervention.

References

Achenbach, T. M. (1991). *Manual for the Child Behavior Checklist 4/18 and Profile 1991.* Burlington, VT: University Associates in Psychiatry.

Aman, C. J., Roberts, R J., & Pennington, B. F. (1998). A neuropsychological examination of the underlying deficit in attention–deficit/hyperactivity disorder: Frontal lobe versus right parietal lobe theories. *Developmental Psychology, 34,* 956–969.

American Psychiatric Association. (1987). *Diagnostic and statistical manual of mental disorders* (3rd ed., rev.). Washington, DC: Author.

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

Arnold, L. E. (1996). Sex differences in ADHD: Conference proceedings. *Journal of Abnormal Child Psychology, 24,* 555–569.

Barkley, R. A. (1997a). *ADHD and the nature of self-control.* New York: Guilford Press.

Barkley, R. A. (1997b). Behavioral inhibition, sustained attention, and executive function: Constructing a unified theory of ADHD. *Psychological Bulletin, 121,* 65–94.

Barkley, R. A. (1998). *Attention-deficit/hyperactivity disorder: A handbook for diagnosis and treatment* (2nd ed.). New York: Guilford Press.

Barkley, R. A., Grodzinsky, G., & DuPaul, G. J. (1992). Frontal lobe functions in attention deficit disorder with and without hyperactivity: A review and research report. *Journal of Abnormal Child Psychology, 20,* 163–188.

Bernstein, D. H., & Waber, D. P. (1996). *Developmental scoring system for the Rey–Osterrieth Complex Figure.* Odessa, FL: Psychological Assessment Resources.

Cabeza, R., & Nyberg, L. (1997). Imaging cognition: An empirical review of PET studies with normal subjects. *Journal of Cognitive Neuroscience, 9,* 1–26.

Carte, E. T., Nigg, J. T., & Hinshaw, S. P. (1996). Neuropsychological functioning, motor speed, and language processing in boys with and without ADHD. *Journal of Abnormal Child Psychology, 24,* 481–498.

Casey, B. J., Castellanos, F. X., Giedd, J., & March, W. L. (1997). Implication of right frontostriatal circuitry in response inhibition and attention deficit/hyperactivity disorder. *Journal of the American Academy of Child & Adolescent Psychiatry, 36,* 374–383.

Castellanos, F. X., Marvasti, F. F., Ducharme, J. L., Walter, M. A., Israel, M. E., Krain, A., et al. (2000). Executive function oculomotor tasks in girls with ADHD. *Journal of the American Academy of Child & Adolescent Psychiatry, 39,* 644–650.

Chen, W. J., Faraone, S. V., Biederman, J., & Tsuang, M. T. (1994). Diagnostic accuracy of the Child Behavior Checklist scales for attention-deficit/hyperactivity disorder: A receiver-operating characteristic analysis. *Journal of Consulting & Clinical Psychology, 62,* 1017–1025

Clark, C., Prior, M., & Kinsella, G. J. (2000). Do executive function deficits differentiate between adolescents with ADHD and oppositional defiant/conduct disorder? A neuropsychological study using the Six Elements Test and Hayling Sentence Completion Test. *Journal of Abnormal Child Psychology, 28,* 403–414.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences.* Mahwah, NJ: Erlbaum.

Conners, C. K. (1995). *Conners' Continuous Performance Test computer program: User's manual.* Toronto, Ontario, Canada: Multi-Health Systems.

Conners, C. K. (2000). Attention-deficit/hyperactivity disorder—Historical development and overview. *Journal of Attention Disorders, 3,* 173–191.

Denckla, M. B. (1974). Motor co-ordination in normal children. *Developmental Medicine & Child Neurology, 16,* 729–741.

Denckla, M. B. (1985). Revised neurological evaluation for subtle signs. *Psychopharmacology Bulletin, 21,* 773–800.

Denckla, M. B., & Rudel, R. (1974). Rapid "automatized" naming of pictured objects, colors, letters, and numbers by normal children. *Cortex, 10,* 186–202.

Doyle, A. E., Biederman, J., Seidman, L., Weber, W., & Faraone, S. V. (2000). Diagnostic efficiency of neuropsychological test scores for discriminating boys with and without attention-deficit/hyperactivity disorder. *Journal of Consulting and Clinical Psychology, 68,* 477–488.

