

Neuropsychological Analyses of Comorbidity Between Reading Disability and Attention Deficit Hyperactivity Disorder: In Search of the Common Deficit

Erik G. Willcutt
Department of Psychology
University of Colorado at Boulder

Bruce F. Pennington
Department of Psychology
University of Denver

Richard K. Olson
Department of Psychology
University of Colorado at Boulder

Nomita Chhabildas
The Children's Hospital
Denver, CO

Jacqueline Hulslander
Department of Psychology
University of Colorado at Boulder

Measures of component reading and language skills, executive functions, and processing speed were administered to groups of children with attention deficit hyperactivity disorder (ADHD; $n = 113$), reading disability (RD; $n = 109$), both RD and ADHD ($n = 64$), and neither RD nor ADHD ($n = 151$). Groups with RD exhibited

pronounced deficits on all measures of component reading and language skills, as well as significant weaknesses on measures of verbal working memory, processing speed, and response inhibition. Groups with ADHD exhibited weaknesses on all response-inhibition and processing speed tasks and were impaired on some measures of component reading skills and verbal working memory. The group with comorbid RD and ADHD exhibited the combination of the deficits in the RD-only and ADHD-only groups, providing evidence against the phenocopy and cognitive subtype hypotheses as explanations for the co-occurrence of RD and ADHD. Slow and variable processing speed was characteristic of all 3 clinical groups, suggesting that measures of this domain may be useful for future studies that search for the common genes that increase susceptibility to RD and ADHD.

Reading disability (RD) and attention deficit hyperactivity disorder (ADHD) are two of the most common disorders of childhood, each occurring in approximately 5% of the population as seen in *Diagnostic and Statistical Manual of Mental Disorders–Text Revision* (4th ed. [DSM–IV–TR]; American Psychiatric Association, 2000). ADHD and RD also co-occur significantly more frequently than expected by chance; 25% to 40% of individuals with ADHD also meet criteria for RD (e.g., Dykman & Ackerman, 1991; Semrud-Clikeman et al., 1992), whereas 15% to 40% of individuals with RD meet criteria for ADHD (Gilger, Pennington, & DeFries, 1992; Shaywitz, Fletcher, & Shaywitz, 1995; Willcutt & Pennington, 2000).

Several competing explanations have been proposed to account for comorbidity of RD and ADHD. We first briefly review data that do not support several of these hypotheses, then we describe four additional hypotheses that have received empirical support. We then summarize the implications of neuropsychological studies for these competing hypotheses and provide an overview of current knowledge regarding the neuropsychological correlates of RD, ADHD, and comorbid RD+ADHD.

ARTIFACTUAL EXPLANATIONS OF COMORBIDITY BETWEEN RD AND ADHD

Before attempting to understand the etiological underpinnings of comorbidity between disorders, it is important to rule out the possibility that the observed comorbidity is an artifact that is caused by a biased sampling procedure or measurement problem. For example, artifactual comorbidity could occur due to ascertainment biases in clinic-referred samples, common method variance in the measures used to define the disorders, symptom overlap between the disorders, or rater bias.

Most of these artifactual hypotheses can be rejected based on existing data. RD and ADHD co-occur more frequently than expected by chance in both samples ascertained from clinics (e.g., Semrud-Clikeman et al., 1992) and nonreferred samples recruited from the community (e.g., Fergusson & Horwood, 1992; Willcutt &

Pennington, 2000), indicating that this comorbidity is not restricted to clinic-referred samples. Because RD is assessed by cognitive tests, whereas ADHD is assessed by behavioral ratings, the relation between RD and ADHD cannot be explained by shared method variance. Similarly, the symptoms of RD and ADHD as defined in the *Diagnostic and Statistical Manual of Mental Disorders* (4th ed. [DSM-IV]; American Psychiatric Association, 1994) and the *DSM-IV-TR* do not overlap.

The rater-bias hypothesis is somewhat more difficult to test, and the possibility remains that parents or teachers may be more likely to endorse ADHD symptoms if they know that the child is experiencing difficulty with reading. However, previous studies indicate that in addition to higher ratings of inattention symptoms by parents and teachers, children with RD report greater difficulties with attention than children without RD on self-report measures (Willcutt, Chhabildas, & Pennington, 1998). Although the rater-bias hypothesis cannot be conclusively rejected based on these results, these data suggest that it is not likely to provide a sufficient explanation for all cases of comorbidity between RD and ADHD.

COMPETING EXPLANATIONS FOR TRUE COMORBIDITY

Numerous competing hypotheses have been proposed to explain nonartifactual comorbidity between two disorders (e.g., Caron & Rutter, 1991; Neale & Kendler, 1995). Explanations that have been supported in some previous studies of RD and ADHD include the cross-assortment hypothesis (e.g., Faraone et al., 1993), the phenocopy hypothesis (e.g., Pennington, Groisser, & Welsh, 1993), the cognitive subtype hypothesis (e.g., Rucklidge & Tannock, 2002), and the common etiology hypothesis (e.g., Willcutt, DeFries, et al., 2003). The cross-assortment hypothesis suggests that an individual with RD is more likely to have a child with an individual with ADHD than would be expected by chance based on population base rates of RD and ADHD. In a family study of the biological relatives of children with ADHD, Faraone et al. (1993) found that comorbidity between learning disabilities and ADHD was best explained by cross-assortment. This result, however, was not replicated in later studies (Doyle, Faraone, DuPre, & Biederman, 2001; Friedman, Chhabildas, Budhiraja, Willcutt, & Pennington, 2003), suggesting that cross-assortment is not likely to explain the majority of cases of comorbid RD and ADHD.

Pennington et al. (1993) described results in a small sample of children with RD and ADHD that suggested that RD might lead to the phenotypic manifestation of ADHD in the absence of the etiological influences typically associated with ADHD in isolation. This phenocopy hypothesis suggests that a child might appear to be inattentive or hyperactive in the classroom due to frustration elicited by difficulties with reading, rather than as a consequence of the neurocognitive difficulties

that are typically associated with ADHD in the absence of RD. More recent data from larger samples, however, have generally failed to support the phenocopy hypothesis (e.g., Nigg, Hinshaw, Carte, & Treuting, 1998; Rucklidge & Tannock, 2002; Seidman, Biederman, Monuteaux, Doyle, & Faraone, 2001; Willcutt, Pennington, et al., 2001).

The cognitive subtype hypothesis suggests that comorbid RD+ADHD is a third disorder that is due at least in part to etiological factors that are distinct from those that increase susceptibility to RD or ADHD alone. Therefore, this hypothesis predicts that the comorbid group will exhibit a different pattern of external correlates than would be expected based on the additive combination of the correlates of each disorder when they occur separately. Rucklidge and Tannock (2002) found that the comorbid group performed significantly worse than the RD-only and ADHD-only groups on measures of color naming, providing some support for this hypothesis. In contrast, other studies have found that the RD+ADHD group exhibited the additive combination of the deficits associated with each individual disorder (e.g., Pisecco, Baker, Silva, & Brooke, 2001; Swanson, Mink, & Bocian, 1999; Willcutt, Pennington, et al., 2001), suggesting that additional research is needed.

Finally, a series of studies tested if the relation between RD and ADHD is attributable to common etiological influences that increase susceptibility to both disorders. Twin studies indicated that RD and ADHD are each highly heritable and polygenic (e.g., DeFries & Alarcón, 1996; Faraone, Doyle, Mick, & Biederman, 2001; Fisher & DeFries, 2002; Gayán & Olson, 2001; Levy, Hay, McStephen, Wood, & Waldman, 1997; Willcutt, Pennington, & DeFries, 2000a; Willcutt, *in press*), and bivariate twin analyses suggested that comorbidity between RD and ADHD is due primarily to common genetic influences (Light, Pennington, Gilger, & DeFries, 1995; Stevenson, Pennington, Gilger, DeFries, & Gillis, 1993; Willcutt, DeFries, et al., 2003; Willcutt, Pennington, & DeFries, 2000b). Based on these results, genetic linkage studies have begun to search for chromosomal regions that may contain a gene or genes that increase risk for both disorders (Loo et al., *in press*; Willcutt et al., 2002, 2003). In the first of these studies, Willcutt, Pennington, et al. (2002; Willcutt, DeFries, et al., 2003) reported that the well-replicated quantitative trait locus for RD on Chromosome 6p21 (e.g., Cardon et al., 1994, 1995) also increases susceptibility to ADHD. Similarly, Loo et al. (*in press*) screened the entire genome and found that regions of Chromosomes 16p and 17q may contain genes that increase risk for both RD and ADHD. Thus, although several of these results await independent replication, existing data provide the strongest support for the hypothesis that comorbidity between RD and ADHD is due at least in part to a common genetic etiology.

Although these previous studies suggest that common genetic influences contribute to comorbidity of RD and ADHD, the mechanisms of these common genes are unknown. Because a single gene can influence multiple facets of neurocognitive development (e.g., Falconer & MacKay, 1996), it is possible that the same

genetic influences could increase risk for RD due to their effects on a first neural system and increase risk for ADHD due to the impact of the gene on a second, distinct pathophysiological substrate. In this case there would be no common neuropsychological deficit in groups with RD and ADHD. Alternatively, a more parsimonious model would suggest that the common genes lead to a developmental change in a single pathophysiological substrate and that this change then increases risk for both disorders. In this model all three groups (i.e., RD, ADHD, and RD+ADHD) should be characterized by at least some common neuropsychological weaknesses, with each individual's final phenotype determined by the other genetic and environmental influences that affect that individual.

ADHD Symptom Dimensions

A final factor that complicates interpretation of previous studies of comorbidity of RD and ADHD is the distinction between symptoms of inattention and hyperactivity-impulsivity in *DSM-IV*. Phenotypic analyses suggest that deficits in reading achievement and more general academic difficulties are more strongly associated with inattention than hyperactivity-impulsivity (e.g., Chhabildas, Pennington, & Willcutt, 2001; Lahey & Willcutt, 2002; Molina, Smith, & Pelham, 2001; Willcutt & Pennington, 2000; Wolraich, Feurer, Hannah, Baumgaertel, & Pinnock, 1998). Similarly, twin studies suggest that whereas the phenotypic correlation between reading deficits and inattention symptoms is primarily explained by common genetic influences, these common genes play a smaller role in the correlation between reading deficits and hyperactivity-impulsivity symptoms (Willcutt, Pennington, et al., 2003; Willcutt et al., 2000b). In contrast, Willcutt, DeFries, et al. (2003) found that bivariate linkage to Chromosome 6p was somewhat stronger for reading difficulties and hyperactivity-impulsivity symptoms than for reading difficulties and inattention symptoms. This sample was relatively small ($n = 48-83$ sibling pairs), however, and a much larger sample would be needed for this difference to reach statistical significance. Nonetheless, these inconsistent results suggest that the relations between reading difficulties and the *DSM-IV* ADHD symptom dimensions and subtypes are likely to be complex and, therefore, warrant additional research.

NEUROPSYCHOLOGY OF RD AND ADHD

Although several different types of studies are required to test definitively between all possible competing explanations for comorbidity (e.g., Neale & Kendler, 1995), neuropsychological methods can be used to test several of the competing models described most frequently in the literature (e.g., Pennington et al., 1993; Rucklidge & Tannock, 2002; Willcutt, Pennington, et al., 2001). Therefore, in this

section we briefly summarize current knowledge regarding the neuropsychology of RD and ADHD, then turn in the subsequent section to studies that used neuropsychological methods to examine the etiology of comorbidity between RD and ADHD.

Neuropsychological Profile of RD

Studies of individuals with and without reading difficulties suggest that *phonological decoding* (PD), defined as the ability to translate sequences of printed letters into the corresponding sounds, plays a central role in both normal and abnormal reading development (e.g., Pennington, 2002; Wagner & Torgesen, 1987). The unique contribution of PD to most cases of RD has been suggested by the presence of significant group deficits in PD when older children with RD are compared to younger readers without RD who are reading at the same level (Olson, 1985; Rack, Snowling, & Olson, 1992). Moreover, twin studies have shown that there are strong genetic influences on the group deficit in PD (Gayán & Olson, 2001; Olson, Wise, Conners, Rack, & Fulker, 1989) and that these genes also influence the group deficit in word reading (Gayán & Olson, 2001; Olson, Forsberg, & Wise, 1994). Deficits in PD and word reading are in turn linked to genetic influences on deficits in the oral language skill of phoneme awareness, defined as the ability to recognize and manipulate the phonemic constituents of speech (Gayán & Olson, 2001; Olson et al., 1994). Problems with phoneme awareness are regarded by many as the most proximal cause of most cases of RD (cf. Wagner, Torgesen, & Rashotte, 1994).

Although deficits in groups with RD are most pronounced on measures of phoneme awareness and other facets of phonological processing, recent studies suggest that individuals with RD also have weaknesses in several other neurocognitive domains. These include difficulty accessing the orthographic representation of words from the lexicon (Gayán & Olson, 2001), weaknesses in other areas of speech and language processing (e.g., Olson, 1994; Pisecco et al., 2001), slower verbal naming speed (e.g., Compton, Olson, DeFries, & Pennington, 2002; Denckla & Rudel, 1976; Semrud-Clikeman, Guy, Griffin, & Hynd, 2000; Tannock, Martinussen, & Frijters, 2000), and weaknesses in executive domains such as verbal working memory (e.g., Roodenrys, Koloski, & Grainger, 2001; Swanson, Mink, & Bocian, 1999; Willcutt, Pennington, et al., 2001), set shifting (Weyandt, Rice, Linterman, Mitzlaff, & Emert, 1998), planning (e.g., Klorman et al., 1999), and response inhibition (e.g., Purvis & Tannock, 2000; Willcutt, Pennington, et al., 2001).

Neuropsychological Profile of ADHD

A large body of research indicates that groups with ADHD differ significantly from groups without ADHD on a variety of neurocognitive measures (see reviews by Barkley, 1997a, 1997b; Nigg, 2000, 2001; Pennington, 2002; Pennington &

Ozonoff, 1996). Based on similarities between ADHD symptoms and the behavioral sequelae of frontal lobe injuries, several authors have proposed that ADHD is attributable to a core deficit in some facet of executive functions (EFs), defined as *cognitive functions* that serve to maintain an appropriate problem-solving set to attain a future goal (e.g., Barkley, 1997a, 1997b; Pennington & Ozonoff, 1996).

