Genetic and Environmental Etiologies of Reading Disabilities: Analysis of data from the Colorado Learning Disabilities Research Center

by

Raven Astrom

RN, Front Range Community College, Westminster, Colorado, 1993

B.A., University of Colorado, Boulder, 2003

M.A., University of Colorado, Boulder, 2010

A thesis submitted to the

Faculty of the Graduate School of the

University of Colorado in partial fulfillment

Of the requirement for the degree of

Doctor of Philosophy

Department of Psychology and Neuroscience

2011

This thesis entitled:

Genetic and Environmental Etiologies of Reading Disabilities:
Analysis of data from the Colorado Learning Disabilities Research Center

written by Raven Leigh Astrom
has been approved for the Department of Psychology and Neuroscience

---

Michael C. Stallings, Ph.D.

Sally J. Wadsworth, Ph.D.

Erik G. Willcutt, Ph.D.

Mark A. Whisman, Ph.D.

Matthew B. McQueen,Ph.D.

Date:_____

The final copy of this thesis has been examined by the signatories, and we find that both the content and the form meet acceptable presentation standards of scholarly work in the above mentioned discipline.

IRB Protocol # 1096.30, 0201.30

Astrom, Raven Leigh (Ph.D., Psychology)

Genetic and Environmental Etiologies of Reading Disabilities:
Analysis of data from the Colorado Learning Disabilities Research Center

Thesis directed by Associate Professor Michael C. Stallings

Specific reading disability (RD) or dyslexia is often defined as an unexpected problem with learning to read despite having normal intelligence, no sensory impairments, and the opportunity to learn from reasonable instruction (Lyon et al., 2003). Research has shown that deficits in reading are both stable and heritable suggesting that genetic influences may be largely continuous throughout development.

This dissertation employs data from twins and their nontwin siblings in the Colorado Learning Disabilities Research Center (DeFries et al., 1997) and the Longitudinal Twin Study of Reading Disability (Wadsworth et al., 2007) to investigate factors which contribute to the heritability and stability of reading difficulties. Early twin studies compared "concordance" rates in pairs of identical and fraternal twins as a test for genetic etiology. However, DeFries and Fulker (1985, 1988) proposed fitting a multiple regression model to data from selected twin pairs to more rigorously assess genetic and environmental influences on extreme scores.

First, data from twin and nontwin siblings were fitted to DF multiple regression models to investigate the heritable nature of reading deficits in addition to examining evidence for a "special twin environment". Second, twin data was employed to examine the etiology of genetic and environmental influences on the stability of reading deficits. Third, we examined heritability and stability utilizing data from a larger sample of twin pairs and their nontwin siblings. Our

fourth study examined the differential etiology of genetic and environmental influences for reading disability as a function of gender.

Results from the first study indicated that reading deficits are substantially heritable; in addition, there were significant results for a special twin environment. Findings from the second study indicated that reading deficits were not only heritable, but also highly stable. Our third study suggested that reading disabilities are heritable and stable for both twins and their nontwin siblings. There was no finding for a special twin environment influencing the stability of reading deficits. The fourth study examined the etiology of the heritability and stability of reading deficits as a function of gender. Results were highly heritable and stable, however, for this sample there were no significant gender differences. Implications for these findings are discussed.

Dedication

To my husband, David Astrom

Acknowledgments

First and foremost, I want to convey my sincere gratitude to Michael C. Stallings, for supporting me and my research interests under difficult circumstances. I also wish to acknowledge the extraordinary support and guidance from Sally J. Wadsworth. In addition to Drs. Stallings and Wadsworth, I would like to thank Erik Willcutt, Mark Whisman and Matthew McQueen, for their support and education in the field of Behavior Genetics. It's impossible to succeed in graduate school without the support of your fellow graduate students and as such, I would like to thank Josh Bricker, Jesse Hawke, Melissa Munn-Chernoff, Rohan Palmer, and Joanna Vandever – my life is richer for having known each and every one of you.

Finally, none of this would have been possible without the support of my amazing children; Dusty, Shawn, Krissy, and Samantha – thank you for always believing in me. To my immediate family and my new children, you have all played an important part in helping me to achieve this goal – thank you so very much. Beyond my family, I am blessed with incredible friends – your constant support when I didn't think I could do anymore – has been such a gift – thank you. And to my husband, David – thank you for telling me that there was no crying in genetics – your constant love and support made all the difference.

This work would not have been possible without the generous funding of the Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD) grant HD027802. In addition to funding from the National Institute on Deafness and other Communication Disorders (NIDCD) grant DC005190. I am also grateful for the support by institutionaltraininggrantHD007829.

Contents

Chapter

List of Tables

Table

List of Figures

Figure

List of Appendices

Appendices

Chapter 1

Introduction

Specific reading disability (RD) or dyslexia is often defined as an unexpected problem with learning to read despite having normal intelligence, no sensory impairments, and the opportunity to learn from reasonable instruction (Lyon, Shaywitz, & Shaywitz, 2003). Approximately 80% of children diagnosed with learning disability are principally affected by poor reading ability. Reading performance is a normally distributed trait, with prevalence rates of dyslexia in school-age children ranging between 7% and 15% regardless of normal intelligence and adequate educational opportunity (Lyon, Fletcher, & Barnes, 2002; Pennington, Peterson, & McGrath, 2009; Shaywitz, Escobar, Shaywitz, Fletcher, & Makuch, 1992). There have been numerous studies of reading performance and dyslexia that have employed various methodologies to investigate both normal reading development and reading performance. Although reading difficulties have long been recognized as having significant negative effects on children and adults, evidence for reading disability as a valid diagnostic construct was slow to emerge (Lyon & Chhabra, 1996). The ability to establish an unbiased inclusionary definition and classification system for reading disabilities has been a significant challenge for the scientific community (Fletcher & Morris, 1986).

One aim of this dissertation is to discuss the importance of research in examining reading abilities and disabilities. We will discuss issues related to poor literacy and provide an overview of the current trends in reading intervention and outcomes. Further, we will provide an in depth review of the background and historical definitions of reading disability as well as to present the current working definition for RD. Following will be an overview of family and twin studies which have contributed to the understanding of reading difficulties, including evidence from

twin and nontwin sibling studies for the genetic and environmental etiologies of reading disability. In addition, we will discuss characteristics of reading disabilities specific to gender. Finally, we will present an overview of DeFries-Fulker (DF) multiple regression analysis and its uses for behavioral genetic studies of RD and to present a novel extension of DF analysis which provides a test of special twin environment.

