

Colorado longitudinal twin study of reading disability

Sally J. Wadsworth · John C. DeFries ·
Richard K. Olson · Erik G. Willcutt

Received: 10 January 2007 / Accepted: 10 September 2007 / Published online: 4 December 2007
© The International Dyslexia Association 2007

Abstract The primary objectives of the present study are to introduce the Colorado Longitudinal Twin Study of Reading Disability, the first longitudinal twin study in which subjects have been specifically selected for having a history of reading difficulties, and to present some initial assessments of the stability of reading performance and cognitive abilities in this sample. Preliminary examination of the test scores of 124 twins with a history of reading difficulties and 154 twins with no history of reading difficulties indicates that over the 5- to 6-year interval between assessments, cognitive and reading performance are highly stable. As a group, those subjects with a history of reading difficulties had substantial deficits relative to control subjects on all measures at initial assessment, and significant deficits remained at follow-up. The stability noted for all cognitive and achievement measures was highest for a composite measure of reading, whose average stability correlation across groups was 0.80. Results of preliminary behavior genetic analyses for this measure indicated that shared genetic influences accounted for 86% and 49% of the phenotypic correlations between the two assessments for twin pairs with and without reading difficulties, respectively. In addition, genetic correlations reached unity for both groups, suggesting that the same genetic influences are manifested at both time points.

Keywords Longitudinal · Twin · Reading disability · Cognitive ability · Etiology · Genetic · Stability

S. J. Wadsworth (✉) · J. C. DeFries · R. K. Olson · E. G. Willcutt
Institute for Behavioral Genetics, University of Colorado, 447 UCB, Boulder, CO 80309, USA
e-mail: Sally.Wadsworth@Colorado.edu

J. C. DeFries
e-mail: John.DeFries@Colorado.edu

R. K. Olson
e-mail: Richard.Olson@Colorado.edu

E. G. Willcutt
e-mail: Erik.Willcutt@Colorado.edu

J. C. DeFries · R. K. Olson · E. G. Willcutt
Department of Psychology, University of Colorado, Boulder, CO 80309, USA

Introduction

During the past few years, considerable progress has been achieved toward understanding the nature of reading difficulties (RD). For example, results obtained from recent studies have provided compelling evidence regarding their genetic and environmental etiologies (e.g., Fisher & DeFries, 2002; Schumacher et al., 2007; Williams & O'Donovan, 2006), subtypes or dimensions (e.g., Bailey, Manis, Pedersen & Seidenberg, 2004; Griffiths & Snowling, 2002; Olson, Datta, Gayán & DeFries, 1999; Vukovic & Siegel, 2006), gender differences (e.g., Harlaar, Spinath, Dale & Plomin, 2005; Rutter et al., 2004; Wadsworth & DeFries, 2005), and comorbidity with other conditions such as attention-deficit/hyperactivity disorder (ADHD) (e.g., Gayán et al., 2005; McGee, Prior, Williams, Smart & Sanson, 2002; Stevenson et al., 2005; Willcutt et al., 2002, 2003) and antisocial behavior (e.g., Trzesniewski, Moffitt, Caspi, Taylor & Maughan, 2006; Willcutt & Pennington, 2000). In addition, recent longitudinal studies have addressed the stability of RD and rates of change for disabled vs nondisabled readers (e.g., Badian, 1999; Bast & Reitsma, 1998; Raskind et al., 1998; Raskind, Higgins, Goldberg & Herman, 1999; Shaywitz et al., 2003). However, relatively little is known about the etiology of developmental stability and change in reading disability. Genetically informative designs, such as twin, adoption, or molecular genetic family studies, are needed to address these issues, and such studies of reading ability and disability are relatively recent (e.g., Astrom, Wadsworth & DeFries, 2007; Byrne et al., 2007; Harlaar, Dale & Plomin, 2007; Lyytinen et al., 2001; Moffitt, 2002; Petrill et al., 2007; Wadsworth, Corley, Plomin, Hewitt & DeFries, 2006).

Although there are a number of longitudinal studies with genetically informative designs and data related to reading performance, the Colorado Longitudinal Twin Study of Reading Disability (LTSRD) is the first longitudinal study of children specifically selected for reading disability utilizing genetically informative data. In this ongoing study, follow-up assessment of a subset of twin pairs and their siblings who previously participated in the Colorado Learning Disabilities Research Center (CLDRC; DeFries et al., 1997) is conducted approximately 5–6 years after their initial participation. Participants are administered an extensive psychometric test battery, including tests of reading performance, spelling, reading component processes, general and specific cognitive abilities, and behavioral interviews/questionnaires.

The availability of these data from twins at the two time points provides the opportunity to explore many questions related to reading disability and its relations with other variables/traits. Although the goals of the LTSRD are broad, including the assessment of stability and change, longitudinal comorbidity with ADHD and other psychopathology, predictors of growth in reading, and long-term outcome, the current report focuses on stability of performance on a subset of our measures of cognitive ability and reading achievement among those with and without a history of reading difficulties, with particular focus on the stability of reading performance and its etiology.

Stability of reading ability and disability

Results obtained from longitudinal studies indicate that reading deficits are generally stable (e.g., Maughan, Hagell, Rutter & Yule, 1994; Satz, Buka, Lipsitt & Seidman, 1998; Shaywitz et al., 1999; Shaywitz et al., 2003; Spira, Bracken & Fischel, 2005), with similar magnitude of deficits remaining over several years, and with few exceptions, no indication that poor readers catch up to their peers. In addition, some studies have suggested that children with reading difficulties continue to fall farther behind their normally-achieving

peers in reading performance (e.g., Maughan et al., 1994; Bast & Reitsma, 1998; Speece & Ritchey, 2005), a phenomenon referred to as the “Matthew effect” (Stanovich, 1986). Although this effect has often been noted, few studies have tested the hypothesis directly (Bast & Reitsma, 1998). Furthermore, not all studies have observed the effect (e.g., Scarborough & Parker, 2003; Shaywitz et al., 1999; Shaywitz, Holford et al., 1995), nor has it been found for all aspects of reading (Bast & Reitsma, 1998).

Although a number of studies have examined the longitudinal stability of reading disability, relatively little is known about the etiology of this stability or how reading difficulties change over time as a function of genetic and environmental influences (McKinney, 1994). Thus, in a preliminary analysis of data from the LTSRD, Astrom et al. (2007) assessed the etiology of the phenotypic stability of a composite measure of reading performance from 56 twin pairs (18 identical, 38 fraternal) tested at two time points (average age of 10.3 years at initial assessment and 16.1 years at follow-up) and meeting criteria for reading disability at initial assessment. Similar to the findings of studies with non-twin samples of children with RD, scores on the reading composite were highly stable, with a stability correlation of 0.84, and averaging more than two standard deviations (SDs) below those of an unaffected control group ($n=1263$ at initial assessment and $n=93$ at follow-up) at both assessments. In addition, results of bivariate DeFries-Fulker multiple regression analysis (DeFries & Fulker, 1985, 1988) indicated that common genetic influences accounted for approximately 75% of the stability between reading deficits at the two time points.

Although little is known about the etiology of stability of reading *deficits*, a few studies have assessed the etiology of the stability of reading performance within the normal range. For example, to assess the stability of genetic and environmental influences on reading performance measured at 7, 12, and 16 years of age, Wadsworth, Corley, Plomin, Hewitt and DeFries (2006) analyzed data from adoptive and nonadoptive probands and their unrelated and related siblings participating in the Colorado Adoption Project (CAP). Observed stability correlations ranged from 0.58–0.71, with 53% of the observed stability between ages 7 and 12, 86% of that between ages 7 and 16, and 62% of that between ages 12 and 16 attributable to shared genetic influences. Furthermore, genetic correlations among the measures at the three ages were near unity, suggesting that those genetic factors influencing reading performance at age 7 are also operating at ages 12 and 16.

Recently, Harlaar, Dale and Plomin (2007) assessed the stability of genetic influences on reading achievement of participants of the Twins Early Development Study (TEDS), a longitudinal study of a population sample of twins ascertained from records of twin births in England and Wales. The reading achievement of 4,291 twin pairs was evaluated by teacher assessment at ages, 7, 9, and 10. In addition, at age 10, participants completed a web-based test at home. Results from this study confirmed stability of individual differences in reading performance ($r=.59-.63$) and suggested that 68%–77% of the phenotypic stability correlation is genetically mediated. However, unlike the CAP results reported by Wadsworth et al. (2006), genetic correlations were less than 1.0 and “new” genetic influences were observed at ages 9 and 10.

Although analyses of the etiology of stability of reading performance have, in large part, been limited to reading performance within the normal range, the preliminary findings of Astrom et al. (2007) for reading deficits are remarkably similar to those of Wadsworth et al. (2006) and Harlaar et al. (2007) for individual differences. In addition, there is some evidence to suggest that reading disability may represent the lower tail of a normal distribution (e.g., Shaywitz, Escobar, Shaywitz, Fletcher & Makuch, 1992; Plomin & Kovas, 2005), suggesting that the etiologies of stability of reading deficits and reading

performance within the normal range should be similar. However, additional studies of the etiology of stability of both reading deficits and reading performance within the normal range are clearly warranted. The current study provides a preliminary assessment of the stability of reading performance and cognitive ability, and the etiology of stability of reading performance, among a subset of participants in the CLDRC, tested approximately 5–6 years after their initial participation.