Previous theorists have criticized the construct of EFs as weakly defined and overly broad (e.g., Pennington, 1997). Because many putative EF tasks are relatively nonspecific, poor performance on a specific task could be attributable to difficulties in any of several different aspects of the task. Moreover, definitions of EF often include a wide range of tasks, many of which appear to require somewhat different aspects of neurocognitive functioning. The multifactorial nature of EF is demonstrated by factor analyses of several batteries of putative EF measures (Mariani & Barkley, 1997; Miyake, Friedman, Emerson, Witzki, & Howerter, 2000; Pennington, 1997; Willcutt, Pennington, et al., 2001). All of these studies found that EF tasks tap more than one latent dimension of neurocognitive functioning. Although the specific factors varied somewhat among the studies, the overall pattern of results suggests that EF tasks may comprise at least four domains: (a) response inhibition, (b) working memory or updating, (c) set shifting or task switching, and (d) interference control.

Studies of children with a *DSM-IV* diagnosis of ADHD underscore the potential importance of the distinction between these different EF domains (Bedard et al., 2003; Chhabildas et al., 2001; Hinshaw, Carte, Sami, Treuting, & Zupan, 2002; Houghton et al., 1999; Klorman et al., 1999; Nigg, Blaskey, Huang-Pollock, & Rappley, 2002; Rucklidge & Tannock, 2002; Schmitz et al., 2002). A recent meta-analysis found that groups with *DSM-IV* ADHD performed significantly worse than groups without ADHD on measures of EF domains, such as response inhibition, planning or organization, and working memory, as well as measures of domains with less of an executive component, such as processing speed, rapid naming, and fine and gross motor skill (Willcutt et al., in press). In contrast, the ADHD groups were not consistently impaired on measures of set shifting or interference control.

Although these studies suggest that ADHD is associated with significant weaknesses in several EF domains, these results challenge the hypothesis that any specific EF deficit is the core neurocognitive deficit in ADHD. The mean effect size between groups with and without ADHD is moderate for each of the EF measures ($d = 0.4 - 0.6$; Willcutt et al., in press), suggesting that none of these neurocognitive weaknesses is a necessary or sufficient cause of ADHD. Moreover, many of these studies did not assess important covariates, such as IQ and reading achievement, leaving open the possibility that the EF deficits associated with ADHD could be attributable to group differences in IQ or to the association between ADHD and RD (Pennington & Ozonoff, 1996). Thus, in contrast to the consistently strong relation between phonological processing weaknesses and RD, the

neuropsychological profile of ADHD is not as well understood, and the core neuropsychological deficit remains elusive.

USING NEUROPSYCHOLOGY TO UNDERSTAND COMORBIDITY

To fully understand the neuropsychological correlates of RD and ADHD and the etiology of comorbidity between the two disorders, groups with RD only, ADHD only, and both RD and ADHD must be directly compared on the same measures. The common etiology, phenocopy, cross-assortment, and cognitive subtype hypotheses each make several key predictions regarding the relations among the groups on these measures.

Predictions of the Competing Hypotheses Regarding a Double Dissociation Between RD and ADHD

A double dissociation occurs when two disorders are associated with opposite patterns of impairment in two different cognitive domains. Because the phenocopy, cross-assortment, and cognitive subtype hypotheses suggest that RD and ADHD are caused by different etiological influences, all three models predict a significant double dissociation between RD and ADHD. In contrast, although the common genetic etiology hypothesis also predicts a double dissociation between RD and ADHD for all measures that are not influenced by the common genetic effects, it predicts that both the RD-only and ADHD-only groups will exhibit weaknesses in any neuropsychological domains that are influenced by the common genes.

In the first study that used a full 2×2 (RD \times ADHD) design to examine performance on measures of EF and phonological processing, Pennington et al. (1993) found a double dissociation between RD and ADHD. Specifically, the group with ADHD alone was significantly impaired on EF measures but not on measures of phonological processing, whereas the group with RD alone exhibited severe phonological processing deficits but was not impaired on the EF measures. Subsequent studies generally supported this double dissociation, although several suggested that RD may also be associated with mild deficits on at least a subset of EF measures (e.g., Klorman et al., 1999; Nigg et al., 1998; Purvis & Tannock, 2000; Rucklidge & Tannock, 2002; Shaywitz et al., 1995; Willcutt, Pennington, et al., 2001). These mixed results provide some support for all four hypotheses, suggesting that additional research is needed to better understand which neurocognitive weaknesses are specific to RD or ADHD and which neurocognitive difficulties are associated with both disorders.

Predictions of the Competing Hypotheses Regarding the Comorbid Group

Each competing hypothesis makes a different set of predictions regarding the neuropsychological profile of the comorbid group in relation to the groups with RD and ADHD alone. The phenocopy hypothesis suggests that one disorder often causes the symptoms of the second disorder in the absence of the neurocognitive weaknesses associated with the second disorder when it occurs alone. Therefore, it predicts that the comorbid group will exhibit the same pattern of neuropsychological deficits that are present in the group with the first disorder alone. The cognitive subtype hypothesis is less specific; it simply predicts a significant RD \times ADHD interaction on at least some measures, such that the neuropsychological deficits of the comorbid group differ from the simple additive combination of the deficits associated with RD and ADHD when they occur alone. Because the common genetic etiology and cross-assortment hypotheses each suggest that individuals with comorbid RD+ADHD have the risk factors for both RD and ADHD, these models each predict that the comorbid group will exhibit the neuropsychological weaknesses that are associated with each disorder when it occurs alone. In contrast, the common etiology hypothesis predicts that all three clinical groups will exhibit weaknesses on at least some of the same neuropsychological tasks, whereas the cross-assortment hypothesis predicts that the comorbid group will exhibit the additive combination of the distinct neuropsychological weaknesses that are associated with RD and ADHD.

Initial results suggested that the comorbid group exhibited significant phonological processing deficits in the absence of the EF deficits associated with ADHD when it occurred in the absence of RD (Pennington et al., 1993), but most later studies found that the comorbid group exhibited both the EF deficits associated with ADHD and the phonological weaknesses associated with RD (e.g., Nigg et al., 1998; Willcutt, Pennington, et al., 2001). Some recent studies have even suggested that comorbid RD may be a marker for a group of children with ADHD with more severe cognitive deficits (e.g., Purvis & Tannock, 2000; Seidman et al., 2001; Willcutt, Pennington, et al., 2001). Taken together, these results provide minimal support for the phenocopy hypothesis, but they suggest that additional research is needed to test the common genetic etiology, cross-assortment, and cognitive subtype hypotheses.

THIS STUDY

An extensive battery of neuropsychological measures was administered to groups with ADHD only, RD only, ADHD+RD, and neither ADHD nor RD. The sample described in this article is completely independent from the sample used in our previous article (Willcutt, Pennington, et al., 2001). The primary goals of the study were as follows:

1. An exploratory factor analysis (EFA) of the neurocognitive battery was conducted to examine the relations among the neuropsychological variables and simplify interpretation. We predicted that the battery of EF measures would be best described by two or more latent factors, rather than a single factor indicating that the EF tasks tap a unitary construct.

2. The neuropsychological profile of groups with RD, ADHD, RD+ADHD, and neither RD nor ADHD were compared to clarify the pattern of neurocognitive weaknesses associated with RD and ADHD independent of the influence of the other disorder. Based on previous research, it was predicted that RD would be associated with prominent weaknesses on all measures of component reading and language skills, coupled with milder weaknesses on EF and processing speed tasks. In contrast, we anticipated that ADHD would be associated with deficits on EF and processing speed tasks that would be most pronounced on measures of response inhibition, but that ADHD would not be independently associated with deficits on the reading and language measures.

3. To examine the etiology of comorbidity between RD and ADHD, the neuropsychological profile of the three clinical groups was compared. Based on our previous results and other findings in the literature, we predicted that the comorbid group would exhibit the deficits associated with both RD and ADHD, providing further evidence against the phenocopy and cognitive subtype hypotheses as explanations for comorbidity between RD and ADHD.

4. By examining the extent to which each neuropsychological deficit was associated with the three clinical groups, we tested the potential utility of these neuropsychological measures as markers for the common genetic etiology of RD and ADHD. Although we did not have strong a priori predictions for this analysis, some previous studies reported promising results for measures of processing speed (e.g., Rucklidge & Tannock, 2002).

5. Finally, a series of ancillary analyses were conducted to test if the pattern of neuropsychological weaknesses varied as a function of ADHD subtype. Based on initial results in this sample (Chhabildas et al., 2001), we did not anticipate any significant differences between the inattentive and combined subtypes.

METHOD

Participants

Recruitment. Participants completed the measures described in this article as part of the Colorado Learning Disabilities Research Center (CLDRC) twin study, an ongoing study of the etiology of learning disabilities, ADHD, and other related disorders (e.g., DeFries et al., 1997; Willcutt, DeFries, et al., 2003). In collaboration with 22 local school districts, parents of all twins between the ages of 8

and 18 were contacted by letter and invited to participate in the study. After initial parental consent was obtained, two parallel recruitment procedures were conducted independently to identify twin pairs in which at least one of the twins met criteria for ADHD, or at least one of the twins exhibited significant reading difficulties, as well as a comparison sample of twin pairs in which neither twin exhibited either ADHD or reading difficulties.

To identify twin pairs in which at least one twin exhibited significant reading difficulties, parental consent was requested to allow study staff to review each twin's academic records. If either member of a twin pair had a positive history of learning difficulties (e.g., low achievement test scores, referral to a tutor, reports by classroom teachers or school psychologists), both members of the pair were invited to participate in the larger study. To identify twins with ADHD, parents and teachers were asked to complete the Disruptive Behavior Rating Scale (Barkley & Murphy, 1998) to assess symptoms of *DSM-IV* ADHD. If either of the twins met criteria for any *DSM-IV* ADHD subtype based on parent or teacher ratings, the twin pair was invited to participate. The comparison sample was composed of twin pairs from the same school districts in which neither twin met screening criteria for ADHD or reading difficulties. Approximately 35% of the families who were contacted agreed to participate in the initial screening procedure, and 95% of the families in the screening sample agreed to participate in the larger study if invited.

Exclusionary criteria. CLDRC staff conducted a telephone screening interview prior to any testing. Because the focus of the overall project is on the etiology and correlates of familial RD and ADHD, potential participants with a documented brain injury, significant hearing or visual impairment, or other rare genetic or environmental etiology (e.g., Fragile X syndrome, Down syndrome, or other sex chromosome anomalies) were excluded from the sample. In addition, 3 participants were excluded from analyses due to a Full Scale IQ (FSIQ) score below 75 on the Wechsler Intelligence Scale for Children–Revised (WISC–R; Wechsler, 1974).

Diagnostic Measures and Operational Definition of RD and ADHD

Assessment of RD. Academic achievement in reading and mathematics was assessed with the Peabody Individual Achievement Test (PIAT; Dunn & Markwardt, 1970). To simplify interpretation, a normally distributed reading composite score was created based on a previous discriminant function analysis of the PIAT Reading Recognition, Reading Comprehension, and Spelling subtests (DeFries, 1985). A standard score 1.75 *SDs* below the estimated population mean was used as the cutoff score for RD. This cutoff selects approximately 5% of the control sample, consistent with the estimated population prevalence of RD (e.g., *DSM-IV-TR*). Participants were categorized as RD ($n = 173$) if they had a positive

school history of reading problems and scored below this cutoff on the reading discriminant function score.

Assessment of ADHD. The Disruptive Behavior Rating Scale (DBRS; Barkley & Murphy, 1998) was used to obtain parent and teacher ratings of the 18 symptoms of *DSM-IV* ADHD. Each symptom on the DBRS is rated on a 4-point scale (*never or rarely, sometimes, often, and very often*). Items rated as *often* or *very often* were scored as positive symptoms, and items rated as *never or rarely* or *sometimes* were scored as negative symptoms, consistent with the procedure used in previous studies of similar rating scales (e.g., Pelham, Gnagy, Greenslade, & Milich, 1992). Previous results from this sample and others indicate that parent and teacher ratings on the DBRS and other similar scales are internally consistent ($\alpha = 0.92\text{--}0.96$) and have adequate-to-high test-retest reliability ($r = .59\text{--}.89$; e.g., DuPaul, Power, Anastopoulos, & Reid, 1998; Willcutt, Chhabildas, et al., 2001).

The algorithm from the *DSM-IV* field trials for the disruptive behavior disorders was used to combine parent and teacher ratings of ADHD symptoms (Lahey et al., 1994). This procedure codes each symptom as positive if it is endorsed by either the parent or the teacher. Consistent with *DSM-IV* criteria, children were categorized as ADHD only if symptoms were present prior to age 7 and only if these symptoms caused significant functional impairment. Individuals with six or more symptoms of inattention but fewer than six symptoms of hyperactivity-impulsivity were coded as predominantly inattentive type, participants with six or more symptoms of hyperactivity-impulsivity but fewer than six symptoms of inattention were categorized as predominantly hyperactive-impulsive type, and individuals with six or more symptoms on both dimensions were coded as combined type.

A total of 190 participants met criteria for *DSM-IV* ADHD. Consistent with other community samples (e.g., DuPaul et al., 1998; Gaub & Carlson, 1997), the majority of participants who met symptom criteria for *DSM-IV* ADHD met criteria for the inattentive type ($n = 115$), and most of the remaining participants met symptom criteria for the combined type ($n = 62$). Only 13 participants met criteria for the hyperactive-impulsive type. Results from our sample and others call into question the validity of the hyperactive-impulsive type in school-age children (Willcutt et al., in press) and suggest that this subtype is not consistently associated with the neuropsychological weaknesses that characterize the inattentive and combined types (e.g., Chhabildas et al., 2001; Schmitz et al., 2002). In light of these results and the small number of individuals with the hyperactive-impulsive type, individuals in this group were excluded from the analyses described in this article.

Assignment to groups. The use of twins for phenotypic analyses in which each twin is considered as an individual data point presents a methodological difficulty, as the scores of the twins in each pair do not represent fully independent ob-

servations. Therefore, one twin was selected at random from each twin pair in which both twins met inclusion criteria for one of the four groups. Results were virtually identical when analyses were repeated in a sample in which the selected twin was replaced by the co-twin that was excluded from the first set of analyses, suggesting that the random selection of one twin from each of these pairs did not inadvertently bias the results. A dummy code for zygosity was included in all initial models to control for any differences between participants from MZ and DZ pairs, but this code was dropped from all final models because it had no significant impact on any result.