Chapter 2 investigates the etiology of reading deficits in addition to employing a novel extension of DF analysis which incorporates data from both twins and their nontwin siblings to explore aspects of special twin environmental influences. Chapter 3 examines the etiology of stability of reading disabilities from a selected twin sample in the Colorado Learning Disabilities Research Center (CLDRC, DeFries, Filipek, Fulker, Olson, Pennington, Smith, & Wise, 1997) with follow-up data from the Longitudinal Twin Study of Reading Disability (LTSRD, Wadsworth, DeFries, Olson, & Willcutt, 2007). Chapter 4 examines the heritability and stability of reading disability employing a bivariate extension of the DF model and further explores data from twins and their nontwin siblings with regard to special twin environment. Chapter 5 will discuss the differential etiology of stability of reading disability as a function of gender. Finally, chapter 6 summarizes and discusses the results from these studies and considers how they may inform further research in these areas.

1.1 The importance of studying reading disability

The most important activity for children in their beginning school years and the foundation for their later academic achievement is learning to read (Chall, 1983). Early reading development is a strong predictor of later reading development and is fundamental for later academic success (Juel, 1988; Stanovich, 1986). The question of why some children have difficulty learning to read has been the focus of a plethora of research over the past several decades and much has

been learned about the probable and improbable causes of reading disability. Of special interest in the study of reading disability have been children who have at least average intelligence, who do not have general learning difficulties, and whose reading problems are not due to extraneous factors such as auditory or visual deficits, socioeconomic disadvantage, or lack of educational opportunities.

Reading problems in such children are manifested as extreme difficulties in acquiring basic reading sub-skills such as word identification and phonological decoding. Such difficulties have been estimated to occur in approximately 7% to 15% of school age children (Benton & Pearl, 1978; Harris & Sipay, 1990; Shastry, 2007; Shaywitz et al., 1992) and tend to be accompanied by specific deficits in cognitive abilities related to reading and other literacy skills. Deficits in phonological coding continue to characterize dyslexic readers throughout adolescence (Shaywitz, Fletcher, Holahan, Shneider, Marchione, Stuebing, Francis, Pugh, et al., 1999). Investigators in the field strongly assert that there is a deficiency within the phonological module, a specific component of the language system, which underlies dyslexia. Because of impairment within this system, subjects are unable to translate and identify words (Shaywitz, 1998).

Negative effects of reading disability extend well beyond childhood. For example, a 20-year follow-up of students diagnosed with dyslexia found that these individuals achieved a lower socioeconomic status (SES) than their parents, and a greater proportion of the students reported a diagnosis of a mental illness compared with their nondisabled peers (Raskind, Golberg, Higgins, & Herman, 1999). A review of the literature on long-term outcomes for students diagnosed with learning disabilities concluded that these students are less likely to receive or complete postsecondary education than their peers without disabilities and that employment rates are lower

for persons with learning disabilities than for those without a learning disability (Levine & Nourse, 1998). Older dyslexic students may be similar to their unimpaired peers on untimed measures of word recognition, but may continue to suffer from the phonologic deficits that make reading less automatic, more effortful, and slow (Shaywitz & Shaywitz, 2005).

In the 1970's a survey commissioned by the Right to Read Office reported substantial percentages of adults lacking "survival literacy". In other words, these adults had difficulty completing applications for employment due to reading difficulties. Further, essential forms such as driver's license, Medicaid insurance forms, or various technical manuals were unable to be processed appropriately by adults without good literacy skills (Chall, 1983).

Although it's important to note concerns in reading disability for adults, there is evidence for low literacy among high school students showing a steady decline in SAT verbal scores. Freshman reading and vocabulary scores were found to be significantly lower than results on tests originally administered 50 years prior (Chall, 1983). Presently SAT scores are lower than they were in the past five years. More recently, results examined by the National Assessment of Education Progress (NAEP) found that high school reading test scores were significantly lower than the prior decade (Coulson, 1996). One hypothesis for these trends is that education for reading is better for early grades than for later grades. In addition, that there is a better understanding of the processes involved for early reading skills than for later reading skills (Resnick & Weaver, 1979).

In light of a downward trend in reading literacy, the current educational policy climate emphasizes the need to bring "evidence-based progress" to reading instruction (Stanovich & Stanovich, 2003). The highly controversial No Child Left Behind 2001 federal legislation mandated that school-based professionals adopt scientifically based research to equalize reading

disparities among students. One such program, Response to Intervention (RTI) seeks to redefine how reading disabilities are identified and addressed within the public school system. Designed as a prevention model, RTI features multiple tiers of reading interventions designed to support students from the earliest stages of reading development (prekindergarten and kindergarten) following each subject's progress with interventions that are carefully monitored to ensure progress towards specific benchmarks (Vellutino, Scanlon, & Jaccard, 2003). In a 5-year longitudinal study, Vellutino, Scanlon, Small, & Fanuele (2006) assessed 1,373 children at the beginning of kindergarten. Based on a letter-knowledge test, approximately 30% of the subjects were identified as being at risk for early reading difficulties. As a means to further document their status, the subjects were given additional tests to examine phonological awareness; rapid automatized naming of objects, counting by ones, and number identification. Children assigned to the treatment group were provided with small-group early literacy interventions consisting of meeting with a certified teacher twice each week for 30 minute sessions where they focused on various reading tasks. When compared to a school-based comparison group, children assigned to the treatment group performed significantly better overall and were better prepared for First-Grade reading.

In another study by Torgensen, Alexander, Wagner, Rashotte, Voeller, & Conway (2001) sixty reading-disabled children received intense intervention consisting of two 50-minutes periods of reading remediation each day for eight weeks. Notable improvements in generalized reading skills were found in addition to remaining stable for approximately two-years post intervention.

Finally, in order to promote successful reading skills, acknowledging a students' interest in reading can have a profound effect on their learning and motivation (Cole, Teti, & Zahn-

Waxler, 2003, Coles & Hall., 2002) and should be of important consideration when focusing on overall reading improvement for future remediation programs.