Gabrieli, J. D. E., Desmond, J. E., Demb, J. B., Wagner, A. D., Stone, M. V., Vaidya, C. J., & Glover, G. H. (1996). Functional magnetic resonance imaging of semantic memory processes in the frontal lobes. *Psychological Science, 7,* 278–283.

Gaub, M., & Carlson, C. L. (1997). Gender differences in ADHD: A meta-analysis and critical review. *Journal of the American Academy of Child & Adolescent Psychiatry, 36,* 1036–1046.

Gioia, G. A., Isquith, P. K., Guy, S. C., & Kenworthy, L. (2001). *Brief Rating Inventory of Executive Function (BRIEF*). Odessa, FL: Psychological Assessment Resources.

Greenhill, L. L., & Osman, B. B. (Eds.). (2000). *Ritalin: Theory and practice* (2nd ed.). Larchmont, NY: Mary Ann Liebert.

Hinshaw, S. P. (2001). Is the inattentive type of ADHD a distinct disorder? Commentary on Milich, Balentine, and Lynam. *Clinical Psychology: Science & Practice, 8,* 498–501.

Hinshaw, S. P. (2002). Preadolescent girls with attention-deficit/hyperactivity disorder: I. Background characteristics, comorbidity, cognitive and social functioning, and parenting practices. *Journal of Consulting and Clinical Psychology, 70,* 1086–1098.

Hinshaw, S. P., March, J. S., Abikoff, H., Arnold, L. E., Cantwell, D. P., Conners, C. K., et al. (1997). Comprehensive assessment of childhood attention-deficit/hyperactivity disorder in the context of a multisite, multimodal clinical trial. *Journal of Attention Disorders, 1,* 217–234.

Hynd, G. W., Lorys, A. R., Semrud-Clikeman, M., Nieves, N., Huettner, M. I. S., & Lahey, B. B. (1991). Attention-deficit disorder without hyperactivity: A distinct behavioral and neurocognitive syndrome. *Journal of Child Neurology, 6,* 37–43.

Kendler, K. S. (1990). The super-normal control group in psychiatric genetics: Possible artifactual evidence for coaggregation. *Psychiatric Genetics, 1,* 45–53.

Klorman, R., Hazel-Fernandez, L. A., Shaywitz, S. E., Fletcher, J. M., Marchione, K. E., Holahan, J. M., et al. (1999). Executive functioning deficits in attention-deficit/hyperactivity disorder are independent of oppositional–defiant or reading disorder. *Journal of the American Academy of Child & Adolescent Psychiatry, 38,* 1148–1155.

Knights, R. M., & Norwood, J. A. (1979). *A neuropsychological test battery for children*: *Examiner's manual.* Ottawa, Ontario, Canada: Knights Psychological Consultants.

Krikorian, R., & Bartok, J. (1999). Developmental data for the Porteus Maze Test. *Clinical Neuropsychologist, 12,* 305–310.

Lahey, B. B., Applegate, B., McBurnett, K., Biederman, J., Greenhill, L. L., Hynd, G., et al. (1994). *DSM–IV* field trials for attention-deficit/hyperactivity disorder in children and adolescents. *American Journal of Psychiatry, 151,* 1673–1685.

Levin, H. S., Mendelsohn, D., Lilly, M. A., Fletcher, J. M., Culhane, K. A., Chapman, S. B., et al. (1994). Tower of London performance in relation to magnetic resonance imaging following closed head injury in children. *Neuropsychology, 8,* 171–179.

Lezak, M. (1995). *Neuropsychological assessment* (3rd ed.). New York: Oxford University Press.

Lyon, G. R. (1996). Learning disabilities. In E. J. Mash & R. A. Barkley

(Eds.), *Child psychopathology* (pp. 390–435). New York: Guilford Press.

MacLeod, D., & Prior, M. (1996). Attention deficit in adolescents with ADHD and other clinical groups. *Child Neuropsychology, 2,* 1–10.

Marshall, R. M., Hynd, G. W., Handwerk, M. J., & Hall, J. (1997). Academic achievement in ADHD subtypes. *Journal of Learning Disabilities, 30,* 635–642.

McBurnett, K., Pfiffner, L. J., & Frick, P. J. (2001). Symptom properties as a function of ADHD type: An argument for continued study of sluggish cognitive tempo. *Journal of Abnormal Child Psychology, 29,* 207–213.

Mehta, M. A., Owen, A. M., Sahakian, B. J., Mavaddat, N., Pickard, J. D., & Robbins, T. W. (2000). Methylphenidate enhances working memory by modulating discrete frontal and parietal lobe regions in the human brain. *Journal of Neurosciences, 20,* 1–6.

Milich, R., Balentine, A. C., & Lynam, D. R. (2001). ADHD combined type and ADHD predominantly inattentive type are distinct and unrelated disorders. *Clinical Psychology: Science & Practice, 8,* 463–488.

Mitrushina, M. N., Boone, K. B., & D'Elia, L. F. (Eds.). (1999). *Handbook of normative data for neuropsychological assessment.* New York: Oxford University Press.

Mota, V. L., & Schachar, R. J. (2000). Reformulating attention-deficit/hyperactivity disorder according to signal detection theory. *Journal of the American Academy of Child & Adolescent Psychiatry, 39,* 1144–1151.

Naglieri, J. A., & Das, J. P. (1997). *Cognitive Assessment System.* Chicago: Riverside.

Nigg, J. T. (2000). On inhibition/disinhibition in developmental psychopathology: Views from cognitive and personality psychology and a working inhibition taxonomy. *Psychological Bulletin, 126,* 220–246.

Nigg, J. T., Blaskey, L., Huang-Pollock, C., & Rappley, M. D. (2002). Neuropsychological executive functions and ADHD *DSM–IV* subtypes. *Journal of the American Academy of Child & Adolescent Psychiatry, 41,* 59–66.

Nigg, J. T., Hinshaw, S. P., Carte, E. T., & Treuting, J. J. (1998). Neuropsychological correlates of childhood attention-deficit/hyperactivity disorder: Explainable by comorbid disruptive behavior or reading problems? *Journal of Abnormal Psychology, 107,* 468–480.

Oosterlaan, J., Logan, G. D., & Sergeant, J. A. (1998). Response inhibition in AD/HD, CD, comorbid AD/HD+CD, anxious, and control children: A meta-analysis of studies with the Stop Task. *Journal of Child Psychology & Psychiatry, 39,* 411–425.

Osterrieth, P. A. (1944). Le test de copie d'une figure complex [A test of copying of a complex figure]. *Achives de Psychologie, 30,* 206–256.

Pennington, B. F., & Ozonoff, S. (1996). Executive functions and developmental psychopathology. *Journal of Child Psychology & Psychiatry, 37,* 51–87.

Perugini, E. M., Harvey, E. A., Lovejoy, D. W., Sandstrom, K., & Webb, A. H. (2000). The predictive power of combined neuropsychological measures for attention-deficit/hyperactivity disorder in children. *Child Neuropsychology, 6,* 101–114.

Porteus, S. D. (1973). *Porteus Maze Test: Fifty years' application.* Palo Alto, CA: Pacific Books.

Purvis, K. J., & Tannock, R. (2000). Phonological processing, not inhibitory control, differentiates ADHD and reading disability. *Journal of the American Academy of Child & Adolescent Psychiatry, 39,* 485–494.

Rourke, B. P., & Orr, R. R. (1977). Prediction of reading and spelling performance of normal and retarded readers: A 4-year follow-up. *Journal of Abnormal Child Psychology, 5,* 9–20.

Rubia, K., Overmeyer, S., Taylor, E., Brammer, W., Williams, S. C., Simmons, A., & Bullmore, E. T. (1999). Hypofrontality in attention-deficit/hyperactivity disorder during higher-order motor control: A study with functional MRI. *American Journal of Psychiatry, 156,* 891–896.

Sami, N., Carte, E. T., Hinshaw, S. P., & Zupan, B. A. (2002). *Performance of girls with ADHD and comparison girls on the Rey–Osterrieth Complex Figure.* Unpublished manuscript, University of California, Berkeley.

Seidman, L. J., Biederman, J., Faraone, S. V., Weber, W., Mennin, D., & Jones, J. (1997). A pilot study of neuropsychological function in girls with ADHD. *Journal of the American Academy of Child & Adolescent Psychiatry, 36,* 366–373.

Seidman, L. J., Biederman, J., Faraone, S. V., Weber, W., & Oulette, C. (1997). Toward defining a neuropsychology of attention-deficit/hyperactivity disorder: Performance of children and adolescents from a large clinically referred sample. *Journal of Consulting and Clinical Psychology, 65,* 150–160.

Seidman, L. J., Biederman, J., Monuteaux, M. C., Weber, W., & Faraone, S. V. (2000). Neuropsychological functioning in nonreferred siblings of children with attention-deficit/hyperactivity disorder. *Journal of Abnormal Psychology, 109,* 252–265.

Semrud-Clikeman, M., Steingard, R. J., Filipek, P., Biederman, J., Bekken, K., & Renshaw, P. F. (2000). Using MRI to examine brain-behavior relationships in males with attention-deficit disorder with hyperactivity. *Journal of the American Academy of Child & Adolescent Psychiatry, 39,* 477–484.

Sergeant, J., & Scholten, C. A. (1985). On resource strategy limitations in hyperactivity: Cognitive impulsivity reconsidered. *Journal of Child Psychology & Psychiatry, 26,* 97–109.

Shaffer, D., Fisher, P., Lucas, C. P., Dulcan, M. K., & Schwab-Stone, M. E. (2000). NIMH Diagnostic Interview Schedule for Children—Version IV (NIMH DISC–IV): Description, differences from previous versions, and reliability of some common diagnoses. *Journal of the American Academy of Child & Adolescent Psychiatry, 39,* 28–38.

Shaywitz, S. E., Escobar, M. D., Shaywitz, B. A., Fletcher, J. M., & Makuch, R. (1992). Evidence that dyslexia may represent the lower tail of a normal distribution of reading disability. *New England Journal of Medicine, 326,* 145–150.

Sonuga-Barke, E. (1994). Annotation: On dysfunction and function in psychological theories of childhood disorder. *Journal of Child Psychology & Psychiatry, 35,* 801–815.

SPSS, Inc. (1999). *SPSS 10.0 for Windows.* Chicago: Author.

Stokman, C. J., Shafer, S. Q., Shaffer, D., Ng, S., O'Connor, P. A., & Wolff, R. (1986). Assessment of neurological soft signs in adolescents: Reliability studies. *Developmental Medicine & Child Neurology, 28,* 428–439.

Swanson, J. M. (1992). *Assessment and treatment of ADD students.* Irvine, CA: K. C. Press.

Swanson, J. M., McBurnett, K., Wigal, T., & Pfiffner, L. (1993). Effects of stimulant medications on children with attention-deficit disorder: A review of reviews. *Exceptional Children, 60,* 154–161.

Tannock, R. (1998). Attention-deficit/hyperactivity disorder: Advances in cognitive, neurobiological, and genetic research. *Journal of Child Psychology & Psychiatry, 39,* 65–99.

Tranel, D., Anderson, S. W., & Benton, A. (1994). Development of the concept of "executive function" and its relationship to the frontal lobes. In F. Boller & J. Grafman (Eds.), *Handbook of neuropsychology* (Vol. 9, pp. 125–148). New York: Elsevier.

Van der Meere, J. J. (1996). The role of attention. In S. Sandberg (Ed.), *Hyperactivity disorders of childhood* (pp. 111–148). Cambridge, England: Cambridge University Press.

Vitiello, B., Ricciuti, A. J., Stoff, D. M., Behar, D., & Denckla, M. B. (1988). Reliability of subtle (soft) neurological signs in children. *Journal of the American Academy of Child & Adolescent Psychiatry, 28,* 749–753.

Waber, D. P., & Holmes, J. M. (1985). Assessing children's copy productions of the Rey–Osterrieth Complex Figure. *Journal of Clinical & Experimental Neuropsychology, 7,* 264–280.

Wasserman, J. D., & Becker, K. A. (1999). *Advances in intellectual* *assessment of children and adolescents: New research on the Das Naglieri Cognitive Assessment System.* Itasca, IL: Riverside.

Wechsler, D. (1991). *Manual for the Wechsler Intelligence Scale for Children* (3rd ed.). New York: Psychological Corporation/Harcourt Brace.

Wechsler, D. (1992). *Manual for the Wechsler Individual Achievement Test.* New York: Psychological Corporation/Harcourt Brace.

Woodcock, R., & Johnson, M. B. (1989/1990). *Woodcock–Johnson Psychoeducational Battery—Revised.* Denton, TX: DLM Resources.

Zametkin, A. J., Nordahl, T. E., Gross, M., King, A. C., Semple, W. E., & Rumsey, J. (1990). Cerebral glucose metabolism in adults with hyperactivity of childhood onset. *New England Journal of Medicine, 323,* 1361–1366.

Received February 21, 2001
Revision received November 6, 2001
Accepted November 17, 2001 ■