Individuals were assigned to groups based on the criteria for RD and ADHD described previously. After selecting one twin randomly from each pair in which both twins met criteria for one of the four groups, the total sample included 109 individuals with RD only, 113 participants with ADHD only, 64 participants with both RD and ADHD, and 151 participants with neither RD nor ADHD. Approximately 37% of the individuals with either RD or ADHD also met criteria for the other disorder, a rate of comorbidity that is consistent with previous results in an independent subset of our sample (e.g., Willcutt & Pennington, 2000; Willcutt, Pennington, et al., 2001) and in samples from other studies (e.g., Semrud-Clikeman et al., 1992). The rate of comorbid RD was similar in groups with the *DSM-IV* inattentive type (39.1%) and combined type (33.9%). Similarly, other recent studies that found that the means of the inattentive and combined types were not significantly different on a variety of measures of reading achievement (e.g., Carlson, Booth, Shin, & Canu, 2002; Faraone, Biederman, Weber, & Russell, 1998; Hinshaw, 2002; Nigg et al., 2002; Todd et al., 2002).

Descriptive characteristics of the groups. The mean age of the four groups was not significantly different (Table 1). In contrast, the mean socioeconomic status (SES) as measured by the Hollingshead (1975) two-factor inventory was significantly lower in all three clinical groups than in the comparison group. The mean WISC-R Verbal, Performance, and FSIQ scores of the three clinical groups also fell significantly below the mean of the comparison group.

As expected, based on the way the sample was defined, the mean number of ADHD symptoms was higher for the ADHD and RD+ADHD groups than for the RD and comparison groups, and the RD and RD+ADHD groups had lower means on the reading and spelling measures than did the ADHD and comparison groups (Table 1). In addition, in comparison to the group without RD or ADHD, the group with ADHD alone scored significantly lower on the measures of reading and spelling achievement, and the group with RD alone exhibited significantly more symptoms of inattention. These results suggest that at least a subset of individuals with ADHD or RD alone exhibit subclinical manifestations of the other disorder even though they do not meet full criteria for the other diagnosis.

TABLE 1
Descriptive Characteristics

Variables	Comparison (n = 151)		ADHD Only (n = 113)		RD Only (n = 109)		RD + ADHD (n = 64)		F ^a
	M	SD	M	SD	M	SD	M	SD	
Demographic									
Sex	65 male, 86 female		74 male, 39 female		56 male, 53 female		40 male, 24 female		
Age	11.5 3.46 _a	2.5 1.05	11.2 2.98 _b	2.7 1.26	11.0 2.80 _b	2.4 1.15	11.1 2.81 _b	2.5 1.08	0.97 8.92*
Socioeconomic status									
WISC-R									
Full Scale IQ	113.6 _a	10.5	104.3 _b	11.6	96.8 _c	11.1	92.4 _c	11.0	68.98*
Verbal IQ	112.9 _a	11.9	105.1 _b	13.7	95.6 _c	12.1	91.7 _c	12.8	62.69*
Performance IQ	111.1 _a	11.9	103.6 _b	11.4	99.0 _c	12.4	95.4 _c	11.3	34.91*
Academic achievement									
PIAT Math	111.5 _a	12.2	103.9 _b	13.9	93.6 _c	11.8	91.9 _c	11.1	59.13*
PIAT Reading Recognition	109.2 _a	10.0	102.7 _b	8.4	86.3 _c	7.8	85.6 _c	7.3	199.29*
PIAT Reading Comprehension	111.0 _a	10.1	105.8 _b	11.4	89.8 _c	8.6	87.7 _c	8.9	146.42*
PIAT Spelling	107.8 _a	11.2	99.9 _b	10.5	87.0 _c	9.0	85.8 _c	9.4	111.87*
Reading Discriminant score	1.51 _a	1.03	0.67 _b	0.93	-1.30 _c	0.68	-1.43 _c	0.70	300.41*
ADHD Symptoms									
Inattention	0.6 _a	1.7	7.0 _b	2.2	1.6 _c	1.7	7.5 _b	1.7	463.69*
Hyperactivity-impulsivity	0.5 _a	1.0	4.6 _b	3.2	0.8 _a	1.2	4.5 _b	3.2	94.33*
Total symptoms	1.1 _a	1.8	11.6 _b	3.6	2.5 _c	2.3	11.9 _b	3.4	449.94*

Note. Means with different subscripts are significantly different. ADHD = attention deficit hyperactivity disorder; RD = reading disorder; WISC-R = Wechsler Intelligence Scale for Children-Revised; PIAT = Peabody Individual Achievement Test.

^adf = 3, 432.

* $p < .01$.

Procedures

The WISC–R, PIAT, and measures of component reading and language skills were administered in two initial testing sessions completed at the University of Colorado Department of Psychology and Institute for Behavioral Genetics. The EF and processing speed tasks were completed during a third session scheduled approximately 1 month later at the University of Denver Department of Psychology. The order of the tests was counterbalanced in each of the testing sessions. Each session lasted approximately 2.5 hr, and frequent breaks were provided to minimize fatigue and maximize motivation.

All measures at both sites were administered by trained examiners who had previous experience working with children. All examiners were unaware of the diagnostic status of the child and the results of the testing conducted at the other sites. Parents of participants who were taking psychostimulant medication were asked to withhold medication for 24 hr prior to each session of the study to minimize the influence of this intervention on the results. This included 32 children in the ADHD-only group (29.2%), 22 children in the RD+ADHD group (31.3%), 5 children in the RD-only group (4.6%), and 1 child in the comparison group (0.7%).

NEUROPSYCHOLOGICAL MEASURES

The neuropsychological battery was selected to include measures that have been shown to be most strongly associated with RD or ADHD in previous studies. Participants completed measures of several component reading and language skills that strongly predict individual differences in reading ability. Based on previous studies of ADHD, the battery includes tasks that tap several different domains of EF, as well as measures of other domains that may not have as strong an EF component, such as processing speed and response variability. In this section we first provide an overview of each measure, then we describe the results of an initial EFA conducted to simplify interpretation of group differences on the tasks in the battery.

Measures of Component Reading and Language Skills

Previous studies have shown that groups with RD score significantly lower than groups without RD on measures of phoneme awareness and PD (e.g., Gayán & Olson, 2001), and both RD and ADHD are associated with lower scores on measures of orthographic coding (e.g., Willcutt, DeFries, et al., 2003). Therefore, five measures of these constructs were included in these analyses.

Phoneme awareness. The Pig-Latin test (Olson et al., 1989) requires the participant to transform words into their pig latin equivalent. The participant is told

the rules for transforming the words (e.g., move the initial phoneme to the end of the word, and add the long *a* sound) and completes nine practice words read by the examiner. The test trials then require the participant to transform words read by the examiner into their equivalent in pig latin. The primary dependent variable is a weighted percentage correct on the 45 test items, such that participants receive the maximum score for following the first-phoneme rule described in the instructions (e.g., *read-thay* for thread), but they also receive partial credit if they follow a first-letter (*hread-tay*) or onset (*ead-thray*) strategy.

The Phoneme Deletion task (Olson et al., 1994) is based on the Bruce (1964) phoneme-deletion task and the Rosner and Simon (1971) auditory-analysis task. On each of the 68 trials, participants hear a word or pronounceable nonword, which they are asked to repeat (e.g., say *plig*). They are then asked to say the nonword again after removing a specified phoneme (e.g., say *plig* without the *l*). If done correctly, the result is a word (e.g., after dropping the *l*, *plig* becomes *pig*). The primary dependent variable is the percentage of correct trials.

The Lindamood Auditory Conceptualization Test (LAC, Lindamood & Lindamood, 1971) uses colored blocks to represent phonemes and requires the participant to add, remove, or transpose blocks to reflect changes in nonwords spoken by the examiner. For example, the participant might be shown three different colored blocks in a row and told: "If this says *aps*, show me *asp*." The correct response would then be to exchange the position of the second and third blocks. The primary dependent variable is the LAC total score.

Phonological decoding. The Oral Nonword Reading task (Olson et al., 1994) requires the participant to read aloud 45 one-syllable and 40 two-syllable nonwords. The nonwords are presented on the computer one at a time and response latency is measured with a voice key. Responses are recorded for subsequent analyses of errors. The dependent variable is a composite of the response latency and accuracy scores that is standardized based on the mean and standard deviation of the comparison group in the overall CLDRC sample. Test-retest reliability of the composite score is .86 (Olson et al., 1994).

Orthographic coding. The Lexical Decision by Orthography (Olson et al., 1994) task requires participants to distinguish words from nonword letter strings that are identical in sound when pronounced (e.g., *rain-rane*). Because both letter strings sound the same, phonetic codes cannot be used to identify the real word. Instead, the participant must match the orthographic patterns of the target to the visual representation of the word in their lexicon. The task includes 80 trials presented on a computer; the trials are a mixture of relatively easy items (e.g., *cat-kat*) and more complex items (e.g., *pavement-pavemant*). To ensure that subjects must make use of word-specific knowledge rather than general information about orthographic structure, items are balanced to ensure that the distracter item is a plausible

string. The total percentage correct is the dependent variable (split-half $r = .93$; Olson et al., 1994).

Measures of EF and Related Neurocognitive Domains

Response execution–inhibition. The stop-signal task (e.g., Logan, Schachar, & Tannock, 1997; Schachar, Mota, Logan, Tannock, & Klim, 2000) is a computerized measure of inhibitory control that was developed based on the dual-process model of inhibition proposed by Logan and colleagues (e.g., Logan, 1994; Logan et al., 1997). On primary task trials, the letters *X* or *O* are presented in the center of the monitor, and the participant responds by pressing the corresponding key on the keyboard. On stop-signal trials the same visual stimulus appears, but an auditory tone is also presented shortly after the *X* or the *O* appears on the screen. The participant is instructed to press the *X* or *O* key as rapidly as possible for each trial, but is also told to inhibit the key press on each of the trials on which the tone is presented. The version of the task administered as part of this battery uses an iterative tracking procedure such that the delay between the presentation of the visual stimulus and the onset of the stop signal changes after every trial with a stop signal (e.g., Logan et al., 1997). By increasing the stop-signal delay by 50 msec if the participant is able to inhibit and decreasing the delay by 50 msec if the participant is unable to inhibit, this procedure converges on the stop-signal delay at which the participant fails to inhibit on 50% of the trials.

Stop-signal reaction time (SSRT) is the primary measure of inhibition on the stop-signal task. SSRT is estimated by subtracting the mean stop-signal delay from the mean reaction time on the primary task trials. Based on procedures used in other studies of the stop-signal task (e.g., Rucklidge & Tannock, 2002; Schachar et al., 2000), participants were excluded if they inhibited on fewer than 13% or more than 85% of the stop trials or if their SSRT was below 50 msec (2 individuals from the comparison group, 4 individuals from the RD-only group, 5 individuals from the ADHD-only group, 2 individuals from the RD+ADHD group). In addition to SSRT, mean reaction time and variability (standard deviation) of reaction time on the Go trials (trials without a stop signal) were also recorded for each participant to test whether ADHD or RD is associated with slower or more variable response speed. SSRT and reaction time variability consistently discriminate groups with and without ADHD (e.g., Chhabildas et al., 2001; Nigg, 1999; Oosterlaan & Sergeant, 1998; Schachar et al., 2000).

The Gordon Diagnostic System (Gordon, 1983) is a standardized continuous performance test (CPT) that assesses the ability to sustain attention and inhibit inappropriate responses during an extended visual task. On both the Vigilance and Distractibility subtests a single-digit number is presented in the center of the display once per second for 9 min, with target stimuli occurring 45 times during each subtest. On the Vigilance subtest the participant must press a specified button only

after a correct sequence of two digits (1 followed by 9) and then inhibit their response to all other sequences. The primary task on the Distractibility subtest is similar, but additional numbers are also presented on either side of the target stimulus. The participant is told to attend only to numbers in the center column and to ignore the other two columns. The primary dependent variables are errors of commission and omission. Both error types have been shown to discriminate groups with and without ADHD in previous studies (e.g., Chhabildas et al., 2001).

Set shifting. The Wisconsin Card Sorting Test (WCST; Heaton, 1981) requires the individual to sort 128 cards to match either the color, form, or number of shapes on four target cards. After each trial on the examiner-administered WCST, the participant is provided with verbal feedback indicating whether the response is correct or incorrect. After the participant correctly sorts 10 consecutive cards (e.g., matching to the color of the shapes on the target cards), the rule changes so that the sorting rule is based on one of the other properties of the target stimulus. Therefore, WCST performance depends on the ability to maintain a rule in memory and appropriately shift set to a new rule when presented with feedback that the previous rule is no longer correct. The primary dependent variable for this study was total *perseverative errors*, defined as errors that adhere to the previous sorting rule or to an incorrect sorting rule generated by the participant.

The WCST was included in the battery because it discriminates adults with prefrontal cortex damage from adults with lesions in other brain regions (Heaton, 1981). However, many previous studies suggest that WCST scores do not discriminate groups with and without ADHD as well as other EF measures (e.g., Pennington & Ozonoff, 1996; Willcutt, Pennington, et al., 2001), and the reliability of the WCST appears to be somewhat lower than that of the other tasks in our battery ($r = .30-.61$; Pennington, Bennetto, McAleer, & Roberts, 1995).

The Trailmaking Test (e.g., Reitan & Wolfson, 1985) assesses both processing speed and ability to shift cognitive set. Previous studies suggest that groups with ADHD perform significantly worse than groups without ADHD on this task (e.g., Chhabildas et al., 2001; Nigg et al., 2002). Part A of the task requires the participant to connect a series of circles containing numbers in ascending order. In part B each circle contains either a number or a letter. The participant is again instructed to connect the circles in ascending order, but now he or she must alternate between numbers and letters (i.e., 1, A, 2, B, 3, C, ...). Therefore, part B requires the individual to maintain his or her place in both the alphabetical and numerical series while simultaneously remembering whether a letter or number should be next in the series. The primary dependent measure for this study was total time to complete each part of the task. Previous studies have reported high alternate-form reliability ($r = .89-.92$; Charter, Adkins, Alekoubides, & Seacat, 1987) and adequate test-retest reliability ($r = .66-.86$; Goldstein & Watson, 1989) for these scores.

Verbal working memory. The Sentence Span task is a working memory measure that was adapted by Siegel and Ryan (1989) from the procedure developed by Daneman and Carpenter (1980). The participant is instructed to provide the last word for a set of simple sentences read by the examiner (e.g., “I throw the ball up and then it comes. ...”) and is told that he or she will be asked to reproduce the words that he or she provided after all sentences in that set have been completed. The task begins with a block of three two-sentence sets and increases in difficulty by adding one additional sentence per block up to a total of six sentences. The dependent measure is the number of sets completed correctly. This measure has adequate internal reliability (for the composite score, $\alpha = 0.76$) and test–retest reliability ($r = .71$; Kuntsi, Stevenson, Oosterlaan, & Sonuga-Barke, 2001).

The Counting Span task (Case, Kurland, & Goldberg, 1982) is a second measure of verbal working memory that uses a procedure similar to the Sentence Span task. The participant is instructed to count aloud the number of yellow dots dispersed randomly on a set of $8\frac{1}{2} \times 11$ -in. cards. After all cards in a set are completed, the participant is asked to recall, in temporal order, the number of yellow dots that appeared on each of the cards in the set. Similar to the Sentence Span task, there are five blocks with three sets per block (2–6 cards per set). The dependent variable is the total number of correct sets ($\alpha = 0.81$, test–retest $r = .67$; Kuntsi, Stevenson, et al., 2001).

The WISC–R Arithmetic subtest (Wechsler, 1974) requires the participant to solve a series of verbally presented arithmetic problems without using a pencil and paper. To solve each problem correctly, the participant must retain and manipulate in memory the information provided by the examiner. Therefore, in addition to assessing basic math computation abilities, this task provides another measure of verbal working memory. The split-half reliability of the Arithmetic subtest is .77.

The first half of the WISC–R Digit Span subtest (Wechsler, 1974) is a simple short-term verbal memory task. The examiner reads a series of digits that increases in length with each trial, and the participant repeats the digits verbatim. The second half of the task is similar, with the exception that the participant must now repeat the digits in the reverse of the order in which they were presented by the examiner. Because the digits must be retained in memory and manipulated to reverse their order, the digits backward component of the Digit Span task is frequently interpreted as a verbal working memory task (e.g., Rucklidge & Tannock, 2002).

Spatial working memory. The Cambridge Neuropsychological Test Automated Battery (CANTAB) includes a self-ordered spatial working memory test that requires the participant to press the computer touch-screen to find tokens hidden under some of several boxes displayed on the screen. Prior to beginning the task, the participant is told that after each token is found it will never again appear under the same box during that trial. Therefore, to perform the task in the most efficient manner the participants must remember which squares they have chosen and

inhibit responses to these locations. This task has been shown to activate both dorsal and ventral prefrontal regions (Owen, Doyon, Petrides, & Evans, 1996) and is similar to the self-ordered pointing task that has been shown to be sensitive to frontal lesions in adults (Petrides & Milner, 1982). Total errors is the primary dependent variable.

Naming speed–interference control. The Stroop Color and Word Test (Golden, 1978) measures the participant's ability to respond selectively to one dimension of a multidimensional stimulus. On each of the three trials of the Stroop the participant completes as many of the stimuli as possible within 45 sec. The Word trial requires the participant to read the names of colors (*red*, *blue*, and *green*) printed in black ink. On the subsequent Color trial the participant is asked to name the color of nonlinguistic colored patches of red, blue, and green ink. Finally, the stimuli for the Color–Word trial are the words *red*, *blue*, and *green* printed in a different color of ink, and the participant is told to name the color of the ink of each item (and therefore not to read the word). The three trials of the Stroop are hypothesized to assess reading speed, color-naming speed, and interference control. However, although a low number of positive responses on the Color–Word trial could indicate a specific deficit in interference control, a low score on this trial could also reflect a more general deficit in naming speed that cuts across all three trials of the task. Therefore, an interference-control score was operationalized for each participant by subtracting their mean *z* score on the Word and Color trials from their *z* score on the Color–Word trial.

Processing speed. The WISC–R Coding subtest (Wechsler, 1974) and Wechsler Intelligence Scale–Third Edition (WISC–III) Symbol Search subtest (Wechsler, 1991) are more specific measures of processing speed that have been shown to be associated with ADHD in previous studies (e.g., Chhabildas et al., 2001; Hinshaw et al., 2002; Rucklidge & Tannock, 2002). Studies of the psychometric characteristics of the WISC–R and WISC–III indicate that these subtests have adequate reliability (test–retest $r = .72$ for WISC–R Coding and $.74$ for WISC–III Symbol Search) but correlate relatively modestly with FSIQ (Wechsler, 1974, 1991), suggesting that they may tap aspects of processing speed that are at least partially independent of general intelligence.

The Coding subtest requires the participant to rapidly copy symbols associated with specific digits based on a key provided at the top of the page. Although Coding is typically described as a measure of processing speed (e.g., Wechsler, 1991), optimal performance also requires short-term recall of the symbol–digit pairings and adequate fine motor skills. The dependent measure is the total number of correct items after 2 min.

The Symbol Search subtest (Wechsler, 1991) requires the participant to match a symbol to an identical target that is displayed among several distracter stimuli that

share some physical features. The dependent measure is the number of correct items minus the number of incorrect items completed before the 2-min time limit.

DATA ANALYSES

Data Adjustments

As expected, correlational analyses revealed that performance on all neuropsychological variables improved as a linear function of age ($p < .01$ for all measures). Therefore, to control for the influence of age on any of the results, an age-adjusted score was created for each measure by regressing the variable onto age and age squared and then saving the residual score. The distribution of each age-adjusted variable was then assessed for outliers prior to any additional analyses. *Outliers* were defined as scores that fell more than 3 *SDs* from the mean of the overall sample and more than 0.5 *SD* beyond the next most extreme score. After confirming that these outlying scores were entered correctly in the data file, each outlier was adjusted to a score 0.5 *SD* units beyond the next highest score, with multiple outliers rescored to 0.1 *SD* apart. Adjustments were made to Trails A time for one individual in the comparison group, to Trails B time for one individual in the comparison group, two individuals in the RD-only group, and 1 participant in the RD+ADHD group, and to CPT commission errors for 2 individuals in the RD-only group, 3 participants in the ADHD-only group, and 3 participants in the RD+ADHD group. After these adjustments, the distribution of each variable was assessed for significant deviation from normality. A logarithmic transformation was implemented to approximate a normal distribution for variables with skewness or kurtosis greater than 1 (Trails A and B, CPT commissions and omissions, and WCST Perseverative errors).

Factor Analysis of the Neurocognitive Measures

As noted previously, results from several studies suggest that this battery of neuropsychological measures may reflect more than one latent dimension of neurocognitive functioning (e.g., Mariani & Barkley, 1997; Miyake et al., 2000; Pennington, 1997; Willcutt, Pennington, et al., 2001). Therefore, after data-cleaning procedures were completed, an EFA was conducted to clarify the relations among the tasks.

Principal axis extraction and direct oblimin rotation were used to extract factors with eigenvalues greater than one. The direct oblimin rotation was used because it is an oblique rotation that permits the obtained factors to correlate and, therefore, requires fewer a priori assumptions about the relations among the variables than does an orthogonal method of rotation. However, the same number of factors and

similar factor loadings were obtained when a principal components analysis with varimax rotation was conducted, suggesting that these results are robust across different methods of factor extraction and rotation.

The Stroop interference-control score and mean reaction time on primary stop-signal task trials (trials without a stop signal) did not load above .30 on any obtained factor when they were included in the initial factor analysis. Therefore, these two measures were dropped from the final factor analysis summarized in Table 2. Results revealed five factors with eigenvalues greater than one; these factors were labeled Reading and Language Skills, Processing Speed, Verbal Working Memory, Set Shifting, and Response Inhibition–Execution. All five reading and language component processes loaded on the first factor, measures of verbal naming speed and general processing speed loaded on the second factor, all measures of verbal short-term and working memory loaded together on the third factor, and the measures of response inhibition (CPT commission errors and SSRT) and response execution (CPT omission errors and Go trial reaction time) loaded together on the fifth factor.

The CANTAB spatial working memory task loaded primarily on the set-shifting factor rather than the verbal working memory factor. This may be due to the fact that in addition to the working memory demands of the task, it also requires the individual to shift set to avoid responding to a stimulus that was previously rewarded (i.e., to avoid going back to the same box in which a reward previously appeared). The tasks on the set-shifting factor also may load together because they require more extensive visual processing than many of the other tasks in the battery.

Primary Analyses

Due to the relatively high number of comparisons necessary to examine the performance of the four groups on the extensive battery of neuropsychological measures, an alpha of .01 was adopted as the threshold for statistical significance, and p values between .05 and .01 are described as marginally significant. To test for a double dissociation between RD and ADHD on each of the neuropsychological measures, a 2×2 (RD \times ADHD) factorial analysis of variance (ANOVA) was conducted for each individual cognitive measure and each of the five factor scores. If the initial ANOVA revealed a significant main effect of RD or ADHD or a significant RD \times ADHD interaction, planned post hoc comparisons were conducted among the four groups with a Bonferroni correction for multiple comparisons.

Implications of group differences in intelligence. The *DSM-IV* criteria for RD specify that reading achievement scores must fall significantly below both the score typical of other children the same age and the score that would be expected based on the individual's overall cognitive ability. However, a growing

TABLE 2
Principal Axis Factor Analysis of Neurocognitive Measures

Measure	Factor Loadings on the Five Extracted Factors				
	Reading and Language Skills	Processing Speed	Verbal Working Memory	Set Shifting	Response Inhibition-Execution
Lindamood	.51	—	—	—	—
Orthographic Coding	.47	—	—	—	—
Nonword Reading	.87	—	—	—	—
Phoneme Deletion	.83	—	—	—	—
Pig Latin	.87	—	—	—	—
Stroop Word	.35	.60	—	—	—
Stroop Color	—	.63	—	—	—
Stroop Color-Word	—	.55	.23	—	—
WISC-R Coding	—	.51	—	—	—
WISC-III Symbol Search	—	.42	—	—	.25
Counting Span	—	—	.56	—	—
Sentence Span	—	—	.55	—	—
WISC-R Arithmetic	.26	—	.45	—	—
WISC-R Digits Forward	.21	—	.49	—	—
WISC-R Digits Backward	—	—	.61	—	—
CANTAB Spatial Working Memory	—	—	—	.56	—
Trails Part A	—	.39	—	.36	—
Trails Part B	—	.20	—	.51	—
WCST Perseverative Errors	—	—	—	.46	—
CPT commission errors	—	—	—	—	.66
CPT omission errors	—	—	—	—	.65
Stop-Signal go-trial reaction time variability	—	—	—	—	.63
Stop-Signal reaction time	—	—	—	—	.53
Eigenvalue	8.17	1.75	1.42	1.28	1.07
Percentage of variance explained	35.81	7.58	6.17	5.61	4.61

Note. Em dash (—) indicates factor loading less than .20. Loadings in bold type indicate primary factor loading. WISC-III = Wechsler Intelligence Scale for Children-Third Edition; WISC-R = Wechsler Intelligence Scale for Children-Revised; CANTAB = Cambridge Neuropsychological Test Automated Battery; WCST = Wisconsin Card Sorting Test; CPT = Continuous Performance Test.

number of studies indicate that the same etiological factors and neurocognitive deficits are associated with RD, with and without an IQ discrepancy, suggesting that the inclusion of an IQ discrepancy as a diagnostic criterion adds little to the external validity of the diagnosis of RD (e.g., Pennington, Gilger, Olson, & DeFries, 1992; Siegel, 1989).

Nonetheless, the mean FSIQ scores of groups with and without RD are nearly always significantly different even when RD is defined by an age-discrepancy criterion alone (e.g., Rucklidge & Tannock, 2002), and the mean FSIQ of groups with *DSM-IV* ADHD typically falls 0.75 – 1.0 *SD* below the mean of a comparison group without ADHD (e.g., Chhabildas et al., 2001; Hinshaw, 2002; Lahey et al., 1998; Nigg et al., 2002). Therefore, it is often unclear whether neuropsychological weaknesses are associated with RD or ADHD per se or are attributable to these significant group differences in intelligence. Similarly, because few previous studies have controlled for internalizing and externalizing symptoms that frequently co-occur with RD and ADHD, it is also not clear whether the neurocognitive correlates of RD and ADHD may be attributable in part to these comorbid symptoms.

Based on the significant associations between RD and ADHD and these potential confounding variables, some researchers argue that FSIQ and symptoms of comorbid disorders should always be statistically controlled to ensure that neuropsychological impairments associated with RD or ADHD cannot be explained more parsimoniously by group differences on these correlated variables (e.g., Lahey et al., 1998; Werry, Elkind, & Reeves, 1987). On the other hand, ADHD symptoms or reading difficulties may directly cause a child to perform poorly on standardized tests of intelligence (e.g., Barkley, 1997a) or may precipitate the development of these comorbid disorders. In these cases, it would not be appropriate to control for these variables, as this would remove variance that is associated with ADHD or RD.

These issues have not been resolved conclusively, and the optimal approach is likely to vary depending on the specific research question. Therefore, with the exception of the exclusion of participants with FSIQ scores below 75 from all analyses, neither IQ nor symptoms of comorbid mental disorders were considered in the diagnostic algorithms used to define the groups. Instead, we directly tested if the relations between RD and ADHD, and the neuropsychological measures are explained by group differences in intelligence or comorbid symptomatology.

The WISC-R (Wechsler, 1974) was used to assess the FSIQ of participants 16 years of age or younger, and the Wechsler Adult Intelligence Scale-Revised (WAIS-R; Wechsler, 1981) was administered to participants who were 17 or 18 years of age. The *Diagnostic and Statistical Manual of Mental Disorders* (3rd ed., rev. [*DSM-III-R*]; American Psychiatric Association, 1987) parent-report version of the Diagnostic Interview for Children and Adolescents (DICA-P; Reich & Welner, 1988) was administered to parents to assess symptoms of oppositional defiant disorder, conduct disorder, generalized anxiety disorder, and major depres-

sive disorder. If initial analyses revealed a significant main effect of RD or ADHD on a variable, a parallel analysis of covariance (ANCOVA) was conducted with FSIQ, symptoms of other disorders, SES included as dimensional covariates, and sex included as a dichotomous categorical covariate. Because comorbid symptomatology, SES, and sex did not change the significance of the primary independent variables for any neurocognitive measure, these variables were dropped from the final models described in this article.

Dimensional analyses. Comparisons among the group means test if individuals who meet criteria for ADHD or RD exhibit neuropsychological weaknesses that are independent of the influence of the other categorical diagnosis. However, a growing body of research suggests that categorical diagnostic cutoffs for RD and ADHD artificially dichotomize what may be a continuous dimension of liability (e.g., Barkley, 1998; DeFries & Alarcón, 1996; Levy et al., 1997; Willcutt et al., 2000a). Consistent with this hypothesis, the RD-only group exhibited a higher number of symptoms of ADHD than the comparison group, and the group with ADHD alone scored significantly lower than the comparison group on measures of reading achievement. Therefore, any neurocognitive weaknesses attributed to RD or ADHD could potentially be explained by subclinical elevations of the other disorder rather than by the primary disorder per se (e.g., Nigg et al., 1998). To test this possibility, a series of multiple-regression analyses were conducted in which the score on each neurocognitive measure was regressed onto the continuous-reading discriminant function score and the total number of *DSM-IV* inattention and hyperactivity-impulsivity symptoms.