Methods

Sample

The subjects of the current study participated in the ongoing CLDRC approximately 5–6 years before follow-up assessment. Thus, a brief description of the CLDRC and its ascertainment procedures follow. To minimize the possibility of referral bias (Vogel, 1990), twin pairs in this ongoing study (e.g., DeFries et al., 1997), have been ascertained systematically through 27 cooperating school districts in the state of Colorado. Without regard to reading or ADHD status, all twin pairs within each district are identified, and permission is then sought from parents to review the school records of both members of each pair for evidence of reading problems (e.g., low reading achievement test scores, referral to a reading therapist, reports by classroom teachers, school psychologists, parents, etc.) or ADHD symptoms (meeting DSM-IV symptom criteria for ADHD based on the combination of parent and teacher ratings). If either member of a twin pair has a positive history of reading problems or ADHD symptoms, both members of the pair are invited to complete 2 days of testing in our laboratories at the Institute for Behavioral Genetics and the Department of Psychology, University of Colorado, Boulder, and the Department of Psychology, University of Denver. Participants are administered an extensive psychometric test battery that includes cognitive and achievement tests and measures of reading and language processes, ADHD symptoms, other psychopathology, and executive functions. A discriminant function score (DISCR) is computed for each subject employing discriminant weights estimated from an analysis of data from the Reading Recognition, Reading Comprehension, and Spelling subtests of the Peabody Individual Achievement Test (PIAT; Dunn & Markwardt, 1970) obtained from an independent sample of 140 reading-disabled and 140 control non-twin children (DeFries, 1985; DeFries et al., 1997; Wadsworth & DeFries, 2005; Meng et al., 2005; Willcutt et al., 2002). For an individual to be classified as reading disabled, he or she must have a positive history for reading problems and also be classified as affected by the discriminant score. Additional diagnostic criteria include an IQ score of at least 80 on either the Verbal or Performance Scale of the Wechsler Intelligence Scale for Children-Revised (WISC-R; Wechsler, 1974) or Wechsler Adult Intelligence Scale-Revised (WAIS-R; Wechsler, 1981); no evidence of neurological problems and no uncorrected visual or auditory acuity deficits. A comparison group of control twins is also tested. Control twin pairs are matched to probands on the basis of age, gender, and school district. For a twin pair to be included in the control sample, both members of the pair must have a negative history for reading problems. Additional selection criteria may be imposed, depending on the analyses conducted. Selected items from the Nichols and Bilbro (1966) questionnaire are used to determine zygosity of same-sex twin pairs. In ambiguous cases, zygosity of the pair is confirmed by analysis of blood or buccal samples. Participating twin pairs range in age from 7.7 to 20.5 years.

The Colorado Longitudinal Twin Study of Reading Disability The subjects of the current study participated in the ongoing CLDRC between September, 1996 and September, 2001. Subjects were contacted by mail and invited to participate in a day of testing in our laboratory at the Institute for Behavioral Genetics, University of Colorado, Boulder. Those returning contact information update forms were contacted by phone to provide families with more detailed information about the project and schedule participants for testing. Approximately 55% of those recontacted have participated in follow-up testing (364 of 665 contacted through May, 2006). Although some twin pairs were ascertained for a history of ADHD symptoms, for the current analyses, twin pairs in which one or both members had a history of ADHD symptoms but no reading difficulties were excluded from the analyses. Thus, the sample for the current analyses includes 106 MZ twins and 172 DZ twins ranging in age from 8.1 to 15.9 years (average age of 10.4 years) at initial assessment, and from 12.9 to 23.9 years (average age of 16.0 years) at follow-up.

Measures

In the LTSRD, we employ some of the most diagnostically revealing measures from the CLDRC (DeFries et al., 1997), including measures of ADHD symptoms, other psychopathology, and outcome. Careful attention has been given to providing comparable measures across ages from middle childhood to early adulthood. For those tests for which revisions have been published, in the follow-up study, we use revisions available at the onset of the study. Moreover, measures of component reading processes that have been developed more recently and used in the CLDRC are also administered. In the interest of space, the tests, along with reliabilities, where available, and indication of whether each test was administered in the CLDRC are listed in Tables 1 (cognitive and academic achievement measures) and 2 (behavioral problems and measures of outcome), grouped by area of assessment. These areas represent an effort to balance the assessment of receptive vs expressive performance and performance in a wide array of cognitive, academic achievement, language and behavioral areas. References to detailed descriptions of the measures are also provided.

For the present study, twin data from a subset of our measures of cognitive ability and reading achievement at the two time points were analyzed. Measures of general cognitive ability included the WISC-R/WISC-III (Wechsler, 1974, 1991) or WAIS-R/WAIS-III (Wechsler, 1981, 1997). In addition, processing speed was measured using the Colorado Perceptual Speed Test (CPS; DeFries et al., 1978) and the Identical Pictures Test (IPT; French et al., 1963), each of which require the rapid processing of visually presented symbolic material, and the Colors, Pictures, Numbers and Letters subtests of the Rapid Automatized Naming test (RAN; after Denckla & Rudel, 1976). Reading achievement was measured using the Reading Recognition (REC), Reading Comprehension (COMP) and Spelling (SPELL), subtests of the Peabody Individual Achievement Test (PIAT) and PIAT-Revised (PIAT-R; Dunn & Markwardt, 1970; Markwardt, 1989), and the discriminant function score (DISCR) derived from these three subtests.

Analyses

Phenotypic analyses Because all phenotypic analyses included data from members of twin pairs, all group comparisons were estimated in Mplus (Muthén & Muthén, 1998–2007)

Table 1 Reading-related measures administered in the LTSRD

Construct/measure	Administered at Initial testing?	Reliability	Reference
General cognitive ability			
WISC-III/WAIS-III	Yes ^a	0.96/0.97 ^b	Wechsler (1991, 1997)
Word reading and written spelling			
PIAT-R Reading Recognition	Yes ^c	0.96 ^b	Markwardt (1989)
Timed Oral Reading of Single Words	Yes	0.93 ^b	Olson, Forsberg, Wise & Rack (1994)
TOWRE Sight Word Efficiency	No	0.84 ^b	Torgesen, Wagner & Rashotte (1999)
WRAT-3 Spelling Production	Yes	0.93	Wilkinson (1993)
Phonological decoding			
Oral Nonword Reading	Yes	0.86 ^b	Olson, Forsberg, Wise & Rack (1994)
TOWRE Phonemic Decoding	No	0.89 ^b	Torgesen, Wagner, & Rashotte (1999)
Phonological Choice	Yes	0.80 ^d	Olson, Forsberg, Wise & Rack (1994)
Phonological segmentation/manipulation			
Phoneme Segmentation/Transposition	Yes	0.80 ^e	Olson, Forsberg & Wise (1994)
Phoneme Deletion	Yes	0.80 ^e	Olson, Forsberg & Wise (1994)
Orthographic coding			
Lexical Decision by Orthography	Yes	0.93 ^f	Olson, Forsberg, Wise & Rack (1994)
PIAT-R Spelling	Yes ^c	0.88 ^b	Markwardt (1989)
Reading comprehension			
PIAT-R Reading Comprehension	Yes ^c	0.88 ^b	Markwardt (1989)
WIAT-II Reading Comprehension	No	0.95 ^f	Wechsler (2002)
Processing speed			
Colorado Perceptual Speed (Revised)	Yes	0.81 ^b	Defries, Plomin, Vandenberg & Kuise (1981)
Identical Pictures	Yes	0.82 ^b	French, Ekstrom & Price (1963)
Rapid Automated Naming	Yes	0.84 ^b	Denckla and Rudel (1976)

^a WISC-R/WAIS-R^b Test-retest reliability^c PIAT^d Correlation with oral nonword reading^e Correlation between Phoneme Segmentation and Phoneme Deletion^f Split-half reliability

Table 2 Measures of ADHD, other psychopathology and outcome

Construct/measure	Rater ^a	Administered at initial testing?	Reliability ^b	Reference
DSM-IV ADHD				
DICA-IV	P	Yes	0.88	Reich, Welner & Herjanic (1997)
DBRS rating scale	P, T	Yes	0.79–0.94	Barkley & Murphy (1998)
DBRS impairment scales	P, T	Yes	0.75–0.86	“
DICA-IV impairment scales	P, S	Yes	0.61–0.84	Reich, Welner & Herjanic (1997)
Externalizing behaviors				
DICA-IV ODD and CD	P, S	Yes	0.78	Reich, Welner & Herjanic (1997)
ASEBA rating scales	P, T, S	Yes	0.89–0.92	Achenbach & Rescorla (2001)
Internalizing behaviors				
DICA-IV GAD, MDD, phobia	P, S	Yes	0.74–0.84	Reich, Welner & Herjanic (1997)
ASEBA rating scales	P, T, S	Yes	0.80–0.91	Achenbach & Rescorla, (2001)
CDI	S	Yes	0.84	Kovacs (1988)
RCMAS	S	Yes	0.84–0.92	Reynolds & Richmond (1985)
Outcome information (derived from DICA, DBRS, ASEBA, Study Questionnaire)				
Interventions received				
(Special ed., tutoring, therapy, med.)	P, S	Yes	0.77 ^c –0.91 ^{b,c}	Reich, Welner & Herjanic (1997)
Academic outcome				
General acad. impairment,	P, S, T	Yes	0.77–0.98 ^{b,c}	Reich et al. (1997), Barkley & Murphy (1998)
Grades, subject difficulty,	P, S	Yes	0.89–0.93 ^{b,c}	Achenbach & Rescorla, (2001)
Retained, reading time	P, S	Yes	0.66 ^c	Study questionnaire
Post-HS education/plans	S	No	–	Study questionnaire
ACT/SAT score	P, S	No	–	Study questionnaire
Behavioral outcome				
School/suspensions	P, T, S	Yes	0.83–0.91	Achenbach & Rescorla (2001); Reich et al. (1997)
Delinquency	P, T, S	Yes	–	Achenbach & Rescorla (2001); Reich et al. (1997)
Social outcome				
Global social problems,	P, S, T	Yes	0.74–0.95	Reich et al. (1997) Barkley & Murphy (1998)
sibs, parents, peers	P, S, T	Yes	–	Achenbach & Rescorla (2001)
Occupational outcome				
Type of job/performance	P, S	Yes	0.87–0.93	Achenbach & Rescorla (2001); Study Quest.

^a *DSM-IV* Diagnostic and Statistical Manual of Mental Disorders, 4th Edition (American Psychiatric Association, 1994), *CDI* Children's Depression Inventory, *RCMAS* Revised Children's Manifest Anxiety Scale, *DICA-IV*/DSM-IV Diagnostic Interview for Children and Adolescents, *DBRS* Disruptive Behavior Rating Scale, *ASEBA* Achenbach System of Empirically Based Assessment (CBCL, TRF, YSR forms), *P* Parent, *T* Teacher, *S* self-report

^b Test-retest reliability unless otherwise noted

^c Inter-rater reliability

using the “CLUSTER” and “TYPE = complex” options to obtain standard errors, t values, and p values robust to nonindependence. In addition, for the individual t tests, a Bonferroni correction for multiple planned comparisons was used. Thus, for the 11 means compared at each time point, an alpha level of 0.0045 was adopted.

Phenotypic correlations among the variables, within time point, were calculated separately for twins with a history of reading difficulties, for their co-twins with no history of reading difficulties, and for control twins. In addition, correlations are provided for the full CLDRC control sample ($n=1285$) as a large sample comparison group with which to compare the initial correlations of the different groups on whom follow-up data are available (Table 3).

Sample representativeness To assess the extent to which participants in the follow-up study are representative of the full CLDRC sample, initial IQ and reading composite scores of those retested were compared to scores of those participants who were tested in the CLDRC during the same time period from which the current follow-up sample was recruited (i.e., September, 1996–September, 2001), but have not been retested in the LTSRD, irrespective of the reason.

Phenotypic stability Means and SDs were computed separately for MZ and DZ twins with a positive history of reading difficulties (+HistRD), their co-twins with a negative history of reading difficulties (−HistRD), and control twins. Analysis of test scores was performed using published age standard or scaled scores where available (PIAT, PIAT-R, WISC/WAIS-R, WISC/WAIS-III), as these were less likely to show ceiling effects in the follow-up sample. The remaining measures were age-adjusted based on the full contemporaneous control samples at each time point (i.e., scores from the initial assessment were age-adjusted based on the scores of 1,284 control subjects who participated in the original CLDRC through March 2006, and scores from the follow-up session were age-adjusted based on the scores of the 94 control subjects participating, thus far, in the follow-up study). All age-adjusted and standard scores were standardized based on the means of the full contemporaneous control samples at each assessment. Due to the small sample sizes, the data were not analyzed for gender differences. To facilitate interpretation of these data, unstandardized, raw scores (i.e., not age-adjusted, with the exception of the WISC and WAIS IQ scores) are provided for each assessment, in Tables 4 and 5. It should be noted that the WISC-III and WAIS-III have been more recently normed than the WISC-R and WAIS-R, resulting in slightly lower means at follow-up. In addition, the PIAT and PIAT-R raw subtest scores are not strictly parallel due to the differing numbers of items between the two test editions.

Because the sample sizes of some of the individual groups are still quite small, to obtain reliable correlations, Pearson stability correlations were computed for twins with and without a history of reading difficulties, and for control twins, combining scores of MZ and DZ twins. Before combining the MZ and DZ scores, t tests of the difference between MZ and DZ means and variances were first conducted to determine the appropriateness of combining MZ and DZ scores. As noted above, group comparisons were estimated in Mplus (Muthén & Muthén, 1998–2007) to obtain standard errors, t values, and p values robust to nonindependence. Because MZ and DZ means were not significantly different, scores of MZ and DZ twins were combined within history group for calculation of the stability correlations.

Table 3 Within-time-point intercorrelations among variables at initial (below diagonal) and follow-up (above diagonal) assessments for Control, Positive history, and Negative history groups

	FSIQ	REC	COMP	SPELL	DISCR	RAN COL	RAN PIC	RAN NUM	RAN LET	CPS	IPT
Control subjects with follow-up data (<i>n</i> =94)											
FSIQ	1.00	0.49	0.53	0.42	0.58	0.44	0.27**	0.30**	0.26*	0.39	0.17 ⁺
REC	0.42	1.00	0.49	0.54	0.93	0.18 ⁺	0.02 ⁺	0.18 ⁺	0.15 ⁺	0.33	0.01 ⁺
COMP	0.57	0.68	1.00	0.44	0.70	0.22 ⁺	0.13 ⁺	0.26*	0.27**	0.34	0.21 ⁺
SPELL	0.27	0.68	0.46	1.00	0.76	0.13 ⁺	0.03 ⁺	0.18 ⁺	0.15 ⁺	0.35	0.03 ⁺
DISCR	0.46	0.95	0.78	0.83	1.00	0.21*	0.06 ⁺	0.24*	0.21*	0.40	0.07 ⁺
RAN COL	0.29	0.17 ⁺	0.24	0.10 ⁺	0.19 ⁺	1.00	0.64	0.60	0.67	0.29***	0.20 ⁺
RAN PIC	0.26	0.16 ⁺	0.18 ⁺	0.25*	0.22*	0.56	1.00	0.65	0.63	0.35	0.24*
RAN NUM	0.13 ⁺	0.30	0.18 ⁺	0.22*	0.29***	0.50	0.55	1.00	0.76	0.47	0.25*
RAN LET	0.04 ⁺	0.32	0.16 ⁺	0.19 ⁺	0.28 ⁺	0.36	0.54	0.69	1.00	0.40	0.22*
CPS	0.36	0.42	0.30	0.54	0.49	0.27	0.38	0.42	0.38	1.00	0.37
IPT	0.26	-0.08 ⁺	-0.02 ⁺	0.08 ⁺	-0.02 ⁺	0.35	0.46	0.40	0.32**	0.38	1.00
All control subjects participating in initial testing through May 2006 ^a (<i>n</i> =1284)											
FSIQ	1.00										
REC	0.41	1.00									
COMP	0.51	0.57	1.00								
SPELL	0.30	0.59	0.40	1.00							
DISCR	0.47	0.91	0.72	0.81	1.00						
RAN COL	0.27	0.15	0.13	0.17	0.18	1.00					
RAN PIC	0.24	0.11	0.11	0.13	0.14	0.46	1.00				
RAN NUM	0.14	0.16	0.07*	0.17	0.16	0.53	0.41	1.00			
RAN LET	0.21	0.26	0.14	0.25	0.27	0.50	0.43	0.72	1.00		
CPS	0.37	0.32	0.20	0.42	0.39	0.31	0.22	0.29	0.38	1.00	
IPT	0.42	0.10	0.14	0.13	0.14	0.32	0.30	0.27	0.26	0.50	1.00
Subjects with a positive history of reading difficulties, who have data at follow-up (<i>n</i> =124)											
FSIQ	1.00	0.51	0.45	0.35	0.54	0.13 ⁺	0.15 ⁺	0.11 ⁺	0.11 ⁺	0.33	0.48
REC	0.46	1.00	0.49	0.66	0.93	0.12 ⁺	0.09 ⁺	0.22*	0.29	0.35	0.25**
COMP	0.53	0.73	1.00	0.45	0.69	0.10 ⁺	0.20*	0.18*	0.20*	0.21*	0.15 ⁺
SPELL	0.40	0.69	0.57	1.00	0.83	0.11 ⁺	0.05 ⁺	0.17 ⁺	0.21*	0.47	0.28**
DISCR	0.52	0.95	0.84	0.84	1.00	0.14 ⁺	0.12 ⁺	0.23	0.30	0.43	0.28**
RAN COL	0.11 ⁺	0.12 ⁺	0.14 ⁺	0.09 ⁺	0.13 ⁺	1.00	0.41	0.45	0.40	0.33	0.32
RAN PIC	0.06 ⁺	0.08 ⁺	0.15 ⁺	0.10 ⁺	0.11 ⁺	0.46	1.00	0.43	0.31	0.25	0.29
RAN NUM	0.16 ⁺	0.42	0.35	0.47	0.46	0.32	0.22*	1.00	0.68	0.44	0.41
RAN LET	0.01 ⁺	0.37	0.30	0.38	0.40	0.16 ⁺	0.19*	0.50	1.00	0.45	0.29
CPS	0.34	0.38	0.28	0.55	0.45	0.11 ⁺	0.17 ⁺	0.42	0.40	1.00	0.62
IPT	0.31	0.23*	0.19*	0.43	0.31	0.14 ⁺	0.21*	0.42	0.33	0.63	1.00
Co-twins with a negative history of reading difficulties, who have data at follow-up (<i>n</i> =59)											
FSIQ	1.00	0.37**	0.63	0.22 ⁺	0.50	0.29*	0.11 ⁺	0.29*	0.32*	0.25 ⁺	0.45
REC	0.46	1.00	0.51	0.49	0.94	0.03 ⁺	-0.18 ⁺	0.03 ⁺	-0.09 ⁺	0.08 ⁺	-0.02 ⁺
COMP	0.54	0.80	1.00	0.16 ⁺	0.66	0.11 ⁺	-0.03 ⁺	0.11 ⁺	0.14 ⁺	0.02 ⁺	0.21 ⁺
SPELL	0.32*	0.60	0.46	1.00	0.68	-0.02 ⁺	-0.07 ⁺	-0.04 ⁺	-0.02 ⁺	0.22 ⁺	0.04 ⁺
DISCR	0.50	0.96	0.85	0.77	1.00	0.05 ⁺	-0.15 ⁺	0.03 ⁺	-0.04 ⁺	0.13 ⁺	0.05 ⁺
RAN COL	0.34**	0.34**	0.22 ⁺	0.26*	0.33*	1.00	0.45	0.55	0.52	0.38**	0.37**
RAN PIC	0.31*	0.19 ⁺	0.28*	-0.07 ⁺	0.16 ⁺	0.22 ⁺	1.00	0.43	0.43	0.17 ⁺	0.22 ⁺
RAN NUM	0.21 ⁺	0.22 ⁺	0.05 ⁺	0.23 ⁺	0.21 ⁺	0.29*	0.35**	1.00	0.72	0.37**	0.30*
RAN LET	0.16 ⁺	0.32*	0.17 ⁺	0.30*	0.32*	0.31*	0.51	0.64	1.00	0.46	0.35**
CPS	0.35*	0.40**	0.28*	0.39**	0.42	0.43	0.29*	0.46	0.55	1.00	0.52
IPT	0.45	0.29*	0.27*	0.33*	0.34**	0.47	0.16 ⁺	0.31*	0.41	0.40**	1.00

For all correlations $p < 0.001$, except where otherwise noted.

* $p < 0.05$

** $p < 0.01$

⁺ $p > 0.05$

^a This table is provided as a baseline to compare correlations at initial assessment for control subjects with and without follow-up data—most do not have follow-up data, thus there are no above diagonal elements.

Table 4 Raw means and SDs for cognitive and achievement measures at initial testing

Measure	+HistRD				-HistRD (co-twins)				Control			
	MZ		DZ		MZ		DZ		MZ		DZ	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
FSIQ ^a	99.52	11.33	100.16	10.90	104.64	9.41	113.00	12.00	114.90	10.64	114.67	11.07
REC	37.38	10.53	37.01	9.24	41.79	7.04	49.58	10.56	55.37	10.51	53.27	10.06
COMP	38.86	10.90	38.22	11.18	42.86	9.02	49.56	9.32	53.55	9.93	52.27	8.39
SPELL	37.98	10.11	38.30	8.91	45.21	7.02	49.62	8.86	55.43	9.89	52.71	9.71
RAN COL	20.50	4.28	19.49	4.13	21.21	3.62	22.16	4.49	24.21	4.68	24.96	4.78
RAN PIC	17.54	3.33	17.95	3.06	18.57	3.82	19.22	3.99	20.12	4.00	20.77	3.35
RAN NUM	29.04	5.50	29.08	5.14	31.29	6.21	32.96	7.09	35.10	7.51	34.67	6.29
RAN LET	26.52	6.58	25.72	6.87	31.21	4.56	30.60	6.87	33.21	7.29	33.65	6.31
CPS	14.28	7.43	13.80	5.74	17.00	6.40	19.11	8.17	22.19	8.83	20.35	7.12
IPT	48.62	11.73	46.16	11.27	52.50	10.81	56.31	13.31	58.98	15.31	58.08	14.29
AGE	10.28	1.86	10.11	1.68	10.31	1.81	10.30	1.95	10.67	1.80	10.90	2.13
N	50		74		14		45		42		52	

+HistRD Positive history of reading difficulties, -HistRD negative history of reading difficulties

^aFSIQ scores are based on published scaled scores

Preliminary Genetic Analyses Although the number of MZ and DZ twin pairs is still small for genetic analyses, preliminary analyses were conducted to assess the etiology of the stability of individual differences in reading performance as measured by the discriminant score (DISCR), our most stable measure. For these analyses, a Cholesky decomposition (Figure 1) was fitted separately to data from twin pairs in which at least one member had a history of reading difficulties (+HistRD), and from control twin pairs in which neither member had a history of reading difficulties (CTL). Although structural

Table 5 Raw means and SDs for cognitive and achievement measures at follow-up testing

Measure	+HistRD				-HistRD (co-twins)				Control			
	MZ		DZ		MZ		DZ		MZ		DZ	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
FSIQ ^a	96.38	10.92	96.03	11.63	99.71	12.86	107.13	11.26	111.38	14.32	110.08	9.46
REC ^b	78.30	12.91	79.43	12.35	89.64	6.20	91.58	6.20	93.52	4.81	92.23	4.35
COMP ^b	80.00	10.26	78.85	10.67	85.50	9.69	88.43	8.49	92.07	6.56	92.47	5.32
SPELL ^b	78.69	11.60	79.68	10.93	88.00	5.66	89.41	6.06	92.02	5.38	91.50	4.81
RAN COL	28.38	6.19	27.53	4.31	30.07	3.73	30.96	5.26	31.21	5.15	32.29	4.42
RAN PIC	22.14	3.28	23.30	3.66	24.07	4.14	24.24	3.71	24.88	4.40	25.87	3.53
RAN NUM	38.66	6.50	39.50	6.37	42.07	8.52	43.13	8.38	43.29	8.44	46.04	6.74
RAN LET	36.14	6.16	36.42	6.12	41.14	6.87	41.09	7.09	41.10	6.91	42.40	6.00
CPS	26.30	7.39	27.00	6.36	30.50	6.22	34.59	8.16	37.62	8.85	34.35	6.50
IPT	74.92	11.97	74.64	12.84	80.29	11.65	80.32	11.52	84.33	10.13	85.37	10.37
AGE	15.85	2.30	15.68	1.76	15.81	1.94	15.84	2.03	16.29	2.11	16.66	2.41
N	50		74		14		45		42		52	

+HistRD Positive history of reading difficulties, -HistRD negative history of reading difficulties

^aFSIQ scores are based on published scaled scores; because these newer editions were normed more recently than the WISC-R and WAIS-R administered at the initial assessment, mean follow-up FSIQ scores are slightly lower than initial FSIQ scores.

^bThe numbers of items differ for the PIAT and PIAT-R; thus, the raw scores at initial and follow-up assessment are not directly parallel.

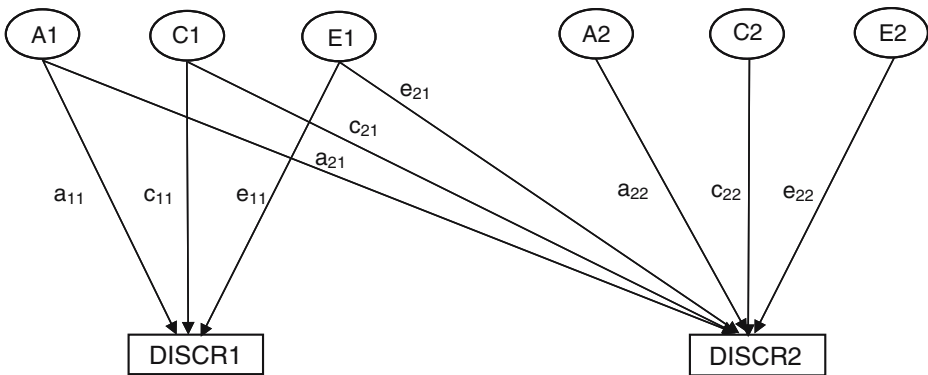


Fig. 1 Developmental genetic/environmental Cholesky

equation modeling (SEM) is generally not appropriate for use with selected samples, when using SEM techniques with CLDRC and LTSRD data, selection criteria are relaxed so that selection is based only on having/not having a history of reading difficulties. As a result, scores of both groups approximate normal distributions, making them acceptable for the application of maximum likelihood estimation procedures.

“Raw” data (i.e., data from individuals, rather than summary data such as covariance matrices) were analyzed using the Mx Statistical Modeling package (Neale, Boker, Xie, & Maas, 2003). The use of raw data, rather than covariance matrices, facilitates the use of all available data. Raw maximum likelihood estimation results in the calculation of twice the negative log-likelihood ($-2LL$) for each twin pair, and the summation of these across all twin pairs. For model comparisons the difference between $-2LL$ for the two models is distributed asymptotically as a chi-square, with degrees of freedom (df) equal to the difference between the number of free parameters estimated for the two models. For the current analyses, the full twin model included a total of 34 estimated parameters, 17 each for the +HistRD and CTL groups: i.e., nine path coefficients and eight means (including the means of twins and co-twins at each of two time points for MZ and DZ groups), required when modeling raw data rather than covariance matrices. The solutions of +HistRD and CTL groups were then equated, and the difference between the $-2LL$'s evaluated to determine if the covariance structures for the +HistRD and CTL groups were significantly different. Additional model comparisons were conducted to determine the significance of individual parameters or groups of parameters.

Results and Discussion

Phenotypic Analyses

Phenotypic correlations among the variables, within each time point, were calculated separately for twins with a history of reading difficulties, for their co-twins with no history of reading difficulties and for control twins (Table 3). Although the intercorrelations varied in magnitude among the groups, some general patterns emerged. In general, the correlations among the reading measures were moderate to high, with our composite reading measure, DISCR, correlating most highly (0.91–0.96) with REC.

Similarly, correlations among measures of processing speed were moderate. Measures of reading were also correlated with measures of processing speed, with CPS correlating most highly with the reading measures. However, RAN numbers and letters correlated moderately with the reading measures among the +HistRD group. With the exception of the -HistRD group, the lowest correlations with reading were obtained for RAN colors and pictures.

Sample Representativeness Control subjects participating in follow-up testing scored higher on their initial tests of IQ and reading achievement than those who did not participate in the follow-up study ($p \leq .02$ for IQ and $p \leq .001$ for reading), with effect sizes (e.s.) of 0.44 and 0.76, respectively. However, a comparison of subjects meeting proband criteria revealed no significant difference between IQ scores of reading-disabled probands retested and scores of those not retested ($p \geq .15$, e.s.=.23) and only a marginally significant difference in reading scores ($p = .05$, e.s.=.33).

Phenotypic Stability Standardized means and SDs at initial testing are presented in Table 6 separately for MZ and DZ twins with a positive history of reading difficulties (+HistRD), their co-twins with a negative history of reading difficulties (-HistRD) and control twins. As expected, both MZ and DZ twins with a positive history of reading difficulties performed more poorly on measures of cognitive ability and reading achievement than did their co-twins with no history of reading difficulties and controls. Both MZ and DZ twins with a positive history scored significantly lower than controls on all measures ($p < .001$). These differences were most pronounced on measures of reading, especially the discriminant score (DISCR), on which those with a history of reading difficulties scored about 1.7 SD units below the full control sample, and more than 2 SD units below those control twins participating in the follow-up study. Co-twins with a negative history also performed more poorly than controls, but these differences

Table 6 Means and SDs for cognitive and achievement measures at initial testing^a

Measure	+HistRD				-HistRD (co-twins)				Control			
	MZ		DZ		MZ		DZ		MZ		DZ	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
FSIQ	-1.25	1.03	-1.19	0.99	-0.78	0.85	-0.03	1.09	0.15	0.97	0.13	1.01
REC	-1.62	1.16	-1.57	1.09	-0.94	1.40	0.10	1.22	0.57	1.13	0.14	1.11
COMP	-1.35	1.14	-1.33	1.32	-0.77	1.38	0.12	1.15	0.37	1.07	0.07	1.14
SPELL	-1.24	0.97	-1.17	0.94	-0.44	0.96	0.16	0.95	0.68	0.86	0.26	1.03
DISCR	-1.73	1.18	-1.67	1.15	-0.89	1.34	0.13	1.14	0.67	1.08	0.18	1.13
RAN COL	-0.41	0.91	-0.60	0.93	-0.25	0.61	0.01	0.94	0.35	0.98	0.47	1.05
RAN PIC	-0.66	0.96	-0.47	0.91	-0.33	0.95	-0.09	1.11	0.05	1.05	0.20	0.99
RAN NUM	-0.53	0.82	-0.47	0.74	-0.17	0.99	0.14	0.80	0.33	0.98	0.16	0.97
RAN LET	-0.79	1.01	-0.87	1.07	0.06	0.90	-0.03	0.94	0.26	0.95	0.27	1.05
CPS	-0.77	0.92	-0.77	0.90	-0.28	1.03	0.17	0.99	0.49	1.00	0.02	0.80
IPT	-0.57	0.86	-0.71	0.95	-0.22	1.23	0.20	0.90	0.20	0.81	0.01	0.87
AGE	10.28	1.86	10.11	1.68	10.31	1.81	10.30	1.95	10.67	1.80	10.90	2.13
N	50		74		14		45		42		52	

Scores are age-adjusted^a and standardized against the mean of 1,284 control subjects in the CLDRC sample. +HistRD Positive history of reading difficulties, -HistRD negative history of reading difficulties

^a FSIQ and PIAT subtests are based on published scaled and age standard scores and have not been age-adjusted within the current sample, but have been standardized as described above.

were significant only among MZ co-twins for some measures of reading ($p < .002$ for REC and SPELL; $p < .001$ for DISCR) and for IQ ($p < .001$). Scores of DZ co-twins with a negative history regressed more toward the mean of the control sample, as might be expected if the trait of interest is genetically influenced. Raw means of initial scores are provided in Table 4.

Similarly, at follow-up (Table 7), those twins with a positive history of reading difficulties scored lower on all measures of cognitive ability and reading achievement than did control twins ($p < .001$ for both MZ and DZ twins). Again, this was especially pronounced for the measures of reading, with the mean DISCR score about 1.6 SD units below the mean of the control twins. These results suggest persistence of reading difficulties 5–6 years after initial assessment. However, it is noteworthy that relative to the control sample, the reading scores of those with a history of reading difficulties improved slightly over the 5–6 year interval, indicating that although difficulties persist, there is no evidence that children with reading difficulties fall further behind their peers in reading performance over time when considering age standard scores. For co-twins with no history of reading difficulties, mean follow-up scores were lower than those of control twins, but this difference was significant only for IQ ($p < .004$) among MZ co-twins. Raw means of follow-up scores are provided in Table 5.

Table 8 provides the stability correlations between measures assessed at the initial and follow-up sessions. As noted previously, scores of MZ and DZ twins have been combined within positive and negative history due to the small sample sizes of some of the groups. In addition to the persistence of deficits among those with a history of reading difficulties, individual differences in performance on measures of reading and cognitive ability are highly stable in all groups, with DISCR evidencing the greatest stability (average correlation of 0.80 across groups). These stability correlations are high in spite of the somewhat restricted ranges within the positive and negative history groups.

Table 7 Means and standard deviations for cognitive and achievement measures at follow-up testing^a

Measure	+HistRD				-HistRD (co-twins)				Control			
	MZ		DZ		MZ		DZ		MZ		DZ	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
FSIQ	-1.21	0.92	-1.24	0.98	-0.93	1.09	-0.30	0.95	0.06	1.21	-0.05	0.80
REC	-1.42	0.78	-1.32	0.87	-0.38	1.05	-0.06	0.99	0.29	1.10	-0.23	0.86
COMP	-1.26	0.87	-1.27	0.95	-0.64	1.07	-0.32	1.09	0.02	0.92	-0.01	1.07
SPELL	-1.42	0.99	-1.32	0.99	-0.46	1.10	-0.23	1.15	0.11	1.05	-0.09	0.96
DISCR	-1.66	0.90	-1.54	0.92	-0.54	1.03	-0.16	1.01	0.23	1.06	-0.19	0.91
RAN COL	-0.65	1.38	-0.87	0.92	-0.33	0.74	-0.12	1.10	-0.13	1.05	0.10	0.95
RAN PIC	-0.79	0.91	-0.46	0.92	-0.28	1.02	-0.24	0.90	-0.12	1.06	0.09	0.95
RAN NUM	-0.74	0.92	-0.64	0.78	-0.31	1.14	-0.16	1.03	-0.19	1.05	0.16	0.93
RAN LET	-0.83	1.02	-0.81	0.92	-0.07	1.08	-0.07	1.05	-0.11	1.04	0.09	0.97
CPS	-1.29	1.28	-1.21	1.22	-0.66	1.31	0.06	1.26	0.35	1.00	-0.28	0.92
IPT	-0.91	1.36	-0.96	1.48	-0.35	1.59	-0.35	1.09	-0.02	0.99	0.02	1.01
AGE	15.85	2.30	15.68	1.76	15.81	1.94	15.84	2.03	16.29	2.11	16.66	2.41
N	50		74		14		45		42		52	

Scores are age-adjusted^a and standardized against the mean of 1,284 control subjects in the CLDRC sample. +HistRD Positive history of reading difficulties, -HistRD negative history of reading difficulties

^a FSIQ and PIAT subtests are based on published scaled and age standard scores and have not been age-adjusted within the current sample, but have been standardized as described above.

Table 8 Correlations between measures at initial and follow-up testing*

Measures	+HistRD	-HistRD (co-twins)	Controls
FSIQ	0.75	0.70	0.72
READING REC	0.77	0.70	0.69
READING COMP	0.57	0.47	0.66
SPELLING	0.62	0.42	0.55
DISCR	0.84	0.77	0.78
RAN COLORS	0.58	0.49	0.62
RAN PICTURES	0.49	0.57	0.51
RAN NUMBERS	0.47	0.56	0.61
RAN LETTERS	0.50	0.42	0.53
CPS	0.72	0.73	0.63
IPT	0.59	0.69	0.40
N	124	59	94

* $p \leq 0.001$ unless otherwise noted

+HistRD Positive history of reading difficulties; - HistRD negative history of reading difficulties

Preliminary Genetic Analyses

To assess the etiology of the stability of reading performance (DISCR), a Cholesky decomposition was fit separately to data from twin pairs with and without a history of reading difficulties. Table 9 presents the solutions for the two groups.

Although genetic parameter estimates are large for both the +HistRD and CTL groups, there appear to be stronger genetic influences on both individual differences in, and on the stability of, reading performance in the +HistRD group than in the CTL group. However, the genetic correlations (r_G) between initial and follow-up measures are, or approximate, unity for both groups ($r_G = (a_{11} \times a_{21}) / (a_{DISCR1} \times a_{DISCR2})$), indicating that the same

Table 9 Results of genetic/environmental Cholesky for twin pairs with a history of reading difficulties (+HistRD) and Control (CTL) pairs

Standardized Genetic Cholesky (c.i.) and components of variance						
,	+HistRD			CTL		
	A1	A2	a ²	A1	A2	a ²
DISCR1	0.89 (0.73, 0.95)		0.79 (0.53, 0.90)	0.60 (0.22, 0.87)		0.36 (0.05, 0.76)
DISCR2	0.84 (0.67, 0.91)	0.00 (-0.32, 0.32)	0.70 (0.45, 0.82)	0.62 (0.26, 0.93)	0.00 (-0.33, 0.33)	0.39 (0.07, 0.87)
Standardized Shared Environmental Cholesky (c.i.) and components of variance						
,	+HistRD			CTL		
	C1	C2	c ²	C1	C2	c ²
DISCR1	0.18 (-0.49, 0.49)		0.03 (0.00, 0.24)	0.58 (-0.79, 0.79)		0.34 (0.00, 0.63)
DISCR2	-0.06 (-0.42, 0.42)	0.00 (-0.31, 0.31)	0.00 (0.00, 0.18)	0.67 (-0.86, 0.86)	0.00 (-0.31, 0.31)	0.45 (0.00, 0.75)
Standardized Nonshared Environmental Cholesky (c.i.) and components of variance						
,	+HistRD			CTL		
	E1	E2	e ²	E1	E2	e ²
DISCR1	0.42 (0.32, 0.57)		0.18 (0.10, 0.33)	0.55 (0.42, 0.70)		0.30 (0.17, 0.49)
DISCR2	0.33 (0.17, 0.53)	0.44 (0.36, 0.52)	0.30 (0.18, 0.50)	0.00 (-0.12, 0.18)	0.40 (0.30, 0.53)	0.16 (0.09, 0.29)

genetic influences are involved in reading performance at both time points. This is consistent with previous findings from adoption studies of individual differences in reading (e.g., Wadsworth et al., 2006). Moreover, given phenotypic stability correlations (r_p) of 0.87 when the scores of the +HistRD and -HistRD co-twins are combined, and 0.78 for the control twins, these shared genetic influences account for 86% and 49% of the phenotypic correlations between the two assessments, i.e., $(a_{11} \times a_{21})/r_p$. Conversely, stronger shared environmental influences are suggested for the CTL group. Furthermore, shared environmental influences appear to contribute to stability in this group, accounting for approximately 50% of the phenotypic stability correlation, whereas they do not in the +HistRD group. However, the confidence intervals for the shared environmental parameters include zero for both groups; thus, a larger sample will be necessary to determine if shared environmental influences are important for individual differences in reading performance as measured by DISCR. Nonshared environmental influences appear to contribute to stability in the +HistRD group, but not in the CTL group.

Because the genetic and shared environmental correlations reached unity for both groups, and the a_{22} and c_{22} paths reached zero (within rounding error), the model was tested for “Cholesky problems” (Carey, 2005) and was also re-parameterized as a correlated factor model re-estimating the genetic and environmental correlations. Using these methods, there was no evidence of significant Cholesky problems, and the correlated factor model also estimated the genetic and shared environmental correlations at 1.0.

Components of variance in reading achievement derived from the full model allowing separate solutions for +HistRD and CTL groups are also presented in Table 9. Here, the group differences are somewhat more apparent, with initial heritability in the CTL group (0.36) less than half that in the +HistRD group (0.79), and shared environmental influences similar in magnitude to heritability estimates for this group (0.34), but with confidence intervals including zero. Genetic and shared environmental influences are similar at the two time points for both groups.

Model comparisons are given in Table 10. When the +HistRD and CTL solutions were equated, a highly significant deterioration in model fit resulted ($\chi^2=27.91$, $df=9$, $p<.001$, model 2), indicating that the covariance structures for the two groups are significantly different. More importantly, however, genetic influences are important for individual differences in the stability of reading for both groups. For neither group could the path from A1 to DISCR at follow-up (path a_{21}) be dropped (models 3 and 4, $p<.001$ and $.02$ for +HistRD and CTL groups, respectively). However, although genetic influences are significant for both groups and the confidence intervals overlap, the genetic covariance structures, taken together, could not be equated (model 5, $p<.05$). In contrast, all shared environmental influences could be omitted from the model (models 6 and 7), indicating

Table 10 Model comparisons for genetic/environmental Cholesky

Model	-2LL	NPAR	χ^2	df	p
1. Full Model	1,306.11	34			
2. Equate +HistRD and CTL groups	1,334.02	25	27.91	9	<0.001
3. Drop A_{21} for +HistRD group	1,322.75	33	16.64	1	<0.001
4. Drop A_{21} for CTL group	1,311.76	33	5.65	1	<0.02
5. Equate A matrices for +HistRD and CTL groups	1,314.10	31	7.99	3	<0.05
6. Drop C for +HistRD group	1,307.91	31	1.80	3	>0.5
7. Drop C for CTL group	1,309.05	31	2.94	3	>0.3
8. Drop E_{21} for +HistRD group	1,328.87	33	22.76	1	<0.001

All sub-models are compared to the full model (model 1).

that the estimates of shared environmental influence do not differ significantly from zero for either group. Nonshared environmental influences also differ for the two groups. Whereas these influences are specific to each occasion for the CTL group, for the +HistRD group, they contribute significantly to stability, and this path could not be dropped from the model (model 8, $p < .001$).

Conclusions

The purposes of the present study were to provide a preliminary description of the cognitive test performance and reading achievement of the current sample of participants in the Colorado Longitudinal Twin Study of Reading Disability, at initial and follow-up sessions, and to provide a preliminary assessment of the stability of reading performance in this sample and the etiology of this stability.

Our preliminary examination of the descriptive statistics and stability correlations for cognitive test performance and reading achievement indicates that over the 5- to 6-year interval between assessments, cognitive and reading performance are highly stable. Those subjects with a positive history of reading difficulties had substantial mean group deficits relative to control subjects on all measures at initial assessment, and significant mean deficits remained at follow-up. However, for reading achievement (as for most measures), there was no increase in the group deficit relative to controls at the follow-up assessment. In fact, the deficit was slightly reduced. Thus, based on these preliminary findings, there is no evidence in this sample of the “Matthew effect” found in some previous studies (e.g., Maughan et al., 1994; Bast & Reitsma, 1998; Speece & Ritchey, 2005; Stanovich, 1986).

The stability noted for all cognitive and achievement measures was most pronounced for the DISCR reading score, whose average stability correlation across groups was 0.80. When these data were subjected to structural equation modeling, allowing separate solutions for twin pairs with and without a positive history of reading difficulties, maximum likelihood estimates indicated that for both groups, genetic influences contributed significantly to the variance in reading performance at each assessment and to the stability of individual differences in reading performance. However, the magnitude of the genetic contribution differed between +HistRD and CTL groups. While genetic correlations reached unity for both groups, suggesting that the same genetic influences are manifested at both time points, shared genetic influences accounted for 86% and 49% of the phenotypic correlations between the two assessments for +HistRD and CTL groups, respectively. In addition, nonshared environmental influences also contributed to stability in the +HistRD group, but not in the CTL group. Whether these differences will be robust to increases in the follow-up sample remains to be seen. Although shared environmental influences were nonsignificant in both groups, estimates of shared environment were moderate for the CTL group and may prove to be important if similar estimates continue to be obtained as the sample size increases.

The current findings of phenotypic stability of reading performance and its etiology are highly similar to those of Harlaar et al. (2007) and Wadsworth et al. (2006), who examined reading scores of population samples of twins and adoptive families. However, whereas the results of the current study and those of Wadsworth et al. (2006) suggested no new genetic involvement beyond the first assessment, results of Harlaar et al. (2007) observed new genetic influence at each age. This difference in results could be due to differences in the ages of the samples or measures used. Whereas the participants in the TEDS sample were

tested at ages 7, 9, and 10, those of the CAP sample were tested at 7, 12, and 16, and those in the current sample comprised a range of ages averaging 10.4 years of age at initial assessment and 16.0 years at follow-up. Thus, it may be that new genetic influences are activated in middle childhood, but are stable by adolescence. If this is the case, the CAP and LTSRD samples would be less likely to find evidence of new genetic influences at these later ages. As the TEDS sample reaches middle to late adolescence and the LTSRD sample size increases, it should be possible to test this hypothesis.

Another difference between these studies concerns the measures of reading and methods of assessment. Whereas reading performance in the CAP and LTSRD was assessed using in-person administration of the PIAT and/or PIAT-R reading and spelling subtests, Harlaar et al. (2007) measured reading performance of TEDS participants using teacher assessments based on group-administered tests at ages 7 and 9, and the addition of a web-based version of the PIAT-R at age 10. These differences in measures and methods of test administration could lead to different findings regarding “new” genetic influences at the later ages, especially when both measures and methods differ between assessments. For example, the focus of the web-based measure added at age 10 may have differed somewhat from that of teacher assessments at ages 7 and 9, resulting in the manifestation of “new” genetic influences at age 10.

Limitations and Issues

Although the similarity between these preliminary results and those of previous studies is encouraging, there are limitations of the present study which suggest that these results should be interpreted with caution. First, based on the comparison of the initial scores of those subjects participating in follow-up assessment and those who have not participated, there is some evidence that the follow-up sample is not entirely representative of the full CLDRC sample. Non-proband subjects, including control subjects, who have participated in follow-up assessment, had significantly higher initial scores on measures of IQ and reading performance than those who have not participated in the follow-up. Clearly, there is the opportunity for self-selection at both assessments. Subjects who perform better on cognitive and achievement tests may be more willing to be tested or may be more likely to volunteer to be research subjects. However, there was no significant difference in scores between those reading-disabled *probands* who have participated and those who have not.

Another sampling issue involves the inclusion of subjects with a prior history of ADHD symptoms. In this study and in the CLDRC, twin pairs are invited to participate if either member of the pair has a history of reading difficulties or ADHD symptoms. Among twins with a positive history of reading difficulties in this sample, 22% had a history of ADHD symptoms vs 10% of those with a negative history of reading difficulties ($p > .05$). This is not surprising given that RD has been shown to be significantly associated with ADHD, as well as with other externalizing behaviors (e.g., McGee et al., 2002; Stevenson et al., 2005; Willcutt et al., 2002, 2003). However, the current analyses excluded any twin pairs in which one or both members had a history of ADHD symptoms, but *no* history of reading difficulties, reducing the effects of this potential confound. As the sample grows, it will be possible to examine these ascertainment groups separately to compare performance of subjects with RD only, RD + ADHD, and ADHD only on our battery of cognitive abilities, reading component processes, achievement and behavioral inventories, both cross-sectionally and longitudinally.

Another limitation involves the use of different versions/editions of our standardized tests: the more recent revisions were used for the follow-up assessment, whereas earlier versions of

the tests were used in the initial assessment. Before standardization against our control samples, scores on these test revisions had slightly lower means on the published age-standard and scaled scores for all groups than did the earlier editions of the tests. However, published correlations between editions are high, and all test scores were standardized against the control samples at each time point, reducing any effect of these differences.

Another measurement issue is that for the WISC and WAIS between initial and follow-up assessments, many of our subjects have aged out of the WISC-R and into the WAIS-III. However, because these are isomorphic, age-adjusted measures, we do not foresee any difficulties in analyzing the two simultaneously.

Finally, the current sample is still small for conducting reliable genetic analyses, particularly when more stringent selection is imposed (e.g., Astrom et al., 2007). Thus, results are preliminary and should be interpreted with caution. As the sample increases, more rigorous tests of the etiology of stability and change in reading performance and in reading disability will be possible. However, it should be noted that use of bivariate or multivariate designs, such as the Cholesky model used in the current study, usually has the beneficial effect of increasing statistical power over simple univariate designs (Schmitz, Cherny & Fulker, 1998). Although power to detect moderate to strong genetic effects under a univariate model is somewhat low, it is increased under a bivariate or multivariate model. For example, given our current individual differences analysis sample of 32 MZ pairs in which at least one member has a history of reading difficulties, and 59 DZ pairs with a history of reading difficulties, the power to detect significant genetic effects of the magnitude obtained for DISCR, under a univariate model, is 0.69. However, under a bivariate model, power to detect significant genetic influence is >0.99. Thus, as our sample grows, power should be sufficient for detecting genetic influences on individual differences even for less heritable measures.

Future Directions

In addition to the kinds of analyses described here, the data are being analyzed in conjunction with data previously collected in the CLDRC to provide some of the first longitudinal assessments of the genetic and environmental etiologies of reading deficits and dimensions/subtypes of reading disability at two ages (e.g., Astrom et al., 2007; Hulslander, Olson & Wadsworth, submitted for review), as well as their longitudinal stability and comorbidity with ADHD (Wadsworth, Willcutt, Betjemann, Olson & DeFries, in preparation). We anticipate that with the increase in sample size, future analyses will confirm our preliminary findings of genetic influence on the stability of reading deficits and on their longitudinal comorbidity with ADHD. As noted above, it will also be possible to compare performance of subjects with RD only, RD + ADHD and ADHD only on our battery of cognitive abilities, reading component processes, achievement, and behavioral inventories. These data may also be used to test novel hypotheses about the longitudinal covariation of reading difficulties with measures of other psychopathology, reading, language, and perceptual processes and specific cognitive abilities (e.g., Betjemann et al., 2007). In addition, our expanded sample will eventually provide power sufficient to test $G \times E$ interactions that may increase susceptibility to RD and/or ADHD and to assess early adult outcomes of children with reading deficits or ADHD symptoms. Moreover, utilizing the genotypic data collected in the CLDRC, univariate and bivariate longitudinal QTL analyses of reading deficits and ADHD, reading component processes and other psychopathology will also be possible.

Given the breadth of measures administered to subjects in the CLDRC and LTSRD, we are in a unique position to explore many developmental issues which are not yet

well understood, related to reading and reading disability from childhood into adulthood. By further illuminating the influence of genes, environment, and how they interact to affect reading difficulties and concomitant behavior problems, and their long-term outcomes, this study affords the opportunity to inform not only future research, but intervention efforts and policy, eventually contributing to a more productive approach to educational, occupational, and mental health services.

Acknowledgments The Colorado Learning Disabilities Research Center is supported by grant HD-27802 from the National Institute of Child Health and Human Development (NICHD). The Longitudinal Twin Study of Reading Disability is supported by grant DC-05190 from the National Institute on Deafness and other Communication Disorders (NIDCD). The continued cooperation of the many families participating in the CLCRC and LTSRD and the work of the staff members of these projects is gratefully acknowledged.

References

- Achenbach, T. A., & Rescorla, L. A. (2001). *Manual for the ASEBA School-age Forms and Profiles*. Burlington, VT: University of Vermont, Research Center for Children, Youth, and Families.
- American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders: DSM-IV*. Washington, DC: American Psychiatric Association.
- Astrom, R. L., Wadsworth, S. J., & DeFries, J. C. (2007). Etiology of the stability of reading difficulties: The longitudinal Twin Study of Reading Disabilities. *Twin Research and Human Genetics*, 10, 434.
- Badian, N. A. (1999). Reading disability defined as a discrepancy between listening and reading comprehension: A longitudinal study of stability, gender differences, and prevalence. *Journal of Learning Disabilities*, 32, 138–148.
- Bailey, C. E., Manis, F. R., Pedersen, W. C., & Seidenberg, M. S. (2004). Variation among developmental dyslexics: Evidence from a printed-word-learning task. *Journal of Experimental Child Psychology*, 87, 125–154.
- Barkley, R. A., & Murphy, K. (1998). *Attention-deficit hyperactivity disorder: A clinical workbook* (2nd ed.). New York: Guilford Press.
- Bast, J., & Reitsma, P. (1998). Analyzing the development of individual differences in terms of Matthew effects in reading: Results from a Dutch longitudinal study. *Developmental Psychology*, 34, 1373–1399.
- Betjemann, R. S., Willcutt, E. G., Olson, R. K., Keenan, J. M., DeFries, J. C., & Wadsworth, S. J. (2007). Etiology of the stability of reading comprehension and its relation with word reading. *Reading and Writing: An International Journal* (in press).
- Byrne, B., Samuelsson, S., Wadsworth, S., Hulslander, J., Corley, R., DeFries, J. C., et al. (2007). Longitudinal twin study of early literacy development: Preschool through Grade 1. *Reading and Writing: An Interdisciplinary Journal*, 20, 77–102.
- Carey, G. (2005). Cholesky problems. *Behavior Genetics*, 35, 653–65.
- 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., 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.
- DeFries, J. C., & Fulker, D. W. (1985). Multiple regression analysis of twin data. *Behavior Genetics*, 15, 467–473.
- DeFries, J. C., & Fulker, D. W. (1988). Multiple regression analysis of twin data: Etiology of deviant scores versus individual differences. *Acta Geneticae Medicae et Gemellologiae*, 37, 205–216.
- DeFries, J. C., Plomin, R., Vandenberg, S. G., & Kuse, A. R. (1981). Parent-offspring resemblance for cognitive abilities in the Colorado Adoption Project: Biological adoptive, and control parents and one-year-old children. *Intelligence*, 5, 245–277.
- DeFries, J. C., Singer, S. M., Foch, T. T., & Lewitter, F. I. (1978). Familial nature of reading disability. *British Journal of Psychiatry*, 132, 361–367.
- Denckla, M. B., & Rudel, R. G. (1976). Rapid “automatized” naming (R.A.N): Dyslexia differentiated from other learning disabilities. *Neuropsychologia*, 14, 471–479.
- Dunn, L. M., & Markwardt, F. C. (1970). *Examiner’s manual: Peabody individual achievement test*. Circle Pines, MN: American Guidance Service.
- Fisher, S. E., & DeFries, J. C. (2002). Developmental dyslexia: Genetic dissection of a complex cognitive trait. *Nature Reviews Neuroscience*, 3, 767–780.

- French, J. W., Ekstrom, R. G., & Price, L. A. (1963). *Manual for a kit of reference tests for cognitive factors*. Princeton, NJ: Educational Testing Service.
- Gayán, J., Willcutt, E. G., Fisher, S. E., Francks, C., Cardon, L. R., Olson, R. K., et al. (2005). Bivariate linkage scan for reading disability and attention-deficit/hyperactivity disorder localizes pleiotropic loci. *Journal of Child Psychology and Psychiatry*, *46*, 1045–1056.
- Griffiths, Y. M., & Snowling, M. J. (2002). Predictors of exception word and nonword reading in dyslexic children: The severity hypothesis. *Journal of Educational Psychology*, *94*, 34–43.
- Harlaar, N., Dale, P. S., & Plomin, R. (2007). From learning to read to reading to learn: Substantial and stable genetic influence. *Child Development*, *78*, 116–131.
- Harlaar, N., Spinath, F. M., Dale, P. S., & Plomin, R. (2005). Genetic influences on early word recognition abilities and disabilities: a study of 7-year-old twins. *Journal of Child Psychology and Psychiatry*, *46*, 373–384.
- Hulslander, J., Olson, R., & Wadsworth, S. (submitted for review). Longitudinal stability of reading skill profiles.
- Kovacs, M. (1988). *The children's depression inventory manual*. New York: Multi-Health Systems, Inc.
- Lyytinen, H., Ahonen, T., Eklund, K., Guttorm, T. K., Laakso, M.-L., Leinonen, S., et al. (2001). Developmental pathways of children with and without familial risk for dyslexia during the first years of life. *Developmental Neuropsychology*, *20*(2), 539–558.
- Markwardt Jr., F. C. (1989). *Peabody individual achievement test-revised*. Circle Pines, MN: American Guidance Service.
- Maughan, B., Hagell, A., Rutter, M., & Yule, W. (1994). Poor readers in secondary school. *Reading and Writing: An Interdisciplinary Journal*, *6*, 125–150.
- McGee, R., Prior, M., Williams, S., Smart, D., & Sanson, A. (2002). The long-term significance of teacher-rated hyperactivity and reading ability in childhood: Findings from two longitudinal studies. *Journal of Child Psychology and Psychiatry*, *43*, 1004–1017.
- McKinney, J. D. (1994). Methodological issues in longitudinal research in learning disabilities. In S. Vaughn, & C. Bos (Eds.) *Research issues in learning disabilities: Theory, methodology, assessment, and ethics* (pp. 202–230). New York: Springer-Verlag.
- Meng, H., Smith, S. D., Hager, K., Held, M., Liu, J., Olson, R. K., et al. (2005). DCDC2 is associated with reading disability and modulates neuronal development in the brain. *Proceedings of the National Academy of Sciences*, *102*, 17053–17058.
- Moffitt, T. E. (2002). Teen-aged mothers in contemporary Britain. *Journal of Child Psychology and Psychiatry*, *43*, 727–742.
- Muthén, L. K., & Muthén, B. O. (1998–2007). *M-plus user's guide* (4th ed.). Los Angeles, CA: Muthén & Muthén.
- Neale, M. C., Boker, S. M., Xie, G., & Maes, H. H. (2003). *Mx: Statistical modeling*. VCU Box 900126, Richmond, VA 23298: Department of Psychiatry. 6th ed.
- Nichols, R. C., & Bilbro, W. C. (1966). The diagnosis of twin zygosity. *Acta Genetica*, *16*, 265–275.
- Olson, R. K., Datta, H., Gayán, J., & DeFries, J. C. (1999). A behavioral-genetic analysis of reading disabilities and component processes. In R. Klein, & P. McMullen (Eds.) *Converging methods for understanding reading and dyslexia* (pp. 133–151). Cambridge, MA: MIT Press.
- 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 Publishers.
- Olson, R., Forsberg, H., Wise, B., & Rack, J. (1994). Measurement of word recognition, orthographic, and phonological skills. In G. R. Lyon (Ed.) *Frames of reference for the assessment of learning disabilities: New views on measurement issues* (pp. 243–277). Baltimore: Paul H. Brookes Publishing Co.
- Petrill, S. A., Deater-Deckard, K., Thompson, L. A., Schatschneider, C., Dethorne, L. S., & Vandenberg, D. J. (2007). Longitudinal genetic analysis of early reading: The Western Reserve Reading Project. *Reading and Writing*, *20*, 127–146.
- Plomin, R., & Kovas, Y. (2005). Generalist genes and learning disabilities. *Psychological Bulletin*, *131*, 592–617.
- Raskind, M. H., Gerber, P. J., Goldberg, R. J., Higgins, E. L., & Herman, K. L. (1998). Longitudinal research in learning disabilities: Report on an international symposium. *Journal of Learning Disabilities*, *31*, 266–277.
- Raskind, M. H., Higgins, E. L., Goldberg, R. J., & Herman, K. L. (1999). Patterns of change and predictors of success in individuals with learning disabilities: Results from a twenty-year longitudinal study. *Thalamus*, *16*, 40–64.
- Reich, W., Welner, Z., & Herjanic, B. (1997). *Diagnostic Interview for Children and Adolescents—IV*. North Tawanda Falls, NY: Multi-Health System, Inc.

- Reynolds, C. R., & Richmond, B. O. (1985). *Revised children's manifest anxiety scale manual*. Los Angeles: Western Psychological Services.
- Rutter, M., Caspi, A., Fergusson, D., Horwood, L. J., Goodman, R., Maughan, B., et al. (2004). Sex differences in developmental reading disability: New findings from 4 epidemiological studies. *Journal of the American Medical Association*, *291*, 2007–2012.
- Satz, P., Buka, S. L., Lipsitt, L. P., & Seidman, L. (1998). The long-term prognosis of learning disabled children: A review of studies (1954–1993). In B. K. Shapiro, P. J. Accardo, & A. J. Capute (Eds.) *Reading disability: A view of the spectrum* (pp. 223–250). Timonium, MD: York Press.
- Scarborough, H. S., & Parker, J. D. (2003). Children's learning and teachers' expectations: Matthew effects in children with learning disabilities: Development of reading, IQ, and psychosocial problems from grade 2 to grade 8. *Annals of Dyslexia*, *53*, 47–71.
- Schmitz, S., Cherny, S. S., & Fulker, D. W. (1998). Increase in power through multivariate analyses. *Behavior Genetics*, *28*, 357–364.
- Schumacher, J., Hoffmann, P., Schmael, C., Schulte-Körne, G., & Nöthen, M. M. (2007). Genetics of dyslexia: The evolving landscape. *Journal of Medical Genetics*, *44*, 289–297.
- 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*, S50–57.
- Shaywitz, S. E., Escobar, M. D., Shaywitz, B. A., Fletcher, J. M., & Makuch, R. (1992). Evidence that dyslexia may represent the lower tail of a normal distribution of reading disability. *New England Journal of Medicine*, *326*, 144–150.
- Shaywitz, S. E., Fletcher, J. M., Holahan, J. M., Shneider, A. E., Marchione, K. E., Stuebing, K. K., et al. (1999). Persistence of dyslexia: The Connecticut Longitudinal Study at adolescence. *Pediatrics*, *104*, 1351–1359.
- Shaywitz, S. E., Shaywitz, B. A., Fulbright, R. K., Skudlarski, P., Mencl, W. E., Constable, T., et al. (2003). Neural systems for compensation and persistence: Young adult outcome of childhood reading disability. *Biological Psychiatry*, *54*, 25–33.
- Spee, D. L., & Ritchey, K. D. (2005). A longitudinal study of the development of oral reading fluency in young children at risk for reading failure. *Journal of Learning Disabilities*, *38*, 387–399.
- Spira, E. G., Bracken, S. S., & Fischel, J. E. (2005). Predicting improvement after first-grade reading difficulties: The effects of oral language, emergent literacy, and behavior skills. *Developmental Psychology*, *41*, 225–234.
- Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly*, *21*, 360–407.
- Stevenson, J., Langley, K., Pay, H., Payton, A., Worthington, J., Ollier, W., et al. (2005). Attention deficit hyperactivity disorder with reading disabilities: Preliminary genetic findings on the involvement of the ADRA2A gene. *Journal of Child Psychology and Psychiatry*, *46*, 1081–1088.
- Torgesen, J., Wagner, R., & Rashotte, C. A. (1999). *A Test of Word Reading Efficiency (TOWRE)*. Austin, TX: PRO-ED.
- Trzszniewski, K. H., Moffitt, T. E., Caspi, A., Taylor, A., & Maughan, B. (2006). Revisiting the association between reading achievement and antisocial behavior: New evidence of an environmental explanation from a twin study. *Child Development*, *77*, 72–88.
- Vogel, S. A. (1990). Gender differences in intelligence, language, visual-motor abilities, and academic achievement in students with learning disabilities: A review of the literature. *Journal of Learning Disabilities*, *23*, 44–52.
- Vukovic, R. K., & Siegel, L. S. (2006). The double-deficit hypothesis: A comprehensive analysis of the evidence. *Journal of Learning Disabilities*, *39*, 25–47.
- Wadsworth, S. J., Corley, R. P., Plomin, R., Hewitt, J. K., & DeFries, J. C. (2006). Genetic and environmental influences on continuity and change in reading achievement in the Colorado Adoption Project. In A. Huston, & M. Ripke (Eds.) *Developmental contexts of middle childhood: Bridges to adolescence and adulthood* (pp. 87–106). New York: Cambridge University Press.
- Wadsworth, S. J., & DeFries, J. C. (2005). Genetic etiology of reading difficulties in boys and girls. *Twin Research and Human Genetics*, *8*, 594–601.
- Wechsler, D. (1974). *Examiners' manual: Wechsler intelligence scale for children-revised*. New York: The Psychological Corporation.
- Wechsler, D. (1981). *Examiners' manual: Wechsler adult intelligence scale-revised*. New York: The Psychological Corporation.
- Wechsler, D. (1991). *Examiners' manual: The Wechsler intelligence scale for children* (3rd ed.). San Antonio, TX: The Psychological Corporation.
- Wechsler, D. (1997). *Examiners' manual: The Wechsler adult intelligence scale* (3rd ed.). San Antonio, TX: The Psychological Corporation.

- Wechsler, D. (2002). *Examiners' manual: The Wechsler individual achievement test* (2nd ed.). San Antonio, TX: The Psychological Corporation.
- Wilkinson, G. S. (1993). *Examiner's manual: Wide range achievement test-3*. Wilmington, DE: Jastak Associates-Wide Range Inc.
- Willcutt, E. G., DeFries, J. C., Pennington, B. F., Smith, S. D., Cardon, L. R., & Olson, R. K. (2003). Genetic etiology of comorbid reading difficulties. In R. Plomin, J. C. DeFries, I. W. Craig, & P. McGuffin (Eds.) *Behavioral genetics in the postgenomic era* (pp. 227–246). Washington, D.C.: American Psychological Association.
- Willcutt, E. G., & Pennington, B. F. (2000). Psychiatric comorbidity in children and adolescents with reading disability. *Journal of Child Psychology and Psychiatry*, *41*, 1039–1048.
- 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)*, *114*, 260–268.
- Williams, J., & O'Donovan, M. C. (2006). The genetics of developmental dyslexia. *European Journal of Human Genetics*, *14*, 681–689.