1.2 History of reading disability and historical definitions

1.2.1 History of reading disability

First recognized in the late 1800's, reading disability was initially termed congenital word-blindness and was identified during this period as "an inability to learn to read easily which was developmental in nature, bore some of the characteristics of acquired reading disorders, and was distinct from developmental disorders of reading which were a result of or related to low intelligence, physical disability, inadequate schooling, or other exogenous factors" (Finucci, 1978). The first account of this disability was noted in the British Medical Journal by Morgan (1896), who published a case of a 14 year old boy who could not learn to read, but had no difficulty learning math. The young man did not recognize written or printed words (Sinclair, 1948).

Less than a decade later, British ophthalmologists and neurologists reported several cases of congenital word-blindness, although during this same time period the condition was also being brought to the attention of school authorities (Finucci, 1978). By 1904 a directive to head teachers of schools under the London County Council resulted in nearly 100 cases of the condition being noted in school records, and it was estimated that 1 in 2000 of all London Elementary School children may be expected to show word-blindness to a considerable degree (Thomas, 1905).

It was around this time that many practitioners believed that there was a hereditary component. For example, Thomas (1905) noted that congenital word blindness "frequently assumes a family type; there are a number of instances of more than one member of the family

being affected, and the mother often volunteers the statement that she herself was unable to learn to read, although she had every opportunity" (p. 381). Warburg (1911) found dyslexia markedly present in the many cases observed and suggested that dyslexia was transmitted by non-affected mothers. Drew (1956) summarized early literature of "heritable dyslexia" and found little reliable evidence for neurological correlates; however, previous studies did indicate a strong familial component for reading disabilities. It was further suggested that the lack of consistent results in prior studies was likely due to difficulties with reliable definitions for "congenital word-blindness".

1.2.2 Historical definitions

Most definitions of dyslexia have emphasized the importance of exclusionary criteria, meaning that in order to receive a diagnosis of developmental reading disability, the observed difficulties in attaining reading skills must not be a secondary symptom of a more primary deficit or problem (e.g., mental retardation, brain damage, sensory problems, deprivation, or low socio-economic status). For example, Kirk & Bateman (1962) characterized reading disability as "a retardation disorder, or delayed development in reading resulting from a psychological handicap caused by a possible cerebral dysfunction and/or emotional or behavioral disturbances. It is not the result of mental retardation, sensory deprivation, or cultural or instructional factors". Critchley (1970) defined developmental dyslexia as "a disorder manifested by difficulty in learning to read despite conventional instruction, adequate intelligence and socio-cultural opportunity".

More recently the Diagnostic and Statistical Manual of Mental Disorders, fourth edition (DSM-IV; American Psychiatric Association [APA], 1994) applied the following diagnostic criteria to dyslexia: "(A) Reading achievement, as measured by individually administered

standardized tests of reading accuracy or comprehension, substantially below that expected given the person's chronological age, measured intelligence, and age-appropriate education; (B) The disturbance in Criterion A significantly interferes with academic achievement or activities of daily living that require reading skills; and (C) If a sensory deficit is present, the reading difficulties are in excess of those usually associated with it (p.50)".

However, critics have suggested that emphasis on exclusionary factors does not allow the formulation of theoretical concepts of the underlying symptoms and characteristics of reading disability and further implies that it may not be accurately diagnosed in a child from a disadvantaged or unconventional background (Lyon, 1995; Lyon & Chhabra, 1996), and does not predict response to intervention. In fact, difficulties in the diagnosis of reading disability have been recognized for more than a century. Nettleship (1901) recognized the need to differentiate between generally "backward" children and those "whose only principal difficulty is real inability to learn to read". Nettleship (1901) also noted that there were substantial difficulties in the diagnosis of word-blindness in children from disadvantageous environments.

Early definitions of reading disability often included unexpected underachievement characterized as a discrepancy between achievement and intellectual ability, despite adequate opportunity to learn and in the absence of sensory impairment or cultural disadvantages. This discrepancy is typically defined in terms of a difference between IQ scores and scores on a test of reading achievement. For example, the World Federation of Neurology defined specific developmental dyslexia as "a disorder manifested by difficulty in learning to read, despite conventional instruction, adequate intelligence, and socio-cultural opportunity" (Critchley, 1970, p.11). Simply by omission, definitions of this nature make certain that learning disabilities are not attributable to such factors as mental retardation (Lyon, 1996; Stanovich, 1986). Rutter and

Yule (1975) distinguished between "reading backwardness," which refers to age-discrepant reading achievement regardless of general intelligence, and "specific reading retardation," which suggests that reading achievement is below what might be expected based on general cognitive ability. Their results suggest that although IQ scores were normally distributed, reading achievement scores did not show the same normal distribution as there was a "hump" indicating the presence of a greater proportion of low reading achievement scores than expected by chance. In addition, recent findings suggest that the use of IQ-discrepancy for identification of reading disabilities is potentially harmful to students as the criteria results in a wait and see attitude (Fletcher, Coulter, Reschly, & Vaughn, (2004) which leads to a delay in suitable interventions. In as much as previous definitions for RD describe the lack of reading ability, it's important to also define reading disabilities by identifying what it is *not*. For example, a definition of reading disabilities should include persistent literacy learning difficulties in otherwise normally developing children and may want to exclude factors such as conditions that began prior to schooling (i.e., severe attentional problems, oral language impairment, emotional and/or behavioral problems, or chronically poor health (Tunmer & Greaney, 2010).

1.2.3 Current definition of reading disability

The present definition of dyslexia is characterized as an unexpected difficulty in reading in individuals who otherwise possess the intelligence and motivation considered necessary for fluent reading, and who also have had reasonable reading instruction. Thus, the current working definition of dyslexia, adopted by both the International Dyslexia Association (IDA) and the National Institute of Child Health and Human Development (NICHD), is as follows:

"Dyslexia is a specific learning disability that is neurobiological in origin. It is characterized by difficulties with accurate and/or fluent word recognition and by poor spelling and decoding

abilities. These difficulties typically result from a deficit in the phonological component of language that is often unexpected in relation to other cognitive abilities and the provision of effective classroom instruction. Secondary consequences may include problems in reading comprehension and reduced reading experience that can impede growth of vocabulary and background knowledge" (Lyon et al., 2003).

In particular, there has been extensive converging evidence relevant to the epidemiology, developmental course, neurobiology, and the cognitive and linguistic characteristics of dyslexia which has accrued since the working definition of dyslexia from 1995. Moreover, the understanding of dyslexia has been informed by a number of studies (see Lyon, 2003) that now provide the opportunity to integrate information about the nature and magnitude of developmental dyslexia.

1.3 Family and Twin studies

1.3.1 Family studies

The familial nature of reading disability was noted early (1896-1917) and a hereditary basis for its occurrence was hypothesized (Finucci, 1978). In 1905, two separate accounts were published of more than one case of reading problems occurring with a family. Thomas (1905), in an initial description of "congenital word-blindness" suggested the likelihood of a familial tendency for the disorder, as the mother of an affected child would indicate that she too had been unable to learn to read. The second description, by Fisher (1905), portrayed a girl whose maternal uncle had a similar problem: "I think we may fairly take this as evidence of a family tendency to imperfection in development of the visual memory centre for words in the cortex of the left angular gyrus (postulated on the basis of an analogy to acquired word-blindness in adults)… such hereditary tendency seems not improbable (p. 316)". Following these initial

reports, doctors began to observe and record other cases of dyslexia occurring within a family (e.g., Hinshelwood, 1907; Stephenson, 1907), and it was suggested by Hinshelwood (1911) that evidence for the hereditary tendency of word-blindness would rapidly increase if observers of cases made careful inquires into the family history of both the present and previous generations. Since these initial reports, results from numerous kinship studies have clearly demonstrated the familial nature of reading disability.

In Hallgren's (1950) classic family study on reading disability, 88% of 112 probands (children diagnosed as affected with reading problems) had one or more relatives who were also affected. Hallgren investigated 276 cases, 112 affected children, and 160 secondary cases (siblings and parents of the affected). Results from Hallgren's study were considered to be the most thorough and provided the first definitive evidence of the familial nature of reading disabilities (Melekian, 1990).

Hallgren hypothesized that reading disability was the result of an autosomal dominant gene, or a single copy of a specific allele that is sufficient to cause the disorder. However, there are issues with the interpretation of Hallgren's family data (DeFries & Alarcón, 1996). A fully penetrant autosomal dominant gene would result in at least one parent being affected. However, in Hallgren's study, both parents were unaffected in 17% of the families examined. Another issue was the parents' disorders were diagnosed primarily based on interview data. Lastly, reading disability may have been over-represented because diagnoses were based in part on family history rather than reading tests, potentially resulting in ascertainment bias.

Finucci, Guthrie, Childs, Abbey, & Childs. (1976) reported results of the first family study of reading disability in which the probands and their relatives underwent psychometric testing. In a small sample of 20 probands (15 males and 5 females), 34 out of 75 first-degree

relatives were diagnosed as having reading disability. In the 16 families of probands in which both parents were tested, 13 out of 16 had one or both parents affected. This is similar to that previously reported by Hallgren (1950). Further examination of the pedigrees indicated that reading disability aggregates in families. In addition, given that both very good and very poor readers occurred within a single family was suggestive of genetic influences rather than a result of familial environment (Finucci et al., 1976). However, because the patterns of inheritance varied in the extended pedigrees included in this study, it was concluded that reading disability is genetically heterogeneous.

The Colorado Family Reading Study (CFRS; DeFries, Singer, Foch, & Lewitter, 1978; Foch, DeFries, McClearn, & Singer, 1977) was initiated in 1973 to assess the etiology of reading deficits. A primary objective of this study was to use data from probands, parents, and siblings to test alternative models of familial transmission (DeFries et al., 1997). Between October 1, 1973 and July 30, 1976, the CFRS administered an extensive psychometric test battery to 125 probands, their parents and siblings, and members of 125 matched control families. Results indicated that the difference between the average scores of probands and controls was significant ($p < .05$) for three principal component scores (symbol-processing speed, spatial reasoning, and a composite measure of reading comprised of three subtests of the PIAT: Reading Recognition, Reading Comprehension, and Spelling). In addition, siblings and parents of probands displayed substantial and highly significant deficits on measures of both reading performance and symbol-processing speed, conclusively demonstrating the familial nature of reading disability.

In a later study, Vogler, DeFries, and Decker (1985) measured the familial risk for dyslexia in the Colorado Family Reading Study (CFRS) sample. Data were obtained from a referred sample of reading-disabled children and their parents. Parents completed a questionnaire

asking them to describe their own reading habits, including whether or not they had had any serious difficulty in learning to read. Results suggested a significantly higher probability of a parent being reading disabled if a child was also reading disabled (x^2 = 44.42, p < .001 for fathers; x^2 = 22.87, p <.001 for mothers). Findings from this family study indicated that there was an increased risk for a child to develop reading difficulties if his or her parent also reported difficulty with reading. In addition, results found that the sex of a parent reporting reading difficulties was not a factor in the increased risk to a child. These risk estimates were somewhat lower than that of Hallgren's (1950) study, but nonetheless were substantially increased and demonstrated familiality for reading disability.

More recently, a study by Pennington & Lefly (2001) examined middle- to upper-middle-class preschool children of high family risk (N = 67) and low family risk (N = 57) for reading disabilities. Families came from two sources, volunteers from Denver area preschools and volunteer families with a history of RD, also from the Denver area. The goal of this study was to use the "natural experiment" provided by familial transmission of RD to obtain answers about early reading development prior to diagnosis. Whereas the design of the study did not allow for separation of genetic and environmental influences, demographic characteristics of the sample suggested that literacy concerns found in the high family risk group were not likely due to environmental factors and more likely due to genetic contribution.

1.3.2 Twin Studies

Twins in a sense are their own unique science experiment, because they offer distinctive and informative ways to study disorders such as reading disability. Examining correlations among family members is useful for initially establishing the familial nature of a behavior but will not distinguish between shared family environmental and shared genetic influences. As a

result, familial resemblance is necessary, but is not sufficient evidence for genetic etiology (DeFries, 1985). Other strategies must be implemented to partition the effects of genetics and the shared environment. Utilizing the twin model to assess the etiology of a behavioral trait is advantageous because the effects of genetic and environmental sources of family resemblance can be estimated separately. Twin studies compare the resemblance of pairs of identical (monozygotic, MZ) twins, who are genetically the same, to the resemblance within pairs of fraternal (dizygotic, DZ) twins, who share half their alleles identical-by-descent (IBD) for segregating genes.

Francis Galton's 1875 study of "The history of twins" marks the beginning of the long and contentious use of twins to test the relative influence of heredity and environment. The first real twin study in which identical and fraternal twins were compared was conducted in 1924 in an attempt to estimate genetic influences on intelligence (Merriman, 1924). Early twin studies used a comparison of MZ and DZ twin concordance rates to assess the extent to which behavioral disorders are due at least in part to genetic influences (DeFries & Light, 1996).

The concept of concordance rates to examine behavioral disorders is relatively straightforward. However, calculated rates of concordance depend on the manner in which the sample was obtained (DeFries & Alarcón, 1996). For example, both members of a twin pair may be affected by the disorder in question (+ +) or only one member of the pair may be affected resulting in two types of discordant pairs (+ - and - +). Twin pairs may be examined by two different forms of selection; single selection where only one member of the twin pair would be ascertained as a proband such as Twin 1 (Thompson & Thompson, 1986), or truncate selection, in which both members of a pair could be selected as probands. As shown in Table 1.1, if we included only the first-born twin (affected) then only the first two rows of twins would be

selected for analysis and the sample result would be C + D and the concordance rate would be calculated as C/(C + D). However, when calculating the pairwise concordance rate using truncate selection (where both members of the twin pair could be tested) the third row from Table 1.1 would be included in the analysis (C + 2D). To adjust for the increased number of discordant pairs in the sample, each concordant pair would be counted twice, first when Twin 1 is the proband and second, when Twin 2 is the proband. Double entry of concordant twin pairs increases the sample to 2C + 2D. The resulting proband-wise concordance rate is equal to that of pairwise concordance with single selection; 2C/ (2C + 2D) = C/ (C + D).

Table 1.1

Concordance of Twin Pairs

Pairs	Twin 1	Twin 2	Number
Concordant	+	+	C
Discordant	+	-	D
Discordant	-	+	D

(after DeFries & Alarcón, 1996)

In a literature review of congenital word-blindness, Zerbin-Rudin (1967) included data from six single-case studies of concordant twin pairs (5 MZ and 1 DZ), 39 pairs from Norrie's (1954) twin study (9 MZ and 30 DZ), and six pairs (3 MZ and 3 DZ) of twins from Hallgren's 1950 family study. The probandwise concordance rates from 17 MZ and 34 DZ twin pairs in this combined sample were 100% and 52% respectively, suggesting that reading difficulties are highly heritable. Similarly, in a sample of 31 MZ and 31 DZ twin pairs obtained through mothers-of-twins clubs, in which at least one member of each pair was classified as affected

based on reading history information (Bakwin, 1973), probandwise concordance rates were 91% for MZ pairs and 45% for DZ pairs (see Table 1.2).

Table 1.2

Probandwise Concordance Rates for Reading Disability

Study	Number of Pairs		Proband-wise Concordance (%)	
	MZ	DZ	MZ	DZ
Zerbin-Rudin (1967)	17	34	100	52
Bakwin (1973)	31	31	91	45
Stevenson et al. (1987)	14-19	27-42	33-59	29-54

(after DeFries & Alarcón, 1996)

One study failed to demonstrate significantly higher concordance rates for identical twins. Stevenson, Graham, Fredman, and McLoughlin (1984, 1987) conducted the first twin study of reading disability in which children were administered standardized tests of intelligence, reading, and spelling. The investigators assessed the genetic contribution to reading disability in 13-year-old MZ and DZ twins ascertained by examining hospital records in five London areas or through primary schools in the London area. Twins were administered the Schonell Graded Word Reading and Spelling Tests and the Neale Analysis of Reading Ability to diagnose reading or spelling 'backwardness' (reading or spelling age below chronological age), or reading or spelling 'retardation' (reading or spelling achievement below what would be predicted form IQ and chronological age). Probandwise concordance rates for various diagnostic criteria were relatively low for both zygosity groups, ranging from 33 to 59% for identical twins and from 29 to 54% for fraternal twins. The investigators (Stevenson et al., 1987) concluded that genetic factors play only a moderate role in general reading 'backwardness' and specific reading

'retardation', and that the genetic etiology for reading disability in children 13 years of age may be less important than in younger children. It's important to consider that there is substantial variation across these studies and the studies are relatively small, however, the concordance rates observed in MZ twin pairs are consistently higher than corresponding rates for DZ twin pairs, in two of the three studies and provide support for partial genetic etiology of reading disability.

Prior to 1982, there were very few twin studies of reading disability. Therefore, a twin study was initiated as part of the Colorado Reading Project (CRP; Decker &Vandenberg, 1985; DeFries, 1985) supported by the National Institute of Child Health and Human Development (NICHD). In this ongoing study, currently supported by the NICHD as the Colorado Learning Disabilities Research Center, MZ and DZ twin pairs in which at least one member of each pair exhibited a positive school history of reading difficulties, and a comparison group of twins with a negative school history, are administered an extensive battery of psychometric tests, which will be discussed at length in chapter 2 of this proposal. In 1995, there were 186 MZ twin pairs and 138 DZ same-sex twin pairs who met criteria for inclusion in the proband sample. The probandwise concordance rates for MZ and DZ twin pairs were estimated at .64 and .35, and significantly different ($p < .001$), providing further evidence that reading disability is due at least in part to heritable influences (DeFries & Alarcón, 1996).

1.3.3 Twin and Sibling studies

In an early study by Zieleniewski, Fulker, DeFries, & LaBuda. (1987), data from identical twins, fraternal twins and nontwin sibling pairs, in which one member of each pair was selected on the basis of low reading scores, were analyzed. Results indicated that tests for special twin environment were nonsignificant, however, Zieleniewski (1987) suggested that inclusion of data from "weaker relationships" (e.g., siblings versus twins) in analyses of the genetic and

environmental etiologies of individual differences may result in somewhat higher standard errors of some parameter estimates.

A more recent study employing data from the Twins Early Development Study (TEDS) suggests that estimates for shared environment were more than twice as large for twins as compared to nontwin siblings (Koeppen-Schomerus, Spinath, & Plomin, 2003). A possibility for this discrepancy may be that twins, who are necessarily the same age, may share more common environmental factors than different age siblings. Although they suggest that the same-age hypothesis does not explain all of the differences between results for twins and siblings. For example, as twins become adults and share less of a family environment you might expect a reduction in the special twin environment; however, this is not always the case. In fact correlations were slightly greater for DZ twins adopted-apart (r_{DZA} = .32 to .34) compared to DZ twins reared together (r_{DZT} = .22). Although the samples sizes were not large enough to provide adequate power to detect differences (Bouchard, Lykken, McGue, Segal, & Tellegen, 1990; Pedersen, Plomin, Nesselroade, & McClern, 1992). In addition, across studies, DZ twin correlations tend to be greater than nontwin sibling correlations, indicating that twin estimates of shared environment are to some extent specific to twins suggesting a "special twin environment" (Medland, Wright, Geffen, Hay, Levy, Martin, & Duffy, 2003; Young, Rhee, Stallings, Corley, & Hewitt, 2006).

1.4 Gender Studies

Sex differences in reading disability have been an ongoing research subject for decades and results are as varied presently as they have been in previous research. Multiple studies have compared reading difficulties for boys and girls examining both mean differences and prevalence rates (DeFries, 1989; Hawke, Wadsworth, & DeFries, 2006; Hawke, Olson, Willcutt,

Wadsworth, & DeFries, 2009; Shaywitz, Shaywitz, Fletcher, & Escobar, 1990; Stevenson 1992). Differences in sex ratios for clinical and referred samples versus research identified samples vary widely. For example, sex ratios for clinical and referred samples range from 2:1 to 15:1 males to females (Finucci & Childs 1981; Vogel, 1990). In contrast, research identified sex ratios are closer to 1:1 (Shaywitz et al. 1990; Stevenson 1992; Wadsworth, DeFries, Stevenson, Gilger & Pennington 1992. Results from Shaywitz et al. (1990) indicate that samples obtained by referral contain an excess of males (sex ratios of 4.25 and 2.38 in second and third grades respectively). In contrast, research identified studies find sex ratios reduced (1.26 and 1.5 respectively). Although ratios were in the direction of an excess of males, the differences for male and females were not significant. Further, investigation of data from reading disabled children and their family members have not been able to replicate an increase in male/female sex ratios in children with disabled mothers (Wadsworth et al., 1992). Differences in the etiology of reading disabilities are suggested when there are sex ratios greater than one. However, it is important to consider that factors which contribute to the etiology of reading disability may be a mix of genetic and environmental influences. For example, if one sex is exposed to low teacher expectations or experiences from parent-child interactions. It may be that the same genetic factors operate in males and females and the observed prevalence differences may be environmentally mediated (Stevenson, 1992).

A study by Wadsworth, Knopik & DeFries (2000) examined the differential etiology of reading disabilities in boys and girls from a research identified sample in the Colorado Reading Project (DeFries, 1985). The ratio of male to female probands was 1.13:1 ($p \geq .07$). A test of differential genetic etiology of reading disability as a function of sex was non-significant ($p \geq .90$). In a more recent study by Hawke et al. (2006), a larger sample from the CRP and CLDRC

examined reading performance data from 264 MZ pairs (129 male pairs and 135 female pairs) and 214 same-sex DZ pairs (121 male pairs and 93 female pairs). The ratio of male to female probands was 1.12:1, a non-significant difference (p = .16). When data were fitted to multiple regression analyses, a test for differential genetic etiology was also not significant ($p \geq .35$).

A study by St. Stauver, Katusic, Barbaresi, Colligan, & Jacobsen (2001) conducted a case-control study to examine whether risk factors (i.e., low birth weight, gender, pre-term birth, parental age) for reading disability varied as a function of gender. Three hundred three reading disabled cases were identified using IQ and achievement test scores collected from school and medical records. Controls consisted of all children who were not identified as reading disabled (n = 4529). Their results suggest that males and females are differentially susceptible to risk factors for RD. For example, males had a 2.5-fold increase of RD compared to females. Additionally, socio-demographic factors, such as paternal age and parental education impacted males and females differentially. For example, increased paternal age decreased the risk for RD in girls, but not for boys. Further, low maternal education level increased the risk of RD for girls, but not for boys. However, for paternal education level, the effect was opposite with an increased risk of RD for boys and not for girls. These differences suggest a differential etiology as a function of gender as evidenced by differential risk for RD. However, the additional criteria to qualify case subjects for this study which included obvious signs of school difficulty based on: 1) grade retention, 2) presence of a special education learning plan, and 3) other supplemental instruction, are not well defined. Additionally, information such as notations made by teachers or comments from parents stating concerns' regarding a child's learning performance was included in the criteria for RD allowing for subjective interpretation of diagnosis and results.

An additional study by Rutter, Caspi, Fergusson, Horwood, Goodman, Maughan, Moffitt, Meltzer, & Carroll (2004) examined results from four previous epidemiological studies evaluating referral bias for reading disability in boys. Results of all four studies suggested that rates of reading disability were significantly higher in boys. Each study, the Dunedin Multidisciplinary Health and Development Study (Silva, 1990), the Christchurch Health and Development Study (Fergusson, Horwood, Shannon, & Lawton, 1989), the Office of National Statistics (Meltzer, Gatward, Goodman, & Ford, 2000), and the Environmental Risk Longitudinal Twin Study (Moffitt, 2002) utilized different criteria for ascertaining their subjects. In addition, subjects were not defined by mean differences, but overall prevalence rates in their results. For example, the non-IQ referenced criteria for reading disability was defined as reading performance 28 months behind population norms on either reading accuracy or reading comprehension. The male-female difference on group tests was 15.9% and 7.2% respectively with males scoring higher. However, it is not indicated if there was a significant mean difference in male-female test scores.

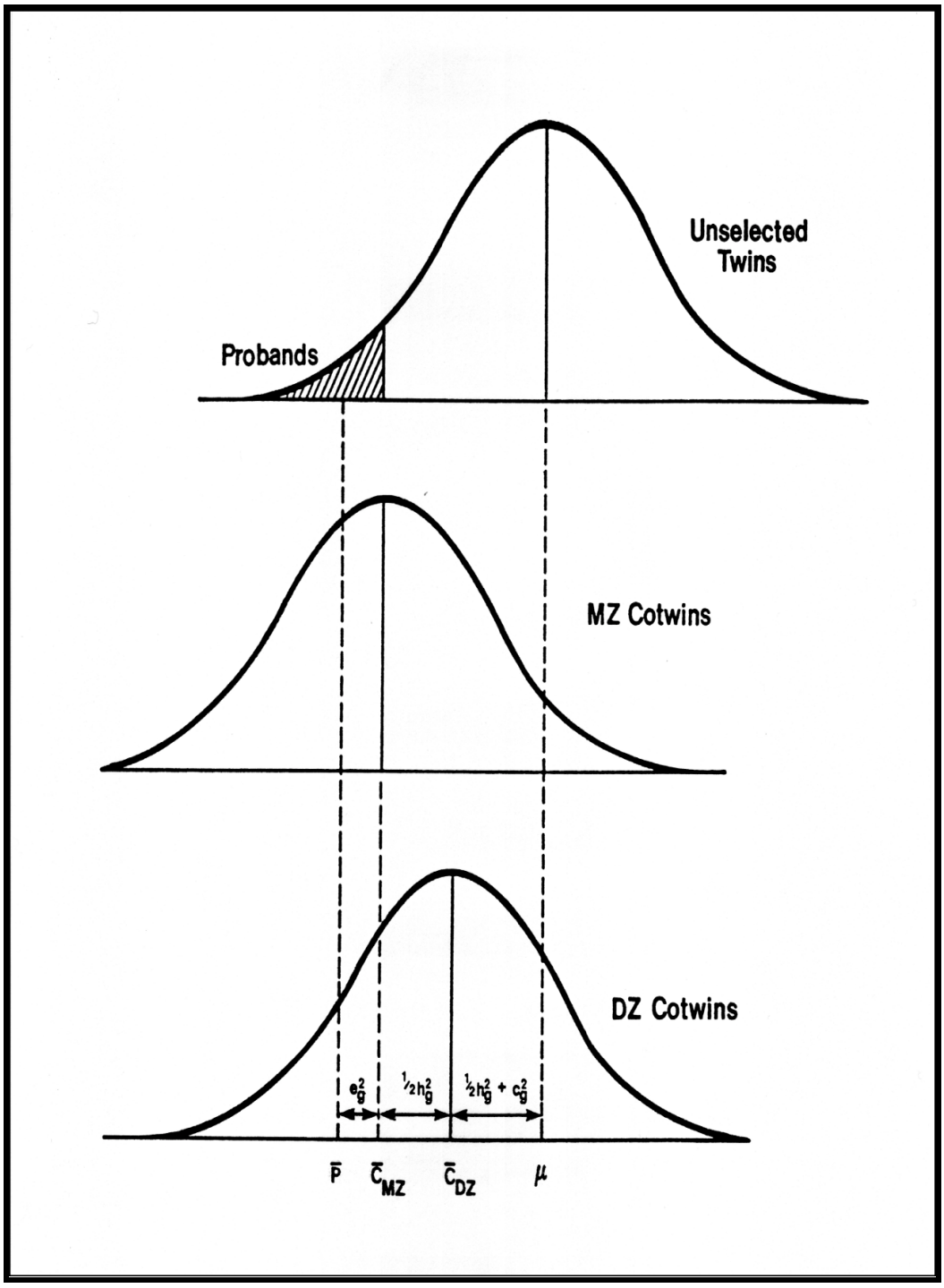
Therefore, males may be over represented in these studies and other similar studies, but these results are unclear as to the genetic and environmental influences contributing to reading disability among males and females.

1.5 DF analysis

When comparing the similarity of monozygotic (MZ) twins, who share all their genes, to dizygotic (DZ) twins, who share half of their segregating genes on average, results provide estimates of the extent to which a disorder is due to genetic or environmental influences. The most straightforward test for genetic influences on a disorder compares the rate of concordance in pairs of MZ versus DZ twins. A comparison of concordance rates for dichotomous variables

(e.g., reading disability versus normal reading performance) in MZ and DZ twin pairs can provide *prima facie* evidence for a genetic etiology (Wadsworth et al., 2000). However, children with reading disability are often diagnosed, at least in part, on the basis of quantitative measures with arbitrary cut-off points (i.e., extreme scores). Transformation of a quantitative measure of reading performance into a categorical variable results in a loss of important information about the continuum of reading variation (DeFries et al., 1997). Therefore, DeFries and Fulker (1985, 1988) proposed that a differential regression of MZ and DZ co-twin means toward the mean of the unselected population provides a more appropriate test of genetic etiology. Figure 1.1 illustrates hypothetical distributions of reading performance data from an unselected population of twins (with a mean μ) and from the MZ and DZ co-twins of the probands selected for extreme scores (DeFries & Fulker, 1988). Proband and co-twin means are symbolized as $\overline{P}$ and $\overline{C}$, respectively. The deficit of probands ($\overline{P} - \mu$) is due to heritable factors (h^2_g) and to environmental influences that are either shared (c^2_g) or not shared (e^2_g) by members of a twin pair (DeFries & Fulker, 1988). When MZ and DZ probands are ascertained because of extreme scores on a continuous measure, the scores of their co-twins are expected to regress toward the mean of the unselected population (μ). To the extent that the deviant scores exhibited by the probands are heritable, this regression toward the population mean should differ for MZ and DZ co-twins if the proband condition is heritable. Because members of MZ twin pairs are genetically identical, whereas members of DZ pairs share only about half of their segregating genes on average, scores of DZ co-twins should regress more toward the mean than those of MZ co-twins if reading disability is due, at least in part, to genetic influences (DeFries & Fulker, 1985, 1988).

Figure 1.1 Regression to the mean of the unselected population



(From DeFries, Fulker and Labuda, 1987)

Therefore, if MZ and DZ proband means are approximately equal, a simple t-test of the difference between MZ and DZ co-twin means ($\overline{C}_{MZ}$ and $\overline{C}_{DZ}$) can be used as a test of genetic etiology. Moreover, with suitable transformation of scores, the extent of differential regression by zygosity provides a direct estimate of the heritability of the extreme group deficit (h^2_g). Selected data are fitted to the following basic multiple regression model:

$$\hat{C}_x = B_1 P_x + B_2 R + A \quad [1]$$

where C is the predicted co-twin's score, P is the proband's score, R is the coefficient of relationship (coded 1.0 for MZ twin and 0.5 for DZ twins), and A represents the regression constant. The B_1 coefficient is the partial regression of the cotwin's score on the proband's score, and a measure of the average MZ and DZ twin resemblance (DeFries & Fulker, 1985; 1988). The B_2 coefficient is the partial regression of the cotwin's score on the coefficient of relationship and equals twice the difference between the MZ and DZ cotwin means after covariance adjustment for any difference between the MZ and DZ proband means. Additionally, when the interaction between the proband's score and the coefficient of relationship is added to the regression equation during a second step in the analysis of these data, direct estimates of heritability (h^2) and the proportion of variance due to common or shared environmental influences (c^2) relevant to the unselected population are also obtained. A more in depth explanation of the basic DeFries-Fulker method is included in chapter 2.

In order to examine twin and sibling data and test for a "special twin environment", a measure of the extent to which shared environmental influences for members of twin pairs and their non-twin siblings differ, a novel extension of the basic DF method (DeFries & Fulker, 1985; 1988) was proposed. The extended model is as follows:

$$\hat{C}_x = B_1 P_x + B_2 R + B_3 S + A \quad [2]$$

where *C* is now the co-twin's or co-sib's predicted score, *P* is the proband's score, *R* is the coefficient of relationship (1.0 for MZ pairs and .5 for both DZ pairs and twin/sib pairs), and *S* is a dummy code for pair type (+.5 for MZ twins, +.5 for DZ twins and -.5 for twin-sib pairs). When this model is fitted to the data and all three partial regression coefficients are estimated simultaneously, B_3 equals the difference between the DZ cotwin and co-sib means and provides a direct test of significance for the difference between environmental influences shared by DZ twin pairs versus those shared by twin-sib pairs, i.e., a test for "special twin environments" (Astrom, Wadsworth, Olson, Willcutt, & DeFries, 2011). The extended DF model is more fully discussed in chapter 2.

1.6 Summary

The investigation into reading disability has been a major focus for researchers for decades, with early theories proposing that a basic deficit in visual processing was at the center of reading difficulties for affected subjects (Orton, 1925). Although the heritable nature of dyslexia has become apparent, the phenotypic definition remains challenging and the genetic basis is very complex (Fisher & DeFries, 2002). The use of genetically sensitive designs, such as twin studies, has lead to a more complete understanding of the etiology of learning abilities. Analytical techniques are utilized to assess proportions of genetic and environmental influences of normally distributed traits. Among these is DeFries-Fulker multiple regression analysis, a versatile and powerful method to investigate extreme scores on a continuous trait. As the definition of specific reading disabilities evolves, so will our ability to examine and understand aspects of this behavioral trait with a variety of methods. However, additional studies of reading performance data from families, twins, and siblings, will be needed to further our research and help us to obtain a better understanding of how genetic and environmental pathways influence

these traits and help to prescribe treatments and appropriate remediation for learning disabled readers.

Chapter 2

Genetic and environmental etiologies of reading difficulties: DeFries-Fulker analysis of reading performance data from twin pairs and their nontwin siblings

2.1 Introduction

The familial nature of reading difficulties has been known for many years. For example, over a century ago, Thomas (1905) noted that "congenital word-blindness frequently assumes a family type" and that the mother often described herself as being unable to learn to read despite sufficient opportunity. Later, in the first large-scale family study of reading disability, Hallgren, (1950) found that 88% of 116 probands (the clinically affected individuals through whom the families were ascertained) had one or more relatives who were also affected. Similar results were obtained from subsequent family studies, thereby confirming the familial nature of reading disability (DeFries, Vogler, & LaBuda, 1986; Finucci & Childs, 1983; Gilger, Pennington, & DeFries, 1991).

In the Colorado Family Reading Study (DeFries et al., 1978), 133 children with reading disabilities, their parents, and siblings were tested on measures of reading and cognitive processes. In addition, 125 control children without reading disabilities, their parents, and siblings were tested on the same measures. Several different genetic analyses of these data were performed over the years (DeFries et al., 1986), and results suggested that familial transmission of reading disability was due at least in part to genetic influences.

Because family members share both genes and environmental influences, family studies do not provide tests of the relative importance of genetic and shared environmental variation. In contrast, twin studies can yield estimates of genetic, shared-family and non-shared environmental influences. Early twin studies compared "concordance" rates in pairs of identical

and fraternal twins as a test for genetic etiology. A pair is considered concordant if both members are affected with the same disorder or discordant if only one member of the pair is affected. Because members of MZ twin pairs are genetically identical, while DZ pairs share, on average, only half of their segregating genes, MZ concordance is expected to be greater than DZ concordance if a condition is heritable. Such differences in MZ and DZ concordance rates were obtained in several early studies of reading disability. However, sample sizes were small and results varied widely (Bakwin, 1973; Stevenson et al., 1987; Zerbin-Rudin, 1967).

Although a comparison of concordance rates in MZ and DZ twin pairs provides evidence for a genetic etiology, reading disability is diagnosed in part on the basis of quantitative measures with arbitrary cut-off points (Stevenson, et al., 1987). Thus, when a continuous measure, such as reading performance, is transformed into a categorical variable (e.g., reading disabled versus non-reading disabled) information pertaining to the range of variation in reading performance is inevitably lost. Consequently, DeFries and Fulker (1985, 1988) proposed fitting a multiple regression model to data from selected twin pairs to assess genetic and environmental influences on deviant scores. This method accounts for variation of continuous variables (e.g., reading performance) and facilitates an analysis of the extent to which deviant scores are due to genetic, shared environmental, and non-shared environmental influences.

An early study from the Colorado Learning Disabilities Research Center (CLDRC; DeFries et al., 1997) assessed the genetic and environmental etiologies of reading disability in a sample of 191 MZ, 143 same-sex DZ, and 99 opposite-sex DZ twin pairs ascertained for reading difficulties. At least one member of each pair was classified as reading disabled based upon a discriminant function score and had a Verbal or Performance IQ of at least 90. A control sample of 170 MZ pairs, 110 DZ same-sex pairs, and 68 opposite-sex DZ pairs was also tested.

Concordance rates were .67 for MZ twins and .38 for DZ twins, a highly significant difference ($p < .001$). Although the MZ and same-sex DZ proband means were highly similar, the MZ co-twin mean regressed only 0.20 standard deviation units on average toward the control mean, whereas that of the DZ co-twins regressed 0.94. When the basic regression model (DeFries & Fulker, 1985, 1988) was fitted to the transformed data, h^2_g (an index of the extent to which reading deficits are due to genetic influences) was 0.56 ($p < .001$), suggesting that more than half of the average reading performance deficit of probands was due to heritable influences. Results obtained from subsequent analyses of data from twin pairs ascertained for reading difficulties have been highly similar (e.g., Harlaar, Dale, & Plomin, 2007; Harlaar, Spinath, Dale, & Plomin, 2007; Wadsworth, Olson, & