

Neuropsychological Functioning of Girls With Attention-Deficit/Hyperactivity Disorder Followed Prospectively Into Adolescence: Evidence for Continuing Deficits?

Stephen P. Hinshaw
University of California, Berkeley

Estol T. Carte
Kaiser Permanente Medical Group

Catherine Fan, Jonathan S. Jassy, and Elizabeth B. Owens
University of California, Berkeley

The current study prospectively followed girls with attention-deficit/hyperactivity disorder (ADHD), along with a matched comparison sample, 5 years after childhood neuropsychological assessments. Follow-up neuropsychological measures emphasized attentional skills, executive functions, and language abilities. Paralleling childhood findings, the childhood-diagnosed ADHD group displayed moderate to large deficits in executive/attentional performance as well as in rapid naming relative to the comparison group at follow up ($M_{\text{age}} = 14.2$ years). ADHD-inattentive versus ADHD-combined contrasts were nonsignificant and of negligible effect size, even when a refined, sluggish cognitive tempo subgroup of the inattentive type was examined. Although ADHD versus comparison group differences largely withstood statistical control of baseline demographics and comorbidities, control of childhood IQ reduced executive function differences to nonsignificance. Yet when the subset of girls meeting diagnostic criteria for ADHD in adolescence was compared with the remainder of the participants, neuropsychological deficits emerged even with full statistical control. Overall, childhood ADHD in girls portends neuropsychological and executive deficits that persist for at least 5 years.

Keywords: attention-deficit/hyperactivity disorder (ADHD), neuropsychology, executive function, longitudinal research, girls

Neuropsychological measures tap a range of functions, including attention, inhibition, motor speed, and linguistic abilities. The construct of executive function (EF) refers to those neuropsychological skills deemed essential for performance of complex human tasks related to planning, set maintenance, set shifting, interference control, and working memory (e.g., Barkley, 1997; Pennington & Ozonoff, 1996). Considerable evidence exists that, in contrast to nondiagnosed comparison individuals, samples of children and adults with attention-deficit/hyperactivity disorder (ADHD) show significant neuropsychological deficits, particularly those linked with EF (see, for example, Hinshaw, Carte, Sami, Treuting, & Zupan, 2002; Klorman et al., 1999; Nigg, Blaskey, Huang-Pollock, & Rappley, 2002; Seidman et al., 2005; Seidman, Biederman, Faraone, Weber, & Ouellette, 1997; see reviews by Barkley, Grodzinsky, & DuPaul, 1992; Hervey, Epstein, & Curry, 2004;

Seidman et al., 2004; Sergeant, Geurts, & Oosterlaan, 2002). Such deficits have appeared on a number of different tests and have been largely independent of comorbid conditions that accompany ADHD (Hinshaw et al., 2002; Seidman et al., 1997). Because EF deficits are particularly likely to predict continuing academic failure in youth with ADHD (Biederman et al., 2004) and because they may reveal a distinct subset of youth with this disorder (Coghill, Nigg, Rothenberger, Sonuga-Barke, & Tannock, 2005), their identification is important both clinically and conceptually.

Several issues are salient regarding the linkage between EF deficits, on the one hand, and ADHD on the other. First, other developmental conditions are characterized by executive dysfunction (e.g., autistic disorder, conduct problems, obsessive-compulsive disorder), meaning that there is no specific, unique linkage between ADHD and executive deficits (see Pennington & Ozonoff, 1996; Weyandt, 2005). Second, effect sizes of neuropsychological and EF deficits in ADHD samples have ranged from small to large depending on the specific tests used as well as the composition of different samples (e.g., Hinshaw et al., 2002; Klorman et al., 1999; Seidman et al., 2005). Thus, there is no uniformity of executive dysfunction with respect to ADHD. Third, and related to the above, EF dysfunction does not characterize all individuals with this disorder (Biederman et al., 2004; Nigg, Willcutt, Doyle, & Sonuga-Barke, 2005). Thus, predicting ADHD status from tests of executive dysfunction has yielded both false positive and false negative classifications (Doyle, Biederman, Seidman, Weber, & Faraone, 2000; Hinshaw et al., 2002; see Coghill et al., 2005, for elucidation). Yet given the strong desire for objective assessment

Stephen P. Hinshaw, Department of Psychology, University of California, Berkeley; Estol T. Carte, Kaiser Permanente Medical Group, San Rafael, California; Catherine Fan, Jonathan S. Jassy, and Elizabeth B. Owens, Institute of Human Development, University of California, Berkeley.

Work on this project was supported by National Institute of Mental Health Grant R01 MH45064. We thank the participating girls and their families for their continuing involvement in our longitudinal study. We also thank our assessors and coders, without whom the present neuropsychological data could not have been collected or prepared.

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

tools in the evaluation of ADHD plus the need to understand its underlying cognitive substrates, assessment of EF remains a priority research area.

Four major gaps in research on EF deficits and ADHD are evident. First, definitions of EF vary across conceptual models and research investigations, making for a lack of comparability. In this article, we emphasize classic notions of this construct with tasks that emphasize planning, inhibitory control, set maintenance, and set shifting.

Second, as with all aspects of research on ADHD, samples have been dominated by male participants, over and above the actual male:female ratio of approximately 3:1 (Gaub & Carlson, 1997; Gershon, 2002). Accordingly, until recently, relatively little has been known about the neuropsychological function of females with ADHD. Two large-sample investigations have yielded evidence for considerable EF deficits in girls with this condition (Hinshaw et al., 2002; Seidman et al., 2006), with such deficits surviving statistical control of comorbidities, demographic characteristics, and overall cognitive functioning (IQ). Indeed, in their review of extant literature on the comparability of executive, inhibitory, and attentional deficits across boys and girls with ADHD, Seidman et al. (2005) concluded that sex differences do not exist regarding EF deficits in ADHD. Seidman et al. (2006) found that neuropsychological deficits were strongest in girls with ADHD and comorbid learning disorders. Overall, because large, diverse samples of girls with ADHD are rare in the field, examination of neuropsychological deficits in female samples is a priority.

Third, understanding the course of neuropsychological deficits (EF deficits in particular) is crucial for those interested in the developmental trajectory of ADHD. Whereas developmental changes exist in core symptomatology, such that hyperactive/impulsive (HI) symptoms decline at a greater rate than inattentive symptoms by adolescence (e.g., Hart, Lahey, Loeber, Applegate, & Frick, 1995), relatively few data are available on the stability of EF deficits. Although cross-sectional investigations reveal no significant differences in EF deficits between children and adolescents with ADHD (see Seidman et al., 2005) and although reviews of the adult literature show that such deficits are present in adults with this disorder (Hervey et al., 2004; Woods, Lovejoy, & Ball, 2002), secular trends or cohort effects constrain interpretations of such nonprospective investigations.

Prospective longitudinal findings include those of Hopkins, Perlman, Hechtman, and Weiss (1979), who found evidence for continuing deficits in a hyperactive sample (relative to comparison individuals) through late adolescence on tests related to impulsivity and set shifting. By young adulthood, however, such differences had diminished (Hechtman, Weiss, Perlman, & Amsel, 1984). Similarly, the New York study of Klein and colleagues found limited evidence for attentional and/or executive deficits by age 18 (see review in Mannuzza & Klein, 1999). Recently, Drechsler, Brandeis, Földényi, Imhof, and Steinhausen (2005) found mixed evidence for the persistence of alertness and inhibitory deficits from late childhood through early adolescence. Finally, in a 13–14-year follow up of a well-characterized, large sample (composed of over 90% men), Fischer, Barkley, Smallish, and Fletcher (2005) found that at least some neuropsychological deficits persisted into young adulthood for the childhood-defined ADHD sample, although the most salient problems at follow up

were shown for those individuals whose ADHD-related symptomatology had persisted across the longitudinal interval from childhood through adulthood. For continuous performance test (CPT) variables of omission and commission errors, the ADHD sample's scores improved from childhood through adulthood but so did the comparison group's scores, meaning that deficits remained in place. Overall, it is still relatively unknown whether neuropsychological deficits in general or executive dysfunctions in particular attenuate over time for youth with ADHD or remain salient, especially in girls.

Fourth, almost no prospective follow up has occurred of youth with the inattentive type of ADHD—whether on cognitive/executive measures or behavioral and impairment-related outcomes in general (Mannuzza & Klein, 1999). Given claims that this variant of ADHD represents a qualitatively distinct condition (Milich, Balentine, & Lynam, 2001), such follow up is a priority, particularly with respect to objective outcomes, which are likely to be less biased than reports from adult informants.

Our research team collected extensive neuropsychological information on a large, well-characterized, and diverse sample of preadolescent girls with ADHD ($n = 140$) plus a matched comparison sample ($n = 88$) ages 6–12 years (see Hinshaw et al., 2002). In childhood, (a) the ADHD group showed clear neuropsychological deficits relative to the comparison group, with effect sizes ranging from medium to large; (b) these deficits largely withstood stringent statistical control of demographic variables, comorbid disorders (disruptive behavior disorders, internalizing disorders, reading disorder), and full-scale IQ (FSIQ); (c) the largest ADHD–comparison differences were revealed on tests tapping EF, including the Rey–Osterrieth Complex Figure (ROCF; see Sami, Carte, Hinshaw, & Zupan, 2003); and (d) ADHD-combined (ADHD-C) versus ADHD-inattentive (ADHD-I) type differences were small in effect size and rarely significant. As expected, individual classification of ADHD diagnostic status from the neuropsychological tests was imperfect, with a preponderance of false positive predictions (for related data, see Doyle et al., 2000).

Regarding our prospective, 5-year follow up of this sample, we ask the following: (a) Do preadolescent girls with ADHD continue to show neuropsychological and executive dysfunction in adolescence, relative to a matched comparison group? (b) Are such deficits robust to statistical control of age, socioeconomic status, comorbidities, and general intelligence yielded from baseline measures during childhood? (c) Do the girls who meet diagnostic criteria for ADHD at follow up show particularly strong EF deficits in adolescence? Our main hypotheses are that neuropsychological (attentional, rapid naming, executive) deficits will be maintained (but attenuated) during adolescence; that these will be robust to statistical control of demographics, comorbidities, and IQ (see Mahone et al., 2002, regarding IQ–neuropsychological performance linkages); and that girls with adolescent ADHD symptomatology will show the largest degree of neuropsychological/executive dysfunction at follow up (Fischer et al., 2005). We also predict that ADHD-C and ADHD-I differences will continue to be small at follow up. Although parallel forms of baseline measures were used, the only measure repeated intact was a test of continuous performance.

Method

Overview of Procedure

As described in Hinshaw, Owens, Sami, and Fargeon (2006), during the school years of 2001–2002, 2002–2003, and 2003–2004, we conducted prospective follow-up investigations of the girls initially investigated in Hinshaw (2002) and Hinshaw et al. (2002). These participants had participated in research summer programs in 1997, 1998, and 1999, respectively. During precamp assessments and the summer programs, we had conducted nearly 3 hours of neuropsychological testing performed in three separate sessions to prevent fatigue and to assess test–retest reliability for certain measures.

For the follow-up evaluations, we selected key measures for a 30-min neuropsychological battery. On this assessment day, any girl taking stimulant medication was asked to participate medication free. Data from caregivers indicated that a small percentage (approximately 10%) of the follow-up ADHD sample may have actually been medicated on the assessment day, but because a number of them took short acting stimulants in the morning and testing was routinely performed during afternoons, any medication benefit may have attenuated. To the extent that stimulants yield improvements in EF, the present results may reflect an underestimate of dysfunction. Testing was conducted by highly trained graduate students in clinical psychology or by bachelor's level research assistants who had undergone extensive training. These assessors were kept blind to the participant's diagnostic status; they comprised, almost exclusively, individuals who had not served as staff for the summer programs. Overall, whereas Hinshaw et al. (2006) presented data on several domains of psychiatric symptoms, on academic and social impairment, and on service use, the current data pertain to the persistence of neuropsychological deficits.

Participants

Hinshaw (2002) provided a complete description of the multi-gated recruitment, screening, and diagnostic procedures used to ascertain the sample of 140 girls with ADHD and 88 age- and ethnicity-matched comparison girls who participated in the baseline neuropsychological assessments and research summer programs. Girls with ADHD were recruited through pediatricians, mental health centers, schools, and direct advertisement, and comparison girls were recruited through pediatricians, community centers, and direct advertisement. Preliminary rating scale criteria were intentionally set with liberal, sex-specific thresholds to prevent premature exclusion of potentially eligible girls, but final study entry depended on the participants' having met full criteria for ADHD through the parent-administered Diagnostic Interview Schedule for Children (4th ed.; DISC-IV; Shaffer, Fisher, Lucas, Dulcan, & Schwab-Stone, 2000). Common comorbidities (oppositional defiant disorder, conduct disorder, anxiety disorders, depression, learning disorders) were allowed. Comparison girls, who were matched at a group level with the ADHD participants and did not significantly differ from the ADHD sample with respect to age or ethnicity, could not meet diagnostic criteria for ADHD via either adult ratings or structured interview criteria. Exclusion criteria were mental retardation, evidence of psychosis or overt

neurological disorder, lack of English spoken in the home, and medical problems prohibiting summer camp participation.

Regarding the types of ADHD sampled (ADHD-C, ADHD-I), Hinshaw (2002) described diagnostic procedures that were based largely on the DISC-IV but also included Swanson, Nolan, and Pelham (SNAP) ratings (Swanson, 1992; this is a widely used adult informant rating scale of ADHD symptomatology) and staff judgments in borderline cases. Each of the 18 *Diagnostic and Statistical Manual of Mental Disorders* (4th ed.; American Psychiatric Association, 1994) ADHD symptoms was considered present if endorsed on the DISC-IV or if the mother or teacher rated it as a 2 (*pretty much*) or 3 (*very much*) on the SNAP. Girls with at least six inattentive and six HI symptoms (with at least four in each domain based on the DISC-IV; see Hinshaw et al., 1997) were designated ADHD-C; girls with at least six inattentive (with at least four based on the DISC-IV) but fewer than six HI symptoms were designated ADHD-I; and girls with fewer than six inattentive and six HI symptoms were designated as not having ADHD. To preserve statistical power for the ADHD-C versus ADHD-I contrast, we did not include girls with the predominantly HI type.

At baseline, the girls spanned the ages of 6–12 years. The sample was ethnically diverse (53% White, 27% African American, 11% Latina, 9% Asian American). The clinic and summer camp procedures yielded multi-informant, multimethod data on both symptoms and a wide range of domains of functional impairment (Hinshaw, 2002).

The follow-up evaluations were performed on 209 of the 228 participants (92%), who ranged in age from 11.3 to 18.2 years ($M = 14.2$ years). Reasons for nonparticipation included (a) family lost to all tracking efforts ($n = 4$), (b) refusal to participate ($n = 5$), and (c) family contacted but scheduling of assessments not possible ($n = 10$). Comparisons of the retained sample versus those lost to attrition revealed that, for 29 of 31 demographic, diagnostic, and symptom variables from baseline, differences were not statistically significant. The only two variables for which significant differences emerged were single-parent status and teacher-reported internalizing symptomatology (for each, baseline rates were higher in the nonretained sample). In short, the retained sample appears representative of the total sample (Hinshaw et al., 2006). Additionally, some assessments occurred via home visits or telephone interviews ($n = 7$), precluding full neuropsychological assessment; some measures were missing because of fatigue or refusal; and in other instances (i.e., Conners' CPT; Conners, 1995), computer failures occurred. Hence, the sample size for our present battery ranges from 186 to 200. For secondary analyses that involve follow-up diagnostic status as the independent variable, we readministered the DISC-IV at the follow-up evaluations (Hinshaw et al., 2006). See Table 1 for descriptive information about the 202 girls who constituted the overall sample for the follow-up neuropsychological battery.

Cognitive and Neuropsychological Measures

The priority was placed on well-established and well-validated neuropsychological tests, most of which were parallel forms of those used at baseline (see Hinshaw et al., 2002) and could be administered in approximately 30 min, given the need to sample multiple domains of functioning in our follow-up battery. The only

Table 1
Demographic Variables and Medication Use for ADHD Subgroups at Follow Up by Baseline Diagnostic Status

| Variable | Comparison (<i>n</i> = 80) | | Inattentive (<i>n</i> = 39) | | Combined (<i>n</i> = 83) | | <i>p</i> ^a |
|---|-----------------------------|-----------|------------------------------|-----------|---------------------------|-----------|-----------------------|
| | <i>M</i> | <i>SD</i> | <i>M</i> | <i>SD</i> | <i>M</i> | <i>SD</i> | |
| Age (months) | 168.3 _a | 19.3 | 175.4 _a | 20.6 | 170.5 _a | 20.5 | .195 |
| Total annual family income ^b | 8.4 _a | 2.8 | 8.3 _a | 2.9 | 7.6 _a | 3.3 | .190 |
| Maternal education ^c | 5.0 _a | 1.0 | 4.7 _a | 0.9 | 4.7 _a | 0.9 | .226 |
| Caucasian (%) | 46.3 _a | | 59.9 _a | | 57.8 _a | | .062 |
| Two-parent household (%) | 67.5 _a | | 73.3 _a | | 56.1 _a | | .140 |
| Any psychotropic medication (%) | 2.5 _a | | 46.2 _b | | 58.8 _b | | .000 |
| Stimulant medication (%) | 0.0 _a | | 28.2 _b | | 44.6 _b | | .027 |

Note. 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.

^a Significance = one-way analysis of variance for continuous variables; Pearson's chi-square statistic for categorical variables.

^b For total annual family income, $1 \leq \$10,000$; $11 \geq \$100,001$.

^c For maternal education, 1 = less than eighth grade; 6 = advanced or professional degree.

two measures that were exact replications of the baseline battery were the Conners' CPT and the Underlining Test (UL; Rourke & Orr, 1977; Rourke & Gates, 1980). Furthermore, to avoid problems of multiple statistical tests, we selected a priori only one or two dependent measures from each test, constituting those with the most established psychometric properties and clinical utility.

Wechsler Intelligence Scale for Children (WISC-III): Digit Span (Wechsler, 1991). This is a widely used measure of auditory working memory that requires participants to immediately recall digit sequences of increasing length either in their original presentation order (Digits Forward) or in their reverse presentation order (Digits Backward). Because of the conceptual importance of rehearsal and interference control for the latter, we analyze scores separately herein. Working memory is considered to be an important component or correlate of EF (see Scheres et al., 2004; Willcutt, Pennington, Olson, Chhabildas, & Hulslander, 2005) and appears to involve both frontostriatal and cerebellar brain regions (Martinussen, Hayden, Hogg-Johnson, & Tannock, 2005). Split-half reliabilities average .85 across the age span of the standardization sample (Wechsler, 1991). We analyzed standard scores ($M = 10$, $SD = 3$) so that mean levels and effect sizes would be clinically interpretable. Note that Digit Span is a supplemental test of the WISC-III, so that FSIQ scores (used as a covariate in our analyses) are independent of Digit Span scores.

Taylor Complex Figure Test (TCFT; Taylor, 1969). We selected this measure as a parallel form of the ROCF, which at baseline had been our most sensitive neuropsychological/executive test for discriminating the ADHD groups from the comparison sample (Hinshaw, 2002; Sami et al., 2003). The ROCF is used to evaluate planning, perceptual organization, and graphomotor abilities (Lezak, 1983; Spreen & Strauss, 1998). The TCFT is the only major alternative to the ROCF in a test-retest situation (Helmes, 2000). We administered only the copy condition of the TCFT because, at baseline, the delayed recall condition yielded weak evidence for ADHD versus comparison differences (Sami et al., 2003).

Our primary dependent measure was the error proportion score, calculated as the number of segments drawn incorrectly (errors) divided by the sum of all segments drawn (correct plus incorrect). The current scoring system was parallel to the one used at baseline

that had been developed from our modifications of the procedures of Bernstein and Waber (1996), as described in Sami et al. (2003). Drawings were scored by a group of three extensively trained, independent raters who were unaware of diagnostic status. For the error proportion score, the intraclass correlations between pairs of the three scorers were .77, .83, and .94, with a mean of .84 on a subsample of 60 drawings.

Conners' CPT (Conners, 1995). In this task of visual attention and set shifting, a target letter is presented on a computer screen and defined as any letter except for X. Participants are instructed to press the computer spacebar key on the presentation of each target letter but to inhibit response to all Xs (i.e., nontargets). This CPT therefore taps the EF of response inhibition by requiring occasional inhibitory responses in the midst of frequent responses to targets. Trials are presented in six blocks, with the interstimulus interval ranging from 1 s, 2 s, to 4 s within each block; stimulus display time is 250 ms. Parallel to the baseline analyses, we chose omission errors (percentage failures to respond to target stimuli out of the total number of targets presented) and commission errors (percentage keypresses for nontargets out of the total number of nontargets presented) as the primary dependent measures. Conners (1995) provided criterion-related validity data for these scores in terms of known-groups differentiation and sensitivity to treatment response.

Rapid Automated Naming Test (RAN; Denckla & Rudel, 1974). Here, children rapidly name repeated items: automatically processed digits (digits), pictured stimuli (objects), and digit-letter pairs (Subtest 1 of global-local; see Hinshaw et al., 2002). We administered the digits subtest as an orienting task; objects and digit-letter pairs were the two subtests used in our analyses. The objects subtest requires effortful semantic processing and retrieval; this type of task appears to require left-prefrontal brain regions (Gabrieli, Desmond, Domb, Wagner, & Stone, 1996). For boys with ADHD, this was the subtest revealing the largest differentiation from comparison boys (Nigg, Hinshaw, Carte, & Treuting, 1998). The digit-letter subtest requires both set shifting and inhibition of a prepotent response for participants to switch between numbers and letters. Our score for each test was an accuracy rate, defined for objects 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. For digit–letter, it was the total number of stimuli named minus the sum of incorrect recognitions plus omissions, with that quantity divided by 60 (i.e., the time in seconds to complete the task). Because the correlation between objects and digit–letter scores was substantial ($r = .67$), we summed these subtests as a single dependent measure.

Underlining Test (Rourke & Gates, 1980; Rourke & Orr, 1977). This is a target cancellation task that measures rapid and accurate visual discrimination, requiring the underlining of target stimuli embedded among distractors at a target:distractor ratio of 1:5. The subtest requiring simple detection of the numeral 4 was used to orient participants. We used three additional subtests: (a) sequenced compound gestalt shapes (gestalt figure; original subtest No. 4), (b) the sequenced consonants *fsbm* (original subtest No. 9), and (c) the sequenced letters *spot* (original subtest No. 11). Whereas the spot subtest is related to sight word recognition, the gestalt figure subtest and the *fsbm* subtest require controlled processing (Posner, 1988) and are unrelated to reading skill (Rourke & Orr, 1977); thus, we selected these two subtests for our analyses. In prior research, the *fsbm* subtest optimally discriminated boys with ADHD from comparison boys (Carte et al., 1996; Nigg et al., 1998). Anterior cingulate prefrontal regions are involved in such tasks (Cabeza & Nyberg, 1997). Because errors (rather than slow speed) characterize participants with ADHD (MacLeod & Prior, 1996), our dependent variable is the single score of correctly identified targets minus misidentified targets and nontargets. Because of only a modest correlation ($r = .31$) between the gestalt and *fsbm* subtests, we examine these as separate dependent measures.

Overall, on the basis of conceptual and empirical formulations (Tranel, Anderson, & Benton, 1994), our measures of EF include digits forward and digits backward from Digit Span, omission and commission errors from the CPT (which we conceptualize as indicators of inattention and impulsivity, respectively), the error proportion score from the TCFT, and number correct minus number incorrect from the UL gestalt and *fsbm* subtests. Our language-related measure is the RAN accuracy score. We reverse scored the CPT and TCFT so that higher scores reflect better performance.

Covariates

To ascertain whether neuropsychological and EF impairments in adolescence are related specifically to the girls' original ADHD status rather than variables associated with ADHD, we repeated our core analyses with stringent statistical control. The first set of covariates includes demographic information (family income and maternal education; see Hinshaw, 2002, for details); participant age (given the 6–7-year age span across the sample); and additional disorders (which constitute comorbidities for the girls with ADHD), dummy coded as 1 versus 0 for the presence versus absence of oppositional defiant disorder or conduct disorder, 1 versus 0 for the presence versus absence of anxiety or mood disorders, and 1 versus 0 for the presence versus absence of reading disorder (see Hinshaw, 2002, for details). In a second set of analyses, we additionally covaried the participant's FSIQ, measured at baseline with the WISC-III. Note that Barkley (1997) recommended conducting neuropsychological analyses within ADHD samples both with and without statistical control of IQ.

Data Analytic Plan

All statistical analyses were performed with SPSS for Windows, Version 12. Because cohort effects could complicate interpretation of findings, we performed an initial set of analyses of variance (ANOVAs) with the independent variable of initial year of participation: 1997, 1998, or 1999. For six of the eight primary dependent measures, there were no effects of year of participation. On the RAN, girls from the 1999 program scored worse than those from 1997 or 1998; on the TCFT, the 1999 sample was worse than the 1997 sample. Because such effects are likely to be of minor importance and because year of participation interacted significantly with ADHD diagnostic group for only one outcome measure (UL gestalt), we include all three cohorts together in our analyses.

Note that correlations among the dependent measures averaged .24 and ranged from .07 to .50, with the largest correlation occurring between Digits Forward and Digits Backward from the WISC-III Digit Span. Our first primary analysis was a multivariate analysis of variance (MANOVA) across the eight dependent measures, with baseline diagnostic status (ADHD-C, ADHD-I, comparison) as the three levels of the independent variable. A significant MANOVA ($\alpha = .05$) afforded the interpretation of follow-up univariate patterns of subgroup differences examined via ANOVAs for each measure along with Tukey's post hoc comparisons of each subgroup contrast. Effect sizes are emphasized, calculated as Cohen's d , with the difference between subtype means as the numerator and the pooled standard deviation as the denominator (Cohen, 1988, designated $d = .2$ as a small effect, $.5$ as medium, and $.8$ or above as large). Next, we completed two multivariate analyses of covariance (MANCOVAs), first partialing the demographic/comorbidity variables and then adding IQ as a covariate. A significant MANCOVA is interpreted as signifying a specific association between childhood ADHD status and outcomes of interest—that is, ADHD status yields significant prediction of adolescent neuropsychological impairment despite stringent statistical control of important covariates. When these analyses were significant, we followed them with analyses of covariance and post hoc pairwise contrasts. For the two measures in our battery that were identical at baseline and follow up (CPT omissions and commissions), repeated measures analyses afforded examination of the tendencies toward improvement or decrement across time for all subgroups. Finally, primary analyses were repeated with diagnostic status at follow up (ADHD-C, ADHD-I, and no ADHD) as the independent variable. Overall, we had good statistical power (over .7) to ascertain group differences of small to medium size between participants with ADHD and comparison girls and adequate power to ascertain group differences of medium size between the ADHD-C and ADHD-I groups (Faul & Erdfelder, 1992).

Results

The overall MANOVA yielded a statistically significant finding for the effect of baseline diagnostic status on neuropsychological performance, $F(16, 342) = 2.76, p < .001$. Table 2 presents the results of the eight univariate ANOVAs. All eight of the outcome measures yielded significant results. For six dependent measures (Digits Forward, Digits Backward, CPT omissions, TCFT, UL gestalt, UL *fsbm*), both ADHD types scored worse than the com-

Table 2
Neuropsychological Performance for ADHD Subgroups at Follow Up by Baseline Diagnostic Status

| Measure | Comparison | | | Inattentive | | | Combined | | | F | p ^a |
|--|------------|-------------------|------|-------------|---------------------|------|----------|-------------------|------|------|----------------|
| | n | M | SD | n | M | SD | n | M | SD | | |
| WISC-III Digits Forward | 80 | 9.8 _a | 2.1 | 38 | 8.8 _b | 2.2 | 82 | 8.3 _b | 2.2 | 9.78 | .000 |
| WISC-III Digits Backward | 80 | 6.3 _a | 1.8 | 38 | 5.5 _b | 1.9 | 82 | 5.1 _b | 2.0 | 9.34 | .000 |
| CPT percent omissions (rev) ^b | 78 | 97.5 _a | 4.6 | 36 | 93.7 _b | 10.6 | 72 | 94.0 _b | 8.7 | 4.84 | .009 |
| CPT percent commissions (rev) ^b | 78 | 64.0 _a | 18.7 | 36 | 53.6 _a | 25.9 | 72 | 56.6 _a | 23.0 | 3.56 | .030 |
| TCFT error proportion score (rev) ^b | 77 | 0.83 _a | 0.09 | 39 | 0.77 _b | 0.10 | 80 | 0.77 _b | 0.13 | 8.22 | .000 |
| RAN correct minus incorrect | 76 | 1.9 _a | 0.4 | 40 | 1.8 _{a, b} | 0.4 | 81 | 1.7 _b | 0.4 | 8.22 | .000 |
| UL gestalt correct minus incorrect | 79 | 20.6 _a | 7.3 | 39 | 15.7 _b | 7.2 | 82 | 17.2 _b | 7.9 | 6.80 | .001 |
| UL fsbm correct minus incorrect | 79 | 14.6 _a | 4.1 | 39 | 12.3 _b | 3.7 | 82 | 11.9 _b | 4.4 | 8.83 | .000 |

Note. In a given row, means with different subscripts differ significantly on the basis of Tukey's post hoc comparisons. 0 = comparison; 1 = inattentive; 2 = combined; WISC-III = Wechsler Intelligence Scale for Children (3rd ed.); CPT = Continuous Performance Test; TCFT = Taylor Complex Figure Test; RAN = Rapid Automatized Naming Test; UL = Underlining Test.

^a Significance based on one-way analyses of variance following significant omnibus multivariate analyses of variance.

^b Reverse scored so that for all variables, higher means indicate better performance.

parison group but did not differ themselves (see contrasts in Table 2). For RAN, the comparison girls performed better than girls with ADHD-C, but the ADHD-I group did not differ from either group. For CPT commissions, no post hoc contrasts were significant. Table 3 (left column) reveals that, for these measures of attentional and executive functioning, effect sizes for the contrasts of (a) ADHD-C versus comparison and (b) ADHD-I versus comparison were medium in magnitude (*d* ranging from .33 to .66). Effect sizes for ADHD-C versus ADHD-I contrasts (which were never significant, as just noted) were not even small in most instances.

When the first set of covariates (demographics, comorbidity) was included, the MANCOVA was just at the traditional level of statistical significance, $F(16, 322) = 1.67, p = .05$. Univariate analyses of covariance revealed that four dependent measures continued to show significance: Digits Forward, $F(2, 195) = 3.74, p < .05$; TCFT error proportion, $F(2, 191) = 3.62, p < .05$; UL gestalt, $F(2, 195) = 6.15, p < .01$; and UL fsbm, $F(2, 195) = 5.09, p < .01$. Effect sizes for the adjusted means regarding ADHD-C versus comparison and ADHD-I versus comparison contrasts decreased slightly for Digits Forward (.35-.52), and TCFT (.45-.50)

but were maintained for the UL subtests (.42-.72; see Table 3). However, when the MANCOVA was reconducted with FSIQ in the set of covariates, the results were now nonsignificant, $F(16, 318) = 1.09, p = .36$. Thus, control of baseline IQ eliminated the effect of childhood ADHD on follow-up neuropsychological performance.

For the two automated measures that were identical to those administered at baseline, CPT omissions and commissions, we conducted a repeated measures MANOVA, with the independent variables of baseline diagnostic status (ADHD-C, ADHD-I, and comparison) and time (baseline, follow up). These revealed significant effects of time, diagnostic status, and their interaction. In separate repeated measures ANOVAs examining these effects on omissions and commissions, there was a main effect of time on omissions, $F(1, 177) = 12.88, p < .001$, signifying that all participants improved, on average, from baseline to follow up. As expected, the main effect of diagnostic status was significant as well, $F(2, 177) = 7.92, p < .01$, indicating that the girls with ADHD performed worse than the comparison girls, reflecting the primary analyses. The Diagnostic Status \times Time interaction was

Table 3
Neuropsychological Performance Effect Sizes at Follow Up by Baseline Diagnostic Status

| Measure | No covariates | | | Demographic and comorbidity set ^a | | | Full set ^b | | |
|--|---------------|--------|--------|--|--------|--------|-----------------------|--------|--------|
| | ES 0-1 | ES 0-2 | ES 1-2 | ES 0-1 | ES 0-2 | ES 1-2 | ES 0-1 | ES 0-2 | ES 1-2 |
| WISC-III Digits Forward | .44 | .66 | .22 | .35 | .52 | .16 | .25 | .40 | .15 |
| WISC-III Digits Backward | .42 | .65 | .22 | .26 | .38 | .12 | .07 | .19 | .13 |
| CPT percent omissions (rev) ^c | .48 | .44 | .04 | .32 | .16 | .16 | .19 | .04 | .15 |
| CPT percent commissions (rev) ^c | .47 | .33 | .13 | .45 | .22 | .23 | .45 | .18 | .27 |
| TCFT error proportion score (rev) ^c | .52 | .59 | .07 | .50 | .45 | .05 | .31 | .26 | .04 |
| RAN correct minus incorrect | .34 | .62 | .28 | .21 | .31 | .10 | .20 | .29 | .09 |
| UL gestalt correct minus incorrect | .62 | .44 | .18 | .72 | .42 | .30 | .43 | .12 | .31 |
| UL fsbm correct minus incorrect | .52 | .61 | .09 | .53 | .54 | .01 | .45 | .46 | .01 |

Note. ES = effect size, calculated as Cohen's *d*; 0 = comparison; 1 = inattentive; 2 = combined; WISC-III = Wechsler Intelligence Scale for Children (3rd ed.); CPT = Continuous Performance Test; TCFT = Taylor Complex Figure Test; RAN = Rapid Automatized Naming Test; UL = Underlining Test.

^a Covariates included baseline family income, child age, reading disorder, disruptive disorder, and internalizing disorder.

^b Covariates included demographic and comorbidity set plus child full scale IQ.

^c Reverse scored so that for all variables, higher means indicate better performance.

not significant $F(2, 177) = 2.49, p < .10$, indicating the pattern of change across time was similar across the diagnostic groups.

For commissions, however, not only was the main effect of time significant, $F(1, 177) = 41.70, p < .001$, again signifying that all groups improved from baseline to follow up, but the Diagnostic Status \times Time interaction was also significant, $F(2, 177) = 5.30, p < .01$. [The diagnostic status effect was not significant, $F(2, 177) = 2.32, p = .102$.] A plot of this interaction revealed that whereas the girls with ADHD-I were relatively stable in their performance across time, the ADHD-C and comparison groups both showed notable improvement.

As noted in the data analytic plan, we reconducted our primary analyses with the independent variable of follow-up diagnostic status. Here, on the basis of diagnoses that were determined in adolescence, we found that 40 girls met criteria for ADHD-C, 47 for ADHD-I, and 115 did not meet criteria for ADHD. In other words, as described in detail in Hinshaw et al. (2006), over half of the girls with ADHD-C at baseline no longer met diagnostic criteria for this subtype at the follow-up evaluations, compared with approximately one third of the girls with ADHD-I at baseline. In other words, as the girls aged, a significant percentage failed to meet official criteria for ADHD, although clear impairments remained in all functional domains investigated (Hinshaw et al., 2006).

With this follow-up designation as the new independent variable, the MANOVA was significant, $F(16, 342) = 3.15, p < .001$. Seven of the eight follow-up ANOVAs were significant: Digits Forward, Digits Backward, CPT omissions, TCFT, UL gestalt, UL fsbm, and RAN (see Table 4). For four of these (Digits Forward, RAN, UL gestalt, UL fsbm), contrasts revealed that each ADHD type was significantly more impaired than the comparison girls were. For CPT omissions, the ADHD-I type was intermediate between the ADHD-C and comparison groups but not significantly different from either. For TCFT, the comparison and ADHD-I groups did not differ, but the ADHD-C girls performed worse than both other groups. For Digits Backward, no contrasts were significant. Effect sizes associated with significant contrasts ranged from medium to large.

With inclusion of covariates, both the demographic plus comorbidity set, $F(16, 322) = 2.70, p < .001$, and the set that also included FSIQ, $F(16, 318) = 2.36, p < .01$, were also significant, with the four

outcome measures (Digits Forward, TCFT, RAN, and UL gestalt) maintaining significance with inclusion of all covariates. As shown in Table 5, the comparison versus ADHD-C contrasts remained medium in effect size. A few of the comparison versus ADHD-I contrasts shrank to small effect sizes with inclusion of covariates, whereas others actually increased (e.g., UL gestalt).

Discussion

Our chief aim was to determine whether girls with carefully ascertained ADHD in childhood would continue to display neuropsychological deficits (and EF deficits in particular) 5 years later. By using well-established measures of working memory, planning, set maintenance, and set shifting—considered as EF in the literature—as well as indicators of attentional processing, impulse control, and rapid naming, we found that the girls diagnosed with ADHD during childhood continued to display significant, medium effect deficits, in contrast with our matched comparison sample, in our 5-year prospective assessments. ADHD-C versus ADHD-I differences did not emerge for any measure, however, with small effect sizes for these contrasts. Many EF differences between girls with ADHD and the comparison girls survived stringent statistical control of age, socioeconomic/demographic indicators, and comorbid diagnostic status, measured at baseline.

For CPT omissions and commissions, objective measures that were repeated identically at baseline and follow up, main effects were significant, indicating better performance for all girls, on average, during adolescence. For commissions, the comparison girls and those with ADHD-C showed relatively greater improvement than the girls with ADHD-I. Regarding the overall analyses, controlling IQ scores from childhood eliminated the ADHD-comparison differences; however, when we performed a secondary analysis examining the cognitive patterns of those girls who met diagnostic criteria for ADHD during the follow-up assessments, strong EF differences emerged between this subset and those who did not meet criteria, and many differences survived full statistical control, even of IQ. Overall, girls with ADHD continue to display noteworthy neuropsychological deficits in adolescence, 5 years

Table 4
Neuropsychological Performance for ADHD Subgroups at Follow Up by Follow Up Diagnostic Status

| Measure | No ADHD | | | Inattentive | | | Combined | | | F | p ^a |
|--|---------|-------------------|------|-------------|----------------------|------|----------|-------------------|------|-------|----------------|
| | n | M | SD | n | M | SD | n | M | SD | | |
| WISC-III Digits Forward | 115 | 9.7 _a | 2.1 | 47 | 8.2 _b | 2.3 | 38 | 7.9 _b | 2.0 | 14.33 | .000 |
| WISC-III Digits Backward | 115 | 6.0 _a | 1.9 | 47 | 5.3 _a | 1.9 | 38 | 5.2 _a | 2.1 | 3.48 | .033 |
| CPT percent omissions (rev) ^b | 107 | 96.5 _a | 6.2 | 46 | 94.6 _{a, b} | 9.7 | 33 | 92.7 _b | 9.4 | 3.30 | .039 |
| CPT percent commissions (rev) ^b | 107 | 62.0 _a | 20.7 | 46 | 54.5 _a | 24.0 | 33 | 56.1 _a | 23.6 | 2.23 | .110 |
| TCFT error proportion score (rev) ^b | 111 | 0.82 _a | 0.09 | 47 | 0.78 _a | 0.10 | 38 | 0.72 _b | 0.14 | 10.78 | .000 |
| RAN correct minus incorrect | 112 | 1.9 _a | 0.4 | 46 | 1.7 _b | 0.4 | 39 | 1.6 _b | 0.4 | 7.07 | .001 |
| UL gestalt correct minus incorrect | 114 | 20.3 _a | 7.8 | 47 | 15.4 _b | 7.1 | 39 | 15.7 _b | 6.2 | 10.21 | .000 |
| UL fsbm correct minus incorrect | 114 | 14.1 _a | 3.8 | 47 | 11.8 _b | 4.8 | 39 | 11.5 _b | 4.3 | 8.19 | .000 |

Note. In a given row, means with different subscripts differ significantly on the basis of Tukey's post hoc comparisons. 0 = comparison; 1 = inattentive; 2 = combined; WISC-III = Wechsler Intelligence Scale for Children (3rd ed.); CPT = Continuous Performance Test; TCFT = Taylor Complex Figure Test; RAN = Rapid Automatized Naming Test; UL = Underlining Test.

^a Significance based on one-way analyses of variance following significant omnibus multivariate analyses of variance.

^b Reverse scored so that for all variables, higher means indicate better performance.

Table 5
Neuropsychological Performance Effect Sizes at Follow Up by Follow Up Diagnostic Status

| Measure | No covariates | | | Demographic and comorbidity set ^a | | | Full set ^b | | |
|--|---------------|--------|--------|--|--------|--------|-----------------------|--------|--------|
| | ES 0-1 | ES 0-2 | ES 1-2 | ES 0-1 | ES 0-2 | ES 1-2 | ES 0-1 | ES 0-2 | ES 1-2 |
| WISC-III Digits Forward | 0.66 | 0.79 | 0.13 | 0.54 | 0.71 | 0.17 | 0.50 | 0.66 | 0.16 |
| WISC-III Digits Backward | 0.34 | 0.41 | 0.07 | 0.29 | 0.41 | 0.12 | 0.20 | 0.28 | 0.08 |
| CPT percent omissions (rev) ^c | 0.24 | 0.48 | 0.24 | 0.19 | 0.52 | 0.33 | 0.14 | 0.42 | 0.28 |
| CPT percent commissions (rev) ^c | 0.34 | 0.27 | 0.07 | 0.18 | 0.07 | 0.25 | 0.16 | 0.07 | 0.23 |
| TCFT error proportion score (rev) ^c | 0.31 | 0.82 | 0.51 | 0.23 | 0.74 | 0.51 | 0.17 | 0.64 | 0.48 |
| RAN correct minus incorrect | 0.50 | 0.55 | 0.05 | 0.52 | 0.52 | 0.00 | 0.50 | 0.48 | 0.02 |
| UL gestalt correct minus incorrect | 0.63 | 0.60 | 0.04 | 0.81 | 0.82 | 0.01 | 0.72 | 0.68 | 0.05 |
| UL fsbm correct minus incorrect | 0.52 | 0.60 | 0.08 | 0.43 | 0.41 | 0.62 | 0.42 | 0.34 | 0.06 |

Note. ES = effect size, calculated as Cohen's *d*; 0 = no ADHD; 1 = inattentive; 2 = combined; WISC-III = Wechsler Intelligence Scale for Children (3rd ed.); CPT = Continuous Performance Test; TCFT = Taylor Complex Figure Test; RAN = Rapid Automatized Naming Test; UL = Underlining Test.
^a Covariates included follow-up family income, child age, reading disorder, disruptive disorder, and internalizing disorder.

^b Covariates included demographic and comorbidity set plus baseline child full scale IQ.

^c Reverse scored so that for all variables, higher means indicate better performance.

after initial ascertainment, with adolescent ADHD status more specifically related to neuropsychological dysfunction than childhood ADHD status.

The largest prospective investigation including neuropsychological/executive measures is that of Fischer et al. (2005). With a sample of over 90% men, followed for an average of 13 years from childhood through early adulthood, they found evidence for continuing neuropsychological deficits, chiefly restricted to those who maintained their ADHD status across the longitudinal interval. In all, our findings document the continuing neuropsychological deficits of girls with ADHD across a 5-year span, which were independent of comorbidities and demographic factors, even though attentional performance showed improvements for all participants across time. Our effects were particularly robust for the girls meeting criteria for ADHD in adolescence. Whether such deficits will persist into adulthood must await subsequent prospective follow up.

A perennial question related to investigations of youth with ADHD is whether IQ scores should be covaried when contrasting such samples with clinical or nondiagnosed comparison groups. We found that control of baseline IQ eliminated the EF, attentional, and rapid naming differences found at follow up when we contrasted our childhood-diagnosed groups. Yet the adolescent-diagnosed group comparisons survived such control. The issue is contentious. On the one hand, it would be important to understand neuropsychological and executive differences between girls with ADHD and comparison girls that occur even when controlling for the 10-point-or-more IQ superiority of the comparison girls at baseline (see Hinshaw et al., 2002, for rationale and presentation of baseline data, which revealed that during childhood all neuropsychological deficits survived control of IQ). However, it may well be that controlling for intelligence scores in such designs is tantamount to overcontrol (see, for example, Miller & Chapman, 2001), given that IQ deficits are an inherent part of the ADHD construct. Although Fischer et al. (2005) opted not to covary IQ scores in their recent report for just this reason, we performed our analyses both ways—that is, with and without covarying IQ—because of the potential importance of examining neuropsychological differences that are truly independent of overall cognitive functioning (see, for example, Barkley, 1997).

Although control of baseline IQ (in addition to demographics and comorbidities) eliminated group differences at follow up when diagnostic status was configured from baseline measures, such control did not eliminate EF and other neuropsychological deficits in the ADHD group when the independent variable was diagnostic status as ascertained at follow up. Thus, parallel to the report of Fischer et al. (2005), we found that our most robust findings in terms of persisting neuropsychological deficits pertained to ADHD status contemporaneous with the follow-up testing. Longer term follow up will be crucial to explore this issue further.

Parallel to our childhood findings, we found little evidence for ADHD-C versus ADHD-I differences in neuropsychological performance. In fact, in both the childhood-diagnosed and adolescent-diagnosed analyses, these “type” differences were never significant except for one instance for the adolescent-diagnosed girls, even though power calculations reveal that, given our follow-up sample sizes, we have statistical power of nearly .6 to detect effects of medium size (see Faul & Erdfelder, 1992). Despite contentions that the inattentive type of ADHD is a separate, qualitatively distinct variant of this disorder, particularly with respect to cognitive performance (Milich, Balentine, & Lynam, 2001), we found no such evidence, at least with respect to the variables in our battery. Whether other cognitive measures would yield such effects must await additional research.

Two comments on the ADHD-C versus ADHD-I contrasts are in order. First, during the year prior to the follow-up assessment, 45% of the former group received medication treatment for ADHD, in contrast to 28% of the latter, a difference that just missed statistical significance (see Hinshaw et al., 2006, for details). Thus, it is conceivable that the tendency for more of the girls with the combined type of ADHD to receive medication treatment served to attenuate neuropsychological dysfunction, although we reiterate that nearly all participants performed the follow-up battery off of medication. Second, it is possible that the ADHD-I group includes a smaller subset of “truly” inattentive participants who exhibit sluggish cognitive tempo (McBurnett, Piffner, & Frick, 2001). Paralleling our article describing baseline neuropsychological performance (Hinshaw et al., 2002), we formed a subset of the ADHD-I participants at baseline who (a) displayed very few HI symptoms and (b) revealed high levels of parent- and teacher-

rated sluggish cognitive tempo. However, just as in childhood, this “refined” inattentive subgroup did not reveal any different pattern of neuropsychological performance from that of the remainder of the ADHD-I group.

Executive dysfunction may be a continuing problem for youth with ADHD (or a subset of such youth) as they develop. Important in such research will be several issues. First, ascertaining EF deficits through ecologically valid, real-world measures over and above laboratory tasks is a priority (see Lawrence et al., 2004). Second, in causal models of ADHD, it will be important to understand that EF may well serve as one component of a multi-level process (Coghill et al., 2005; Sonuga-Barke, 2002). Third, it would be extremely helpful for different research laboratories to converge with respect to common measures of EF so that replications of findings could be more clearly established. Finally, an important question arises as to the proportions of youth with ADHD who display dysfunctions in EF. We made a preliminary determination of this important variable by classifying as “executive dysfunctional” those participants who scored at or below the 10th percentile of the comparison group’s scores on at least two executive tests (see Nigg, Willcutt, Doyle, & Sonuga-Barke, 2005). We found that 43% of the baseline-defined ADHD-I group and 50% of the ADHD-C group showed such dysfunction, roughly consistent with the literature revealing that approximately half of youth with ADHD show executive dysfunction (Nigg et al., 2005). In addition, by follow up, at which time we characterized smaller groups of girls with adolescent ADHD, 55% of the adolescent-diagnosed girls with ADHD-I and 75% with ADHD-C had executive dysfunction. Further analyses of categorically defined executive dysfunctional subgroups will be the subject of additional articles from our research group.

Repeated measures analyses of the CPT scores first revealed that all participants improved across time. For omissions, the girls with ADHD remained significantly worse at follow up. For commissions, however, the pattern was somewhat more complex: The main effect of time (revealing improvement for all participants) was qualified by a Time \times Diagnostic Status interaction, revealing greater improvement of the comparison girls and those with ADHD-C. Why the ADHD-I group showed more stability, across time, is unknown; longer term follow up is needed to understand developmental trajectories.

Limitations of this investigation include, first, the fact that most of our follow-up measures were not exact replicas of our childhood battery. Thus, except for CPT omissions and commissions, we could not examine precise change over time. Second, the follow-up battery was limited to 30 min of testing for practical reasons related to the need for a multidomain assessment protocol. It would be important to know whether girls with ADHD reveal problems with other measures of EF and for neuropsychological constructs beyond EF and attentional problems per se. Third, our retention rate of 92% for the overall follow up (Hinshaw et al., 2006) was reduced by in-home assessments, equipment failure, and missed tests for selected additional participants. Still, the follow-up sample appears representative of the baseline participants. Finally, ours is largely a clinical sample of girls with ADHD, and it is unknown whether similar persistence of neuropsychological and executive dysfunction would emerge from a community sample.

Overall, our findings are consistent with those of Seidman et al. (2005), who found that girls with ADHD show EF deficits during adolescence (at a rate comparable with that of boys with ADHD), and are consistent with those of Fischer et al. (2005), who have provided the most convincing prospective evidence to date that cognitive, attentional, and neuropsychological problems among youth with ADHD do not abate, relative to a comparison group, even by early adulthood. Those youth with ADHD with persisting ADHD symptomatology over time show the strongest evidence for continuing neuropsychological and executive dysfunction, paralleling findings with other samples primarily consisting of male participants with ADHD (Fischer et al., 2005). It may well be that those children with ADHD with the most persistent forms of underlying symptoms are those whose executive dysfunction is most striking by adolescence; further longitudinal work is a priority, with a wider range of measures of EF. Indeed, understanding how EF deficits relate to other vital aspects of the functioning of youth with ADHD across the lifespan is a continuing objective for high-quality research.

References

- American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.
- Barkley, R. A. (1997). Behavioral inhibition, sustained attention, and executive functions: Constructing a unifying theory of ADHD. *Psychological Bulletin*, *121*, 65–94.
- 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, J. H., & Waber, D. P. (1996). *Developmental scoring system for the Rey–Osterrieth Complex Figure*. Odessa, FL: Psychological Assessment Resources.
- Biederman, J., Monuteaux, M., Seidman, L., Doyle, A. E., Mick, E., Wilens, T., et al. (2004). Impact of executive function deficits and ADHD on academic outcomes in children. *Journal of Consulting and Clinical Psychology*, *72*, 757–766.
- 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.
- Coghill, D., Nigg, J., Rothenberger, A., Sonuga-Barke, E., & Tannock, R. (2005). Whither causal models in the neuroscience of ADHD. *Developmental Science*, *8*, 105–114.
- 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.
- Denckla, M. B., & Rudel, R. G. (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. (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.
- Drechsler, R., Brandeis, D., Földényi, M., Imhof, K., & Steinhausen, H.-C. (2005). The course of neuropsychological functions in children

- with attention deficit hyperactivity disorder from late childhood to early adolescence. *Journal of Child Psychology and Psychiatry*, 46, 824–836.
- Faul, F., & Erdfelder, E. (1992). *GPOWER: A priori, post-hoc, and compromise power analysis for MS-DOS* [Computer software]. Bonn, Germany: Department of Psychology, Bonn University.
- Fischer, M., Barkley, R., Smallish, L., & Fletcher, K. (2005). Executive functioning in hyperactive children as young adults: Attention, inhibition, response perseveration, and the impact of comorbidity. *Developmental Neuropsychology*, 27, 107–133.
- Gabrieli, J. D. E., Desmond, J., Domb, J., Wagner, A., & Stone, M. (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, 1036–1045.
- Gershon, J. (2002). A meta-analytic review of gender differences in ADHD. *Journal of Attention Disorders*, 5, 143–154.
- 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.
- Hechtman, L., Weiss, G., Perlman, T., & Amsel, R. (1984). Hyperactives as young adults: Initial predictors of adult outcome. *Journal of the American Academy of Child Psychiatry*, 23, 250–260.
- Helmes, E. (2000). Learning and memory. In G. Groth-Marnat (Ed.), *Neuropsychological assessment in clinical practice: A guide to test interpretation and integration* (pp. 293–334). New York: Wiley.
- Hervey, A. S., Epstein, J., & Curry, J. (2004). Neuropsychology of adults with attention-deficit/hyperactivity disorder: A meta-analytic review. *Neuropsychology*, 18, 485–503.
- 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., Carte, E. T., Sami, N., Treuting, J. J., & Zupan, B. A. (2002). Preadolescent girls with attention-deficit/hyperactivity disorder: II. Neuropsychological performance in relation to subtypes and individual classification. *Journal of Consulting and Clinical Psychology*, 70, 1099–1111.
- Hinshaw, S. P., March, J., Abikoff, H., Arnold, L., Cantwell, D., Conners, C., 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.
- Hinshaw, S. P., Owens, L., Sami, N., & Fargeon, S. (2006). Prospective follow-up of girls with attention-deficit/hyperactivity disorder into adolescence: Evidence for continuing cross-domain impairment. *Journal of Consulting and Clinical Psychology*, 74, 489–499.
- Hopkins, J., Perlman, T., Hechtman, L., & Weiss, G. (1979). Cognitive styles in adults previously diagnosed as hyperactives. *Journal of Child Psychology and Psychiatry*, 20, 209–216.
- Klorman, R., Hazel-Fernandez, L. A., Shaywitz, S. E., Fletcher, J. M., Marcinone, 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.
- Lawrence, V., Houghton, S., Douglas, G., Durkin, K., Whiting, K., & Tannock, R. (2004). Executive function and ADHD: A comparison of children's performance during neuropsychological testing and real-world activities. *Journal of Attention Disorders*, 7, 137–149.
- Lezak, M. D. (1983). *Neuropsychological assessment* (2nd ed.). New York: Oxford University Press.
- MacLeod, D., & Prior, M. (1996). Attention deficit in adolescents with ADHD and other clinical groups. *Child Neuropsychology*, 2, 1–10.
- Mahone, E. M., Hagelthorn, K. M., Cutting, L. E., Schuerholz, L. J., Pellitier, S. F., Rawlins, C., et al. (2002). Effects of IQ on executive function measures in children with ADHD. *Child Neuropsychology*, 8, 52–65.
- Mannuzza, S., & Klein, R. (1999). Adolescent and adult outcomes in attention deficit/hyperactivity disorder. In H. Quay & A. Hogan (Eds.), *Handbook of disruptive behavior disorders* (pp. 279–294). New York: Kluwer Academic/Plenum.
- Martinussen, R., Hayden, J., Hogg-Johnson, S., & Tannock, R. (2005). A meta-analysis of working memory impairments in children with attention-deficit/hyperactivity disorder. *Journal of the American Academy of Child & Adolescent Psychiatry*, 44, 377–384.
- McBurnett, K., Piffner, 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.
- 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 and Practice*, 8, 463–488.
- Miller, G. A., & Chapman, J. P. (2001). Misunderstanding analysis of covariance. *Journal of Abnormal Psychology*, 110, 40–48.
- Nigg, J. T., Blaskey, L. G., Huang-Pollock, C. L., & Rappley, M. D. (2002). Neuropsychological executive functions and DSM-IV ADHD 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.
- Nigg, J. T., Willcutt, E. G., Doyle, A. E., & Sonuga-Barke, E. J. S. (2005). Causal heterogeneity in attention-deficit/hyperactivity disorder: Do we need neuropsychologically impaired subtypes? *Biological Psychiatry*, 57, 1224–1230.
- Pennington, B. F., & Ozonoff, S. (1996). Executive functions and developmental psychopathology. *Journal of Child Psychology & Psychiatry & Allied Disciplines*, 37, 51–87.
- Posner, M. I. (1988). Structures and function of selective attention. In T. Boll & D. K. Bryant (Eds.), *Clinical neuropsychology and brain functions: Research, assessment, and practice* (pp. 173–203). Washington, DC: American Psychiatric Association.
- Rourke, B. P., & Gates, R. D. (1980). *The Underlining Test (preliminary norms)*. Windsor, Ontario, Canada: Department of Psychology, University of Windsor.
- Rourke, B. P., & Orr, R. (1977). Prediction of the reading and spelling performances of normal and retarded readers: A four-year follow-up. *Journal of Abnormal Child Psychology*, 5, 9–20.
- Sami, N., Carte, E. T., Hinshaw, S. P., & Zupan, B. A. (2003). Performance of girls with attention deficit/hyperactivity disorder on the Rey-Osterrieth Complex Figure: Evidence for executive function deficits. *Child Neuropsychology*, 9, 237–254.
- Scheres, A., Oosterlaan, J., Geurts, H., Morein-Zamir, S., Meiran, N., Schut, H., et al. (2004). Executive functioning in boys with ADHD: Primarily an inhibition deficit. *Archives of Clinical Neuropsychology*, 19, 569–594.
- Seidman, L., Biederman, J., Monuteaux, M., Valera, E., Doyle, A., & Faraone, S. (2005). Impact of gender and age on executive functioning: Do girls and boys with and without attention deficit hyperactivity disorder differ neuropsychologically in preteen and teenage years? *Developmental Neuropsychology*, 27, 79–105.
- Seidman, L., Doyle, A., Fried, R., Valera, E., Crum, A., & Matthews, L. (2004). Neuropsychological functioning in adults with attention-deficit/hyperactivity disorder. *Psychiatric Clinics of North America*, 27, 261–282.
- Seidman, L. J., Biederman, J., Faraone, S. V., Weber, W., & Ouellette, C. (1997). Toward defining a neuropsychology of attention-deficit/hyper-

- activity 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., Valera, E. M., Monuteaux, M. C., Doyle, A. E., & Faraone, S. V. (2006). Neuropsychological functioning in girls with attention-deficit/hyperactivity disorder with and without learning disabilities. *Neuropsychology*, *20*, 166–177.
- Sergeant, J. A., Geurts, H., & Oosterlaan, J. (2002). How specific is a deficit of executive functioning for attention-deficit/hyperactivity disorder. *Behavioural Brain Research*, *130*, 3–28.
- 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.
- Sonuga-Barke, E. J. (2002). Psychological heterogeneity in AD/HD—A dual pathway model of behaviour and cognition. *Behavioural Brain Research*, *130*, 29–36.
- Spreen, O., & Strauss, E. (1998). *A compendium of neuropsychological tests: Administration, norms, and commentary* (2nd ed.). New York: Oxford University Press.
- Swanson, J. M. (1992). *Assessment and treatment of ADD students*. Irvine, CA: K.C. Press.
- Taylor, L. B. (1969). Localization of cerebral lesions by psychological testing. *Clinical Neurosurgery*, *16*, 269–287.
- 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.
- Wechsler, D. (1991). *Wechsler Intelligence Scale for Children* (3rd ed.). San Antonio, TX: Psychological Corporation.
- Weyandt, L. (2005). Executive function in children, adolescents, and adults with attention deficit hyperactivity disorder: Introduction to the special issue. *Developmental Neuropsychology*, *27*, 1–10.
- Willcutt, E. G., Pennington, B., Olson, R., Chhabildas, N., & Hulslander, J. (2005). Neuropsychological analyses of comorbidity between reading disability and attention deficit hyperactivity disorder: In search of the common deficit. *Developmental Neuropsychology*, *27*, 35–78.
- Woods, S. P., Lovejoy, D. W., & Ball, J. D. (2002). Neuropsychological characteristics of adults with ADHD: A comprehensive review of initial studies. *Clinical Neuropsychology*, *16*, 12–34.

Received September 27, 2005

Revision received September 5, 2006

Accepted September 18, 2006 ■

Correction to Prescott, Newton, Mir, Woodruff, and Parks (2006)

In the article “A New Dissimilarity Measure for Finding Semantic Structure in Category Fluency Data With Implications for Understanding Memory Organization in Schizophrenia,” by Tony J. Prescott, Lisa D. Newton, Nusrat U. Mir, Peter W. R. Woodruff, and Randolph W. Parks (*Neuropsychology*, 2006, Vol. 20, No. 6, pp. 685–699), the Web site address provided for the supplemental data was incorrect. The correct address is <http://dx.doi.org/10.1037/0894-4105.20.6.685.supp>.

DOI: 10.1037/0894-4105.21.2.273