DSM-IV ADHD subtypes. To examine differences between the inattentive and combined types, comparisons among the four groups were conducted when the ADHD and RD+ADHD groups included only those individuals who met criteria for the inattentive type (i.e., comparison group vs. RD-only vs. inattentive type-only vs. RD+inattentive type), then repeated with only the individuals who met criteria for the combined type. Consistent with our previous analyses of this sample (Chhabildas et al., 2001), the pattern of results was virtually identical for the two subtypes. Therefore, to simplify interpretation both subtype groups were included in a single group with ADHD for the analyses described in this article.

RESULTS

Table 3 presents the unadjusted means of the four groups on each neuropsychological measure and the five factor scores. Results revealed a significant main effect of RD for all measures except WCST perseverative errors and Stroop interference control. Similarly, the ADHD main effect was significant for most individ-

TABLE 3
Unadjusted Means of the Four Groups on Measures of Neuropsychological Functioning

Variables	Comparison (n = 151)		ADHD Only (n = 113)		RD Only (n = 109)		RD + ADHD (n = 64)		Main Effects		Group Differences That Remain Significant When IQ is Controlled ^b
	M	SD	M	SD	M	SD	M	SD	ADHD	RD	
	Component Reading and Language Skills										
Lindamood	86.6 _a	13.6	77.8 _b	18.4	64.2 _c	19.1	59.6 _c	21.3	16.27*** ^b	152.03***	R, R + A < A, C
Nonword Reading	0.12 _a	0.95	-0.44 _b	1.16	-2.54 _c	1.15	-2.56 _c	1.19	14.75*** ^b	467.07***	R, R + A < A, C
Orthographic Coding	86.2 _a	9.1	77.4 _b	11.7	72.2 _c	11.8	70.0 _c	13.5	31.58***	305.44***	R, R + A < A < C
Phoneme Deletion	80.1 _a	15.0	72.5 _b	17.2	50.3 _c	23.6	50.2 _c	22.8	7.29** ^b	341.70***	R, R + A < A, C
Pig Latin	83.2 _a	12.7	75.6 _b	17.9	53.3 _c	28.3	47.0 _c	30.7	10.91*** ^b	201.77***	R, R + A < A, C
Factor score	0.75 _a	0.53	0.31 _b	0.62	-0.72 _c	0.75	-0.81 _c	0.81	16.46*** ^b	299.28***	R, R + A < A, C
Processing Speed											
Stroop Word	83.6 _a	14.4	75.4 _b	16.1	67.2 _c	16.5	65.8 _c	17.9	16.99***	119.52***	R, R + A < A < C
Stroop Color	58.2 _a	12.3	51.1 _b	14.6	49.2 _b	12.4	48.9 _b	14.9	18.95***	33.70***	A, R, R + A < C
Stroop Color-Word	34.0 _a	10.3	27.6 _b	10.7	26.5 _b	8.5	25.8 _b	9.0	24.86***	38.58***	A, R, R + A < C
Stroop interference-control	0.01	0.91	0.17	0.90	-0.12	0.70	-0.11	0.83	1.81	4.01	None
WISC-R Coding	10.9 _a	2.6	8.4 _b	2.8	8.3 _{bc}	2.7	7.3 _c	2.9	49.96***	55.00***	A, R, R + A < C
WISC-III Symbol Search ^c	13.4 _a	2.7	10.8 _b	3.2	10.2 _b	2.8	9.6 _b	3.7	17.91***	30.75***	A, R, R + A < C
Factor score	0.58 _a	0.71	-0.08 _b	0.69	-0.14 _b	0.70	-0.29 _b	0.92	28.93***	36.26***	A, R, R + A < C
Verbal working memory											
Sentence Span	5.8 _a	2.0	4.7 _b	2.1	4.3 _b	1.6	4.1 _b	1.8	9.18** ^{ad}	26.00***	A ^d , R, R + A < A < C
Counting Span	7.4 _a	2.5	6.2 _b	2.4	5.5 _c	1.8	5.1 _c	2.1	14.12***	51.62***	A, R, R + A < C
WISC-R Arithmetic	11.2 _a	2.8	10.2 _b	2.6	8.2 _c	2.4	7.7 _c	2.6	9.13** ^{ab}	118.90***	R, R + A < A < C

WISC-R Digit Span											
Digits Forward	7.3 _a	1.9	6.7 _b	2.0	5.4 _c	1.7	5.6 _c	1.8	4.22	71.40***	R, R + A < A, C
Digits Backward	6.0 _a	1.9	4.9 _b	1.8	4.3 _c	1.5	4.0 _c	1.4	17.20***	70.45***	R, R + A < A < C
Scaled Score	11.1 _{1a}	2.9	9.7 _b	2.6	8.0 _c	2.3	7.9 _c	2.4	11.04*** _b	102.62***	R, R + A < A < C
Factor score	0.58 _a	0.87	-0.09 _b	0.78	-0.61 _c	0.52	-0.64 _c	0.63	22.56***	109.00***	R, R + A < A < C
Set shifting											
Spatial Working Memory ^c	37.2 _a	16.4	45.6 _b	18.4	46.2 _b	19.6	49.1 _b	18.9	4.99** _b	6.44** _b	None
Trailmaking Test Part A	17.3 _a	7.9	20.9 _b	9.1	21.7 _b	8.7	22.7 _b	9.1	8.13**	13.52*** _b	A < C
Trailmaking Test Part B	36.3 _a	19.4	48.7 _b	25.0	52.6 _b	35.0	53.4 _b	38.0	8.96*** _b	18.67*** _b	None
WCST Perseverative Errors	12.3 _a	6.3	15.0 _b	8.4	15.1 _b	8.5	16.9 _b	9.7	5.50** _b	6.69** _b	None
Factor score	0.19 _a	0.34	-0.02 _b	0.49	-0.11 _b	0.51	-0.21 _b	0.64	10.46*** _b	24.64*** _d	R + A < C ^d
Response execution-inhibition											
CPT commission errors	5.1 _a	8.8	12.7 _b	16.0	13.4 _b	20.4	18.9 _c	26.0	18.10***	20.74***	A, R, R + A < C
CPT omission errors	4.9 _a	5.1	8.0 _b	7.3	9.8 _b	7.7	10.2 _b	7.3	9.94** _d	34.95*** _d	R + A < C ^d
Stop-Signal go-trial RT	660.2 _a	117.6	671.6 _a	121.3	706.5 _b	130.0	732.3 _b	125.2	1.29	17.28***	R, R + A < A, C
Stop-Signal no-trial RT SD	178.4 _a	50.5	218.3 _b	67.9	232.9 _b	80.9	239.3 _b	79.2	17.21***	40.49***	A, R, R + A < C
Stop-Signal reaction time	281.2 _a	115.2	340.1 _b	124.8	357.2 _b	151.5	382.8 _b	164.2	15.14***	21.26***	A, R, R + A < C
Factor score	0.42 _a	0.56	-0.22 _b	0.76	-0.45 _{bc}	0.76	-0.63 _c	0.68	30.09***	82.91***	A, R, R + A < C

Note. Means with no common subscripts are significantly different when Full Scale IQ (FSIQ) is not controlled. A less-than sign always indicates poorer performance regardless of the original scaling of the measure. ADHD = attention deficit hyperactivity disorder; RD = reading disorder; WISC-R = Wechsler Intelligence Scale for Children-Revised; WISC-III = Wechsler Intelligence Scale for Children-Third Edition; WCST = Wisconsin Card Sorting Test; CPT = Continuous Performance Test; RT SD = reaction time standard deviation.

^aC = comparison; R = RD only; A = ADHD only; R + A = RD + ADHD. ^bIndicates that the main effect is no longer significant when FSIQ is covaried. ^cMeasure administered to a subset of the total sample; control, *N* = 79; ADHD only, *N* = 44; RD only, *N* = 44; RD + ADHD, *N* = 21. ^dMarginally significant when FSIQ was covaried (*p* < .05).

p* < .05. *p* < .01. ****p* < .001.

ual measures and all five factor scores. RD \times ADHD interactions were not significant for any variable, and the comorbid group was at least as impaired as the RD-only and ADHD-only groups on all measures in the battery. Because no interactions were significant, interaction terms were dropped from all final models.

The overall pattern of results suggests that RD and ADHD are both associated with weaknesses on most measures in the battery of neurocognitive tasks. On the other hand, post hoc comparisons among the four groups and the relative size of the RD and ADHD main effects revealed important differences between the neuropsychological correlates of the two disorders. Both groups with RD scored significantly lower than the ADHD-only group on all measures of component reading and language skills, and the effect sizes were much larger for the main effect of RD (e.g., for the Reading and Language skills factor score, $\eta^2 = 0.441$) than the main effect of ADHD ($\eta^2 = 0.046$). Similarly, the mean of the RD-only group was significantly lower than the mean of the ADHD-only group on five of the six verbal working memory tasks, and the RD effect size for the Verbal Working Memory factor score was again substantially larger (for the factor score, $\eta^2 = 0.268$) than the ADHD effect size ($\eta^2 = 0.071$). In contrast, the two groups differed on only one processing speed measure and none of the set-shifting measures, and the RD and ADHD effect sizes were similar for the Processing Speed factor (RD $\eta^2 = 0.108$, ADHD $\eta^2 = 0.099$) and the Set-shifting factor (RD $\eta^2 = 0.058$, ADHD $\eta^2 = 0.042$). Results for the measures on the Response Inhibition–Execution factor were somewhat less consistent; although the means of the RD-only and ADHD-only groups were significantly different on only one measure (Go trial reaction time from the stop-signal task), the mean of the RD-only group also fell in the direction of greater impairment on the other four measures, and the effect size for the factor score was somewhat larger for RD ($\eta^2 = 0.208$) than for ADHD ($\eta^2 = 0.113$).

Group Comparisons Controlling FSIQ

Because FSIQ was significantly related to scores on all measures, a series of 2×2 (RD \times ADHD) ANCOVAs were conducted to test if RD or ADHD was significantly associated with performance on each task after controlling for FSIQ differences among the groups (last column of Table 3). The main effect of RD remained significant for all measures except the set-shifting tasks, suggesting that most of the neurocognitive difficulties associated with RD are not explained by group differences in FSIQ. The significant ADHD main effect on the set-shifting measures was also eliminated when FSIQ was controlled, as were the main effects on several of the reading, language, and verbal working memory measures. In contrast, the main effect of ADHD remained significant for the orthographic coding task, counting span, and all measures of processing speed and response execution–inhibition.

Multiple-Regression Analyses

When FSIQ was not included in the model, the reading discriminant function score was associated with significantly lower scores on all neuropsychological measures with the exception of Stroop interference control (Table 4). When FSIQ was added to the model the association with reading performance remained significant for measures of component reading and language skills, verbal working memory, most measures of processing speed, and SSRT and reaction time variability on the stop-signal task. In contrast, reading performance was no longer significantly associated with the set-shifting measures or CPT commission errors, suggesting that the relation between RD and these variables may be mediated by the association between reading and FSIQ.

The multiple-regression analyses indicated that the neuropsychological weaknesses associated with *DSM-IV* ADHD are primarily attributable to elevations of inattention symptoms rather than hyperactivity–impulsivity symptoms. Inattention was significantly associated with lower scores on all measures of processing speed, as well as lower scores on the other four factors and a subset of the measures in each of these domains. Interestingly, although the inattention symptom count was not significantly associated with SSRT, SSRT increased significantly as a function of the total number of ADHD symptoms (i.e., the sum of inattention and hyperactivity–impulsivity), suggesting that the combined effects of inattention and hyperactivity–impulsivity may lead to inhibitory difficulties on this measure. When FSIQ was added to each model, all associations that were significant in the initial model remained significant with the exception of spatial working memory and WCST perseverative errors, providing further evidence that the association between ADHD and set shifting may be explained more parsimoniously by the association between inattention symptoms and intelligence.

DISCUSSION

This study compared the neurocognitive correlates of RD and ADHD in a large population-based sample of twins. The primary goal of the study was to clarify the nature of the relation between RD and ADHD by testing whether these neuropsychological functions were associated with RD, ADHD, or both disorders, and to identify shared neuropsychological deficits that might provide useful endophenotypes for molecular genetic studies of the etiology of comorbidity between RD and ADHD.

An initial EFA of the neurocognitive tasks revealed five latent dimensions, which were labeled Reading and Language Skills, Processing Speed, Verbal Working Memory, Set-Shifting, and Response Inhibition–Execution. The last three factors suggest that EF tasks reflect more than one latent dimension of

TABLE 4
Multiple Regression Models Predicting Neuropsychological Variables From Reading Scores and Number of ADHD Symptoms

<i>Variables</i>	<i>Overall Model R²</i>			<i>Inattention Symptoms</i>			<i>Hyperactivity-Impulsivity Symptoms</i>			<i>Reading Composite Score</i>			<i>Significant Independent Predictors When FSIQ Was Included in the Model</i>	
	<i>b</i>	<i>SE</i>	<i>t</i>	<i>b</i>	<i>SE</i>	<i>t</i>	<i>b</i>	<i>SE</i>	<i>t</i>	<i>b</i>	<i>SE</i>	<i>t</i>		
<i>Component reading and language skills</i>														
Lindamood	.416	.004	.019	.021	.023	-2.13*	-.049	.023	-.049	.023	.522	.036	14.66***	Reading
Nonword Reading	.656	.012	.015	0.78	.022	-1.21	-.027	.022	-.027	.022	.874	.029	30.36***	Reading
Orthographic Coding	.472	-.041	.016	-2.61**	.022	-1.64	-.036	.022	-.036	.022	.495	.032	15.46***	Inattention, reading
Phoneme Deletion	.499	.034	.018	1.89	.023	-1.69	-.038	.023	-.038	.023	.700	.034	20.29***	Reading
Pig Latin	.427	.023	.024	0.91	.030	-1.29	-.039	.030	-.039	.030	.754	.046	16.51***	Reading
Factor score	.693	.008	.018	0.44	.022	-0.86	-.019	.022	-.019	.022	.498	.019	26.22***	Reading
<i>Processing Speed</i>														
Stroop Word	.347	-.043	.014	-3.02**	.018	1.32	.020	.018	.020	.018	.345	.026	13.03***	Inattention, reading
Stroop Color	.181	-.061	.016	-3.81***	.021	0.91	.018	.021	.018	.021	.183	.030	5.99***	Inattention, reading
Stroop Color-Word	.189	-.067	.016	-4.16***	.021	-0.21	-.004	.021	-.004	.021	.204	.030	6.86***	Inattention, reading
Stroop interference control	.061	-.036	.017	-2.09*	.022	-1.02	-.022	.022	-.022	.022	.057	.030	1.78	None
WISC-R Coding	.220	-.094	.015	-6.34***	.019	1.01	.019	.019	.019	.019	.191	.028	6.43***	Inattention
WISC-III Symbol Search	.208	-.052	.022	-2.38**	.029	-1.17	-.034	.029	-.034	.029	.222	.042	5.26***	Inattention*
Factor score	.285	.099	.014	7.07***	.019	-2.42*	-.046	.019	-2.42*	.019	-.173	.027	-6.45***	Inattention, reading
<i>Verbal working memory</i>														
Counting Span	.218	-.043	.015	-2.81**	.019	0.57	.011	.019	.011	.019	.269	.029	9.24***	Inattention*, reading
Sentence Span	.165	-.016	.016	-1.01	.020	-0.26	-.005	.020	-.005	.020	.250	.030	8.35***	Reading
WISC-R Arithmetic	.380	-.015	.013	-1.16	.017	0.18	.003	.017	.003	.017	.387	.025	15.41***	Reading

WISC-R Digit Span												
Digits Forward	.265	.003	.014	0.21	-.011	.018	-.061	.327	.028	11.67***	Reading	
Digits Backward	.283	-.028	.014	-2.01*	-.017	.019	-0.96	.313	.027	11.47***	Reading	
Scaled score	.357	-.019	.013	-1.46	-.019	.017	-1.12	.372	.026	14.31***	Reading	
Factor score	.478	-.033	.013	-2.54*	-.005	.017	-0.29	.374	.025	14.76***	Inattention*, reading	
Set Shifting												
Spatial Working Memory	.090	.021	.024	0.87	.043	.031	1.39	-.143	.044	-3.21**	None	
Trailmaking Test Part A	.122	.067	.017	4.01***	-.061	.022	-2.77**	-.133	.031	-4.24***	Inattention, reading*	
Trailmaking Test Part B	.095	.046	.017	2.77**	-.036	.020	-1.79	-.158	.032	-5.02***	Inattention*	
WCST Perseverative Errors	.044	.022	.018	1.22	.020	.021	0.98	-.078	.033	-2.35**	None	
Factor score	.139	.031	.010	3.11***	-.017	.012	-1.42	-.099	.018	-5.51**	Inattention*	
Response Execution/Inhibition												
CPT commission errors	.123	.046	.017	2.77**	.008	.021	0.37	-.015	.031	-4.83***	Inattention	
CPT omission errors	.141	.049	.016	2.99***	-.032	.020	-1.62	-.203	.031	-6.64***	Inattention*, reading*	
Stop-Signal go trial RT	.039	.029	.018	1.66	-.029	.022	-1.32	-.098	.033	-2.95**	None	
Stop-Signal go trial RT SD	.169	.044	.017	2.69***	-.001	.020	-0.04	-.223	.031	-7.20***	Inattention, reading	
Stop-Signal reaction time	.117	.029	.017	1.71	.024	.022	1.08	-.175	.034	-5.12***	ADHD ^a , reading	
Factor score	.370	.049	.013	3.79***	-.016	.016	-0.99	-.262	.024	-11.03***	Inattention, reading	

Note. ADHD - attention deficit hyperactivity disorder; FSIQ = Full Scale IQ; WISC-R = Wechsler Intelligence Scale for Children-Revised; WISC-III = Wechsler Intelligence Scale for Children-Third Edition; WCST = Wisconsin Card Sorting Test; CPT = Continuous Performance Test.

^aSignificant when the total ADHD symptom score was used. Not significant for either inattention or hyperactive-impulsive symptoms when entered separately.

* $p < .05$. ** $p < .01$. *** $p < .001$.

neurocognitive functioning, a result that is consistent with previous studies (Mariani & Barkley, 1997; Miyake et al., 2000; Pennington, 1997; Willcutt, Pennington, et al., 2001).

Neurocognitive Correlates of ADHD

These results provide additional support for the hypothesis that ADHD is associated with a significant weakness in response inhibition (e.g., Barkley, 1997a; Nigg, 2000, 2001; Schachar et al., 2000). Categorical comparisons between groups with and without ADHD and multiple-regression analyses of continuous data indicated that ADHD was associated with slower SSRT and an elevated rate of commission errors on the CPT. Moreover, the ADHD main effect was larger for the factor composed of the measures of response inhibition and execution than for the other four factors.

In addition to weaknesses on measures of response inhibition, ADHD was associated with significantly lower scores on the orthographic coding task, and this weakness remained significant even after comorbid RD and group differences in FSIQ were controlled. The ADHD groups also performed more poorly on a subset of the verbal working memory tasks, were slower on most measures of processing speed, and exhibited significantly more variable reaction times on the simple forced-choice primary task trials of the stop-signal task. The slower and more variable response speed in the ADHD groups suggests that children with ADHD may have difficulty sustaining sufficient cognitive activation, a finding that is consistent with the predictions of the cognitive-energetics model (e.g., Sergeant, 2000; Sergeant, Oosterlaan, & van der Meere, 1999).

Taken together, these results support the hypothesis that ADHD is associated with a deficit in response inhibition (e.g., Barkley, 1997a), but suggest that additional weaknesses in both EF and non-EF domains are also important facets of the neuropsychology of ADHD. Moreover, our battery did not include measures of planning, delay aversion, or motor output, all of which have been shown to be associated with ADHD in previous studies (e.g., Aman, Roberts, & Pennington, 1998; Mariani & Barkley, 1997; Douglas, 1999; Solanto et al., 2001; Sonuga-Barke, Taylor, Sembi, & Smith, 1992; Weyandt & Willis, 1994). Future studies are needed that administer a comprehensive battery of EF measures along with measures of each of these domains to a single sample of participants with ADHD. Such studies will clarify the relations among these measures and will facilitate a direct test of the relative importance of each of these domains as part of the overall neuropsychology of ADHD.

Neurocognitive Correlates of RD

As expected, based on previous studies (Gayán & Olson, 2001; Olson et al., 1989, 1994; Rack et al., 1992; Willcutt, Pennington, et al., 2001), the groups with RD ex-

hibited a marked weakness on all measures of component reading and language skills (RD vs. control effect size: $d = 1.55$ – 2.30 ; Cohen, 1988). These results provide further support for the hypothesis that difficulties with phonological processing and related language skills play a critical role in the etiology of RD. In addition to the predicted deficits on the measures of reading and language processes, the RD groups exhibited significant impairment on several EF and processing speed tasks. These weaknesses were most pronounced on measures of processing speed and verbal working memory, but they were also significant for measures of response inhibition such as SSRT and CPT commission errors.

The finding that RD is associated with slower processing speed and shorter verbal working memory span replicates other recent studies that used the same or similar measures (e.g., Purvis & Tannock, 2000; Roodenrys et al., 2001; Rucklidge & Tannock, 2002; Willcutt, Pennington, et al., 2001) and is consistent with studies that used more specific measures of processing speed such as rapid automatized naming (e.g., Compton et al., 2002; Felton, Wood, Brown, Campbell, & Harter, 1987; Rucklidge & Tannock, 2002; Semrud-Clikeman et al., 2000). In contrast, previous studies of the relation between RD and response inhibition are less consistent. Whereas Purvis and Tannock (2000) reported a significant association between RD and SSRT, two other studies failed to find this association (Rucklidge & Tannock, 2002; Willcutt, Pennington, et al., 2001). However, Willcutt, Pennington, et al. (2001) reported that the association between RD and SSRT was significant if FSIQ was not controlled, and the effect size for the mean difference between the RD and comparison group was moderate ($d = 0.5$) in the study by Rucklidge and Tannock (2002), suggesting that this difference might be significant in a larger sample ($n = 12$ individuals with RD in this study). In combination, these findings suggest that additional research is needed to test definitively if RD is associated with mild deficits on response inhibition tasks and, if so, whether these deficits reflect the same underlying processes that lead to disinhibition in individuals with ADHD.

Implications for Comorbidity Between RD and ADHD

The two groups with RD exhibited significantly larger weaknesses on the measures of component reading and language skills than the group with ADHD. Similarly, although both RD and ADHD were associated with verbal working memory difficulties, these weaknesses were more consistent and severe in the two groups with RD. In contrast, the RD-only and ADHD-only groups did not differ on the response inhibition, processing speed, or set-shifting measures. Thus, whereas comparisons among the four groups reveal a clear dissociation between the groups with and without RD on the measures of reading and language skills, none of the neuropsychological tasks were associated specifically with ADHD. In contrast to previous findings that revealed a double dissociation between RD and ADHD on measures of phonological processing versus EFs (e.g., Pennington et al., 1993;

Willcutt, Pennington, et al., 2001), these results suggest that this dissociation may not be complete.

The profile of the RD+ADHD group is consistent with the additive combination of the deficits of the RD-only and ADHD-only groups, and the RD+ADHD group did not exhibit a specific weakness on any measures that were not impaired in at least one of the other two groups. This pattern of results replicates the findings in an earlier independent sample from the CLDRC (Willcutt, Pennington, et al., 2001) and in studies conducted by several other groups (e.g., Dykman & Ackerman, 1991; Klorman et al., 1999; Nigg et al., 1998; Purvis & Tannock, 2000; Seidman et al., 2001). In contrast, these results appear to fail to replicate the finding that the comorbid group is most impaired on measures of letter and color naming (Rucklidge & Tannock, 2002). However, two caveats suggest that additional research is needed to provide a definitive test of this hypothesis. First, our study did not include measures of rapid automatized naming, and these tasks revealed the largest effects in the study by Rucklidge and Tannock. Second, although the mean of the RD+ADHD group was not significantly different from the mean of the RD group on any measure of naming speed, the mean of the comorbid group was nearly always in the direction of greater impairment on these tasks and the other measures that loaded on the processing speed factor.

In summary, these results suggest that the double dissociation between RD and ADHD is not complete, and they indicate that the neuropsychological weaknesses exhibited by the comorbid group reflect the additive combination of the deficits associated with RD and ADHD alone. The failure to find a clean double dissociation between RD and ADHD is inconsistent with the cross-assortment hypothesis, and recent data from family studies provide additional converging evidence to reject this model as an explanation for most cases of comorbid RD+ADHD (Doyle et al., 2001; Friedman et al., 2003). The absence of any significant RD \times ADHD interactions in this sample and several other recent studies (e.g., Rucklidge & Tannock, 2002; Seidman et al., 2001; Willcutt, Pennington, et al., 2001) provide strong converging evidence that the phenocopy hypothesis can be rejected as the primary explanation for comorbidity between RD and ADHD. This result also argues against the cognitive subtype hypothesis, although the possibility remains that the RD+ADHD group could exhibit a unique neuropsychological weakness on a measure that was not included in our battery.

As noted previously, the most parsimonious model of a common genetic etiology for RD and ADHD would include a common risk allele or set of risk alleles at one or more genetic loci that influence a common pathophysiological pathway that increases susceptibility to both RD and ADHD. In this model all three clinical groups (i.e., RD, ADHD, and RD+ADHD) should be characterized by at least some common neuropsychological weaknesses, with each individual's final phenotype determined by the other genetic and environmental influences that affect that individual. In this study the groups with RD only, ADHD

only, and RD+ADHD all exhibited deficits in comparison to the group without RD or ADHD on most measures of response inhibition and processing speed, as well as a subset of the measures of verbal working memory and reading and language skills. Therefore, although this study cannot definitively rule out all other explanations for comorbidity (e.g., Neale & Kendler, 1995), this pattern of results is consistent with the predictions of the common etiology hypothesis.

What is the Common Underlying Deficit in RD and ADHD?

If the common genetic etiology hypothesis is correct, it may be possible to identify a neuropsychological marker, or endophenotype, that reflects the common genetic risk for RD and ADHD. Measures of this neurocognitive weakness could then be used to facilitate molecular genetic studies of RD, ADHD, and their comorbidity. Therefore, in the next section we assess the strengths and weaknesses of the measures in our battery as potential endophenotypes for future studies attempting to find the common genes that influence RD and ADHD.

Because phoneme awareness deficits are not present in ADHD and set-shifting deficits are not associated with either disorder when FSIQ is controlled, these tasks can be rejected as the common neurocognitive deficit in RD and ADHD. The verbal working memory tasks and the response inhibition tasks are also less compelling candidates for the common deficit because the relations between ADHD and working memory and RD and inhibition are relatively weak and inconsistent in this study and others (e.g., Kuntsi, Oosterlaan, & Stevenson, 2001; Purvis & Tannock, 2000; Rucklidge & Tannock, 2002; Swanson et al., 1999).

In an intriguing result that we did not anticipate, both RD and ADHD were associated with deficits on the orthographic coding measure. The ADHD main effect was larger for this measure than for any of the executive or processing speed measures, and our previous results revealed that the evidence of bivariate linkage with ADHD was stronger for orthographic coding than for any other measure in our battery of component reading and language skills (Willcutt et al., 2002; Willcutt, DeFries, et al., 2003). Due to the multiple-choice format of the orthographic coding task, the underpinnings of the association between ADHD and difficulties on this task are not clear. For example, the ADHD deficit may reflect a true deficit in the ability to access the orthographic representation of a word in the lexicon, or it could simply indicate that the multiple-choice format of the test leads to more frequent impulsive incorrect responses in the ADHD group. In either case, additional research is warranted to test if this result can be replicated in other samples and with other measures of orthographic coding, and to examine whether the poor performance of individuals with RD versus ADHD is attributable to dysfunction in the same underlying processes.

Processing speed is the most promising candidate in our battery for a neuropsychological deficit that is common to RD and ADHD. All three clinical groups

scored lower than the comparison group on all measures of processing speed and reaction time variability, the effect size versus the comparison group was large for all three groups (e.g., $d = 0.87$ – 1.06 for the factor score), and each of these differences remained significant when FSIQ was controlled. Moreover, other studies have reported similar group differences on some or all of these measures, as well as more specific aspects of processing speed such as speeded naming of letters, numbers, colors, and objects (e.g., Hinshaw et al., 2002; Houghton et al., 1999; Nigg et al., 2002; Rucklidge & Tannock, 2002; Semrud-Clikeman et al., 2000). These promising results suggest that processing speed may be a worthwhile focus for future research on the neuropsychological correlates of RD and ADHD.

Perhaps the greatest weakness of processing speed as a candidate neuropsychological endophenotype returns full circle to our initial discussion of EFs at the beginning of this article: Similar to the construct of EFs, the construct of processing speed is overly broad and weakly defined. Previous studies have used both verbal and nonverbal measures that varied from little or no executive component to a large working memory load (e.g., shifting set or updating working memory while completing a task as quickly as possible). Therefore, future research is needed to clarify the definition of processing speed and the taxonomy of processing speed tasks by examining the correlations among different measures and comparing and contrasting the relation between these measures and different clinical groups. It will be essential for future studies to consider carefully the component processes that are involved in each task and to develop methods to test whether poor task performance reflects the same underlying dysfunction across development and in groups with RD, ADHD, or other developmental disorders.

Limitations and Future Directions

These results should be interpreted in light of several limitations. Due to the time constraints of the overall study, *DSM-IV* ADHD was defined by parent and teacher ratings on the DBRS rather than a full structured diagnostic interview. Although all participants who were included in the ADHD groups were required to meet full *DSM-IV* criteria for ADHD, the possibility remains that the use of rating scales rather than a full structured interview to assess symptoms of ADHD could influence the results. To test this possibility, the parents of 30 children with ADHD and 30 children without ADHD based on the DBRS also completed the *DSM-IV* version of the DICA (Reich, Welner, & Herjanic, 1997). The concordance between diagnoses derived from the DBRS and the DICA-IV was extremely high (97% agreement; $\kappa = 0.93$), suggesting that these methods yield similar results.

Although previous studies have found few significant differences between twins and nontwins (e.g., Plomin, DeFries, McClearn, & Rutter, 1997), the use of twins for phenotypic comparisons may limit the generalization of these findings to the population at large. In addition, because the CLDRC twin project has been on-

going for nearly 20 years, older versions of the WISC and the PIAT than those currently available have been maintained to allow comparisons to be made across the entire sample. Finally, due to time constraints some EF constructs were assessed by a single measure (e.g., spatial working memory). Therefore, results based on these measures should be interpreted with caution, and it would be useful for future studies to administer multiple measures of all key neuropsychological domains to facilitate the creation of more reliable measures of each latent construct.

CONCLUSIONS

Groups with and without RD and ADHD were compared on an extensive battery of measures of EFs, processing speed, and component reading and language skills. Results indicate that both disorders are associated with weaknesses in multiple neuropsychological domains. Deficits in phonological processing were only present in the RD groups, whereas no deficits were uniquely associated with ADHD. The neuropsychological profile of the comorbid group was consistent with the additive combination of the weaknesses identified in the groups with RD and ADHD alone, suggesting that the phenocopy and cognitive subtype hypotheses are not likely to explain comorbidity between RD and ADHD in this sample. Instead, these results are most consistent with the predictions of the common genetic etiology hypothesis and suggest that slow and variable processing speed may be a promising endophenotype for future linkage and association studies of RD and ADHD.

ACKNOWLEDGMENTS

This research was supported by National Institute of Child Health and Human Development center Grant P50 HD-27802 (center director: J.C. DeFries). We were also supported, in part, during the preparation of this manuscript by National Institutes of Health Grants F32 MH 12100, R01 MH62120, and R01 MH63941 (to Erik G. Willcutt); MH 38820 and MH04024 (to Bruce F. Pennington); and R0138526 (to Richard K. Olson).

We thank the Colorado Learning Disabilities Research Center staff and the school personnel and families that participated in the study.

REFERENCES

- 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.
- American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text revision). Washington, DC: Author.
- Barkley, R. A. (1997a). *ADHD and the nature of self-control*. New York: Guilford.
- 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.
- Barkley, R. A., & Murphy, K. (1998). *Attention-deficit hyperactivity disorder: A clinical workbook* (2nd ed.). New York: Guilford.
- Bedard, A., Ickowicz, A., Logan, G. D., Hogg-Johnson, S., Schachar, R., & Tannock, R. (2003). Selective inhibition in children with attention-deficit hyperactivity disorder off and on stimulant medication. *Journal of Abnormal Child Psychology*, *31*, 315–327.
- Bruce, D. J. (1964). The analysis of word sounds by young children. *British Journal of Psychology*, *34*, 158–170.
- Cardon, L. R., Smith, S. D., Fulker, D. W., Kimberling, W. J., Pennington, B. F., & DeFries, J. C. (1994, October 14). Quantitative trait locus for reading disability on chromosome 6. *Science*, *226*, 276–279.
- Cardon, L. R., Smith, S. D., Fulker, D. W., Kimberling, W. J., Pennington, B. F., & DeFries, J. C. (1995, June 16). Quantitative trait locus for reading disability: A correction. *Science*, *268*, 1553.
- Carlson, C. L., Booth, J. E., Shin, M., & Canu, W. H. (2002). Parent-, teacher-, and self-rated motivational styles in ADHD subtypes. *Journal of Learning Disabilities*, *35*, 104–113.
- Caron, C., & Rutter, M. (1991). Comorbidity in child psychopathology: Concepts, issues and research strategies. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *32*, 1063–1080.
- Case, R., Kurland, M., & Goldberg, J. (1982). Operational efficiency and the growth of short-term memory span. *Journal of Experimental Child Psychology*, *33*, 386–404.
- Charter, R. A., Adkins, T. G., Alekoubides, A., & Seacat, G. F. (1987). Reliability of the WAIS, WMS, and Reitan battery: Raw scores and standardized scores corrected for age and education. *International Journal of Clinical Neuropsychology*, *9*, 28–32.
- Chhabildas, N. A., Pennington, B. F., & Willcutt, E. G. (2001). A comparison of the cognitive deficits in the *DSM-IV* subtypes of ADHD. *Journal of Abnormal Child Psychology*, *29*, 529–540.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Compton, D. L., Olson, R. K., DeFries, J. C., & Pennington, B. F. (2002). Comparing the relationships among two different versions of alphanumeric rapid automatized naming and word level reading skills. *Scientific Studies of Reading*, *6*, 343–368.
- Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior*, *19*, 450–466.
- DeFries, J. C. (1985). Colorado reading project. In D. B. Gray & J. F. Kavanagh (Eds.), *Biobehavioral measures of dyslexia* (pp. 107–122). Parkton, MD: York Press.
- DeFries, J. C., & Alarcón, M. (1996). Genetics of specific reading disability. *Mental Retardation and Developmental Disabilities Research Reviews*, *2*, 39–47.
- DeFries, J. C., Filipek, P. A., Fulker, D. W., Olson, R. K., Pennington, B. F., Smith, S. D., et al. (1997). Colorado Learning Disabilities Research Center. *Learning Disabilities: A Multidisciplinary Journal*, *8*, 7–19.
- Denckla, M. B., & Rudel, R. (1976). Rapid automatized naming (R.A.N.): Dyslexia differentiated from other learning disabilities. *Neuropsychologia*, *14*, 471–479.

- Douglas, V. I. (1999). Cognitive control processes in attention-deficit/hyperactivity disorder. In H. C. Quay & A. E. Hogan (Eds.), *Handbook of disruptive behavior disorders* (pp. 105–138). New York: Kluwer Academic/Plenum.
- Doyle, A. E., Faraone, S. V., DuPre, E. P., & Biederman, J. (2001). Separating attention deficit hyperactivity disorder and learning disabilities in girls: A familial risk analysis. *American Journal of Psychiatry*, *158*, 1666–1672.
- Dunn, L. M., & Markwardt, F. C. (1970). *Examiner's manual: Peabody Individual Achievement Test*. Circle Pines, MN: American Guidance Service.
- DuPaul, G. J., Power, T. P., Anastopoulos, A. D., & Reid, R. (1998). *ADHD Rating Scale-IV*. New York: Guilford.
- Dykman, R. A., & Ackerman, P. T. (1991). ADD and specific reading disability: Separate but often overlapping disorders. *Journal of Learning Disabilities*, *24*, 96–103.
- Falconer, D., & MacKay, T. (1996). *Introduction to quantitative genetics* (4th ed.). London: Longman.
- Faraone, S. V., Biederman, J., Lehman, B., Keenan, K., Norman, D., Seidman, L., et al. (1993). Evidence for the independent familial transmission of attention deficit hyperactivity disorder and learning disabilities: Results from a family genetic study. *American Journal of Psychiatry*, *150*, 891–895.
- Faraone, S. V., Biederman, J., Weber, W., & Russell, R. L. (1998). Psychiatric, neuropsychological, and psychosocial features of *DSM-IV* subtypes of attention-deficit/hyperactivity disorder: Results from a clinically-referred sample. *Journal of the American Academy of Child and Adolescent Psychiatry*, *37*, 185–193.
- Faraone, S. V., Doyle, A. E., Mick, E., & Biederman, J. (2001). Meta-analysis of the association between the 7-repeat allele of the dopamine D4 receptor gene and attention deficit hyperactivity disorder. *American Journal of Psychiatry*, *158*, 1052–1057.
- Felton, R. H., Wood, F. B., Brown, I. S., Campbell, S. K., & Harter, M. H. (1987). Separate verbal memory and naming deficits in attention deficit disorder and reading disability. *Brain and Language*, *31*, 171–184.
- Fergusson, D. M., & Horwood, L. (1992). Attention deficit and reading achievement. *Journal of Child Psychology and Psychiatry*, *33*, 375–385.
- Fisher, S. E., & DeFries, J. C. (2002). Developmental dyslexia: Genetic dissection of a complex cognitive trait. *Nature Reviews: Neuroscience*, *3*, 767–780.
- Friedman, M., Chhabildas, N., Budhiraja, N., Willcutt, E. G., & Pennington, B. F. (2003). Etiology of comorbidity between ADHD and reading disability: Exploration of the assortative mating hypothesis. *American Journal of Medical Genetics (Neuropsychiatric Genetics)*, *120B*, 109–115.
- Gaub, M., & Carlson, C. L. (1997). Behavioral characteristics of *DSM-IV* ADHD subtypes in a school-based population. *Journal of Abnormal Child Psychology*, *25*, 103–111.
- Gayán, J., & Olson, R. K. (2001). Genetic and environmental influences on orthographic and phonological skills in children with reading disabilities. *Developmental Neuropsychology*, *20*, 483–507.
- Gilger, J. W., Pennington, B. F., & DeFries, J. C. (1992). A twin study of the etiology of comorbidity: Attention deficit-hyperactivity disorder and dyslexia. *Journal of the American Academy of Child and Adolescent Psychiatry*, *31*, 343–348.
- Golden, J. C. (1978). *Stroop Color and Word Test*. Chicago: Stoelting.
- Goldstein, G., & Watson, J. R. (1989). Test–retest reliability of the Halstead–Reitan battery and the WAIS in a neuropsychiatric population. *Clinical Neuropsychologist*, *3*, 265–273.
- Gordon, M. (1983). *The Gordon Diagnostic System*. DeWitt, NY: Gordon Systems.
- Heaton, R. K. (1981). *Wisconsin Card Sorting Test manual*. Odessa, FL: Psychological Assessment Resources.
- 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.
- Hollingshead, A. (1975). *Two-factor index of social status*. Unpublished manuscript.
- Houghton, S., Douglas, G., West, J., Whiting, K., Wall, M., Langsford, S., et al. (1999). Differential patterns of executive function in children with attention-deficit hyperactivity disorder according to gender and subtype. *Journal of Child Neurology, 14*, 801–805.
- 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 and Adolescent Psychiatry, 38*, 1148–1155.
- Kuntsi, J., Oosterlaan, J., & Stevenson, J. (2001). Psychological mechanisms in hyperactivity: I. Response inhibition deficit, working memory impairment, delay aversion, or something else? *Journal of Child Psychology and Psychiatry, 42*, 199–210.
- Kuntsi, J., Stevenson, J., Oosterlaan, J., & Sonuga-Barke, E. J. S. (2001). Test–retest reliability of a new delay aversion task and executive function measures. *British Journal of Developmental Psychology, 19*, 339–348.
- Lahey, B. B., Applegate, B., McBurnett, K., Biederman, J., Greenhill, L., Hynd, G. W., et al. (1994). *DSM-IV* field trials for attention deficit hyperactivity disorder in children and adolescents. *American Journal of Psychiatry, 151*, 1673–1685.
- Lahey, B. B., Pelham, W. E., Stein, M., Loney, J., Trapani, C., Nugent, K., et al. (1998). Validity of *DSM-IV* attention-deficit/hyperactivity disorder for young children. *Journal of the American Academy of Child and Adolescent Psychiatry, 37*, 695–702.
- Lahey, B. B., & Willcutt, E. (2002). Validity of the diagnosis and dimensions of attention deficit hyperactivity disorder. In P. S. Jensen & J. R. Cooper (Eds.), *Attention deficit hyperactivity disorder: State of the science* (pp. 1–1–23). New York: Civic Research Institute.
- Levy, F., Hay, D., McStephen, M., Wood, C., & Waldman, I. (1997). Attention-deficit hyperactivity disorder: A category or a continuum? A genetic analysis of a large-scale twin study. *Journal of the American Academy of Child and Adolescent Psychiatry, 36*, 737–744.
- Light, J., Pennington, B., Gilger, J., & DeFries, J. (1995). Reading disability and hyperactivity disorder: Evidence for a common genetic etiology. *Developmental Neuropsychology, 11*, 323–335.
- Lindamood, C., & Lindamood, P. (1971). *Lindamood Auditory Conceptualization Test*. Boston: Reading Resources.
- Logan, G. D. (1994). On the ability to inhibit thought and action: A user's guide to the stop signal paradigm. In D. Dagenbach & T. H. Carr (Eds.), *Inhibitory processes in attention, memory, and language* (pp. 189–239). San Diego: Academic.
- Logan, G. D., Schachar, R. J., & Tannock, R. (1997). Impulsivity and inhibitory control. *Psychological Science, 8*, 60–64.
- Loo, S. K., Fisher, S. E., Francks, C., Ogdie, M. N., MacPhie, I. L., Yang, M., et al. (2004). Genome-wide scan of reading ability in affected sibling pairs with attention-deficit/hyperactivity disorder: Unique and shared genetic effects. *Molecular Psychiatry, 9*, 485–493.
- Mariani, M. A., & Barkley, R. A. (1997). Neuropsychological and academic functioning in preschool boys with attention deficit hyperactivity disorder. *Developmental Neuropsychology, 13*, 111–129.
- Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., & Howerter, A. (2000). The unity and diversity of executive functions and their contributions to complex frontal lobe tasks: A latent variable analysis. *Cognitive Psychology, 41*, 49–100.
- Molina, B. S. G., Smith, B. H., & Pelham, W. E. (2001). Factor structure and criterion validity of secondary school teacher ratings of ADHD and ODD. *Journal of Abnormal Child Psychology, 29*, 71–82.

- Neale, M. C., & Kendler, K. S. (1995). Models of comorbidity for multifactorial disorders. *American Journal of Human Genetics*, *54*, 935–945.
- Nigg, J. T. (1999). The ADHD response inhibition deficit as measured by the Stop Task: Replication with *DSM-IV* combined type, extension, and qualification. *Journal of Abnormal Child Psychology*, *27*, 391–400.
- Nigg, J. T. (2001). Is ADHD a disinhibitory disorder? *Psychological Bulletin*, *127*, 571–598.
- 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. G., Huang-Pollock, C. L., & Rappley, M. D. (2002). Neuropsychological executive functions and *DSM-IV* ADHD subtypes. *Journal of the American Academy of Child and Adolescent Psychiatry*, *41*, 59–66.
- Nigg, J. T., Hinshaw, S. P., Carte, E., & Treuting, 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.
- Olson, R. K. (1985). Disabled reading processes and cognitive profiles. In D. B. Gray & J. F. Kavanagh (Eds.), *Biobehavioral measures of dyslexia* (pp. 215–244). Parkton, MD: York Press.
- Olson, R. K. (1994). Language deficits in “specific” reading disability. In M. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 895–916). New York: Academic.
- Olson, R. K., Forsberg, H., & Wise, B. (1994). Genes, environment, and the development of orthographic skills. In V. W. Berninger (Ed.), *The varieties of orthographic knowledge: I. Theoretical and developmental issues* (pp. 27–72). Dordrecht, The Netherlands: Kluwer Academic.
- Olson, R. K., Wise, B., Conners, F., Rack, J., & Fulker, D. W. (1989). Specific deficits in component reading and language skills: Genetic and environmental influences. *Journal of Learning Disabilities*, *22*, 339–348.
- Oosterlaan, J., & 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 and Psychiatry*, *39*, 411–425.
- Owen, A. M., Doyon, J., Petrides, M., & Evans, A. C. (1996). Planning and spatial working memory: A positron emission tomography study in humans. *European Journal of Neuroscience*, *8*, 353–364.
- Pelham, W. E., Gnagy, E. M., Greenslade, K. E., & Milich, R. (1992). Teacher ratings of *DSM-III-R* symptoms for the disruptive behavior disorders. *Journal of the American Academy of Child and Adolescent Psychiatry*, *31*, 210–218.
- Pennington, B. F. (1997). Dimensions of executive functions in normal and abnormal development. In N. A. Krasnegor & G. R. Lyon (Eds.), *Development of the prefrontal cortex: Evolution, neurobiology, and behavior* (pp. 265–281). Baltimore: Brookes.
- Pennington, B. F. (2002). *The development of psychopathology*. New York: Guilford.
- Pennington, B. F., Bennetto, L., McAleer, O. K., & Roberts, R. J., Jr. (1995). Executive functions and working memory: Theoretical measurement issues. In G. R. Lyon & N. A. Krasnegor (Eds.), *Attention, memory and executive function* (pp. 327–348). Baltimore: Brookes.
- Pennington, B. F., Gilger, J. W., Olson, R. K., & DeFries, J. C. (1992). The external validity of age- versus IQ-discrepancy definitions of reading disability: Lessons from a twin study. *Journal of Learning Disabilities*, *25*, 562–573.
- Pennington, B. F., Groisser, D., & Welsh, M. C. (1993). Contrasting cognitive deficits in attention deficit hyperactivity disorder versus reading disability. *Developmental Psychology*, *29*, 511–523.
- Pennington, B. F., & Ozonoff, S. (1996). Executive functions and developmental psychopathology. *Journal of Child Psychology and Psychiatry*, *37*, 51–87.
- Petrides, M., & Milner, B. (1982). Deficits on subject-ordered tasks after frontal- and temporal-lobe lesions in man. *Neuropsychologia*, *20*, 249–262.

- Pisecco, S., Baker, D. B., Silva, P. A., & Brooke, M. (2001). Boys with reading disabilities and/or ADHD: Distinctions in early childhood. *Journal of Learning Disabilities, 34*, 98–106.
- Plomin, R., DeFries, J. C., McClearn, G. E., & Rutter, P. (1997). *Behavioral genetics* (3rd ed.). New York: Freeman.
- Purvis, K. L., & Tannock, R. (2000). Phonological processing, not inhibitory control, differentiates ADHD and reading disability. *Journal of the American Academy of Child and Adolescent Psychiatry, 39*, 485–494.
- Rack, J. P., Snowling, M. J., & Olson, R. K. (1992). The nonword reading deficit in developmental dyslexia: A review. *Reading Research Quarterly, 27*, 28–53.
- Reich, W., & Welner, Z. (1988). *Revised version of the Diagnostic Interview for Children and Adolescents (DICA-R)*. St. Louis, MO: Department of Psychiatry, Washington University School of Medicine.
- Reich, W., Welner, Z., & Herjanic, B. (1997). *Diagnostic Interview for Children and Adolescents-IV*. North Towanda Falls, NY: Multi-Health System.
- Reitan, R., & Wolfson, D. (1985). *The Halstead-Reitan Neuropsychological Test Battery: Theory and clinical interpretation*. Tucson, AZ: Neuropsychology Press.
- Roodenrys, S., Koloski, N., & Grainger, J. (2001). Working memory function in attention deficit hyperactivity disorder and reading disabled children. *British Journal of Developmental Psychology, 19*, 325–337.
- Rosner, J. & Simon, D. (1971). The Auditory Analysis Test: An initial report. *Journal of Learning Disabilities, 4*, 384–391.
- Rucklidge, J. J., & Tannock, R. (2002). Neuropsychological profiles of adolescents with ADHD: Effects of reading difficulties and gender. *Journal of Child Psychology and Psychiatry, 43*, 988–1003.
- Schachar, R., Mota, V. L., Logan, G. D., Tannock, R., Klim, P. (2000). Confirmation of an inhibitory control deficit in attention-deficit/hyperactivity disorder. *Journal of Abnormal Child Psychology, 28*, 227–235.
- Schmitz, M., Cadore, L., Paczko, M., Kipper, L., Chaves, M., Rohde, L. A., et al. (2002). Neuropsychological performance in DSM-IV ADHD subtypes: An exploratory study with untreated adolescents. *Canadian Journal of Psychiatry, 47*, 863–869.
- Seidman, L. J., Biederman, J., Monuteaux, M. C., Doyle, A. E., & Faraone, S. V. (2001). Learning disabilities and executive dysfunction in boys with attention-deficit/hyperactivity disorder. *Neuropsychology, 15*, 544–556.
- Semrud-Clikeman, M., Biederman, J., Sprich-Buckminster, S., Lehman, B. K., Faraone, S. V., & Norman, D. (1992). Comorbidity between ADDH and LD: A review and report in a clinically referred sample. *Journal of the American Academy of Child and Adolescent Psychiatry, 31*, 439–448.
- Semrud-Clikeman, M., Guy, K., Griffin, J. D., & Hynd, G. W. (2000). Rapid naming deficits in children and adolescents with reading disabilities and attention deficit hyperactivity disorder. *Brain and Language, 74*, 70–83.
- Sergeant, J. (2000). The cognitive-energetic model: An empirical approach to attention-deficit hyperactivity disorder. *Neuroscience and Biobehavioral Reviews, 24*, 7–12.
- Sergeant, J., Oosterlaan, J., & van der Meere, J. (1999). Information processing and energetic factors in attention-deficit/hyperactivity disorder. In H. C. Quay & A. E. Hogan (Eds.), *Handbook of disruptive behavior disorders* (pp. 75–104). New York: Kluwer Academic/Plenum.
- Shaywitz, B. A., Fletcher, J. M., & Shaywitz, S. E. (1995). Defining and classifying learning disabilities and attention-deficit/hyperactivity disorder. *Journal of Child Neurology, 10*, S50–S57.
- Siegel, L. S. (1989). IQ is irrelevant to the definition of learning disabilities. *Journal of Learning Disabilities, 22*, 469–478.
- Siegel, L. S., & Ryan, E. B. (1989). The development of working memory in normally achieving and subtypes of learning disabled children. *Child Development, 60*, 973–980.

- Solanto, M. V., Abikoff, H., Sonuga-Barke, E. J. S., Schachar, R., Logan, G. D., Wigal, T., et al. (2001). The ecological validity of delay aversion and response inhibition as measures of impulsivity in AD/HD: A supplement to the NIMH multimodal treatment study of AD/HD. *Journal of Abnormal Child Psychology*, *29*, 215–228.
- Sonuga-Barke, E. J. S., Taylor, E., Sembi, S., & Smith, J. (1992). Hyperactivity and delay aversion—I. The effect of delay on choice. *Journal of Child Psychology and Psychiatry*, *33*, 387–398.
- Stevenson, J., Pennington, B. F., Gilger, J. W., DeFries, J. C., & Gillis, J. J. (1993). Hyperactivity and spelling disability: Testing for shared genetic aetiology. *Journal of Child Psychology and Psychiatry*, *14*, 1137–1152.
- Swanson, H. L., Mink, J., & Bocian, K. M. (1999). Cognitive processing deficits in poor readers with symptoms of reading disabilities and ADHD: More alike than different? *Journal of Educational Psychology*, *91*, 321–333.
- Tannock, R., Martinussen, R., & Frijters, J. (2000). Naming speed performance and stimulant effects indicate effortful, semantic processing deficits in attention-deficit/hyperactivity disorder. *Journal of Abnormal Child Psychology*, *28*, 237–252.
- Todd, R. D., Sitdhiraska, N., Reich, W., Ji, T. H. C., Joyner, C. A., Heath, A. C., et al. (2002). Discrimination of *DSM-IV* and latent class attention-deficit/hyperactivity disorder subtypes by educational and cognitive performance in a population-based sample of child and adolescent twins. *Journal of the American Academy of Child and Adolescent Psychiatry*, *41*, 820–828.
- Wagner, R. K., & Torgesen, J. K. (1987). The nature of phonological processing and its causal role in the acquisition of reading skills. *Psychological Bulletin*, *101*, 192–212.
- Wagner, R. K., Torgesen, J. K., & Rashotte, C. A. (1994). Development of reading-related processing abilities: New evidence of bidirectional causality from a latent variable longitudinal study. *Developmental Psychology*, *30*, 73–87.
- Wechsler, D. (1974). *Examiner's manual: Wechsler Intelligence Scale for Children-Revised*. New York: Psychological Corporation.
- Wechsler, D. (1981). *Manual for the Wechsler Adult Intelligence Scale-Revised*. San Antonio, TX: Psychological Corporation.
- Wechsler, D. (1991). *Manual for the Wechsler Intelligence Scale for Children, 3rd Edition*. San Antonio, TX: Psychological Corporation.
- Werry, J. S., Elkind, G. S., & Reeves, J. S. (1987). Attention deficit, conduct, oppositional, and anxiety disorders in children: III. Laboratory differences. *Journal of Abnormal Child Psychology*, *15*, 409–428.
- Weyandt, L. L., Rice, J. A., Linterman, I., Mitzlaff, L., & Emert, E. (1998). Neuropsychological performance of a sample of adults with ADHD, developmental reading disorder, and controls. *Developmental Neuropsychology*, *14*, 643–656.
- Weyandt, L. L., & Willis, W. G. (1994). Executive functions in school-aged children: Potential efficacy of tasks in discriminating clinical groups. *Developmental Neuropsychology*, *10*, 1697–1713.
- Willcutt, E. G. (in press). Genetics of ADHD. In D. Barch (Ed.), *Cognitive and affective neuroscience of psychopathology*. Oxford, England: Oxford University Press.
- Willcutt, E. G., Brodsky, K., Chhabildas, N., Shanahan, M., Yerys, B., Scott, A., et al. (in press). The neuropsychology of ADHD: Validity of the executive function hypothesis. In D. Gozal & D. Molfese (Eds.), *Attention deficit hyperactivity disorder: From genes to practice*. Totowa, NJ: Humana Press.
- Willcutt, E. G., Chhabildas, N., & Pennington, B. F. (2001). Validity of the *DSM-IV* subtypes of ADHD. *ADHD Report*, *9*(1), 2–5.
- Willcutt, E. G., Chhabildas, N. A., & Pennington, B. F. (1998, October). *Psychiatric correlates of reading disability*. Poster session presented at the Kansas Conference on Clinical Child Psychology, Lawrence, KS.

- Willcutt, E. G., DeFries, J. C., Pennington, B. F., Olson, R. K., Smith, S. D., & Cardon, L. R. (2003). Genetic etiology of comorbid reading difficulties and ADHD. In R. Plomin, J. C. DeFries, P. McGuffin, & I. Craig (Eds.), *Behavioral genetics in a postgenomic era* (pp. 227–246). Washington, DC: American Psychological Association.
- Willcutt, E. G., Lahey, B. B., Pennington, B. F., Carlson, C. L., Nigg, J. T. & McBurnett, K. (2003). *Validity of attention-deficit/hyperactivity disorder*. Manuscript submitted for publication.
- Willcutt, E. G., & Pennington, B. F. (2000). Comorbidity of reading disability and attention-deficit/hyperactivity disorder: Differences by gender and subtype. *Journal of Learning Disabilities, 33*, 179–191.
- Willcutt, E. G., Pennington, B. F., Boada, R., Tunick, R. A., Oglie, J., Chhabildas, N. A., et al. (2001). A comparison of the cognitive deficits in reading disability and attention-deficit/hyperactivity disorder. *Journal of Abnormal Psychology, 110*, 157–172.
- Willcutt, E. G., Pennington, B. F., & DeFries, J. C. (2000a). Etiology of inattention and hyperactivity/impulsivity in a community sample of twins with learning difficulties. *Journal of Abnormal Child Psychology, 28*, 149–159.
- Willcutt, E. G., Pennington, B. F., & DeFries, J. C. (2000b). A twin study of the etiology of comorbidity between reading disability and attention-deficit/hyperactivity disorder. *American Journal of Medical Genetics (Neuropsychiatric Genetics), 96*, 293–301.
- Willcutt, E. G., Pennington, B. F., Smith, S. D., Cardon, L. R., Gayán, J., Knopik, V. S., et al. (2002). Quantitative trait locus for reading disability on chromosome 6p is pleiotropic for ADHD. *American Journal of Medical Genetics (Neuropsychiatric Genetics), 260–268*.
- Wolraich, M. L., Feurer, I. D., Hannah, J. N., Baumgaertel, A., & Pinnock, T. Y. (1998). Obtaining systematic teacher reports of disruptive behavior disorders utilizing *DSM-IV*. *Journal of Abnormal Child Psychology, 26*, 141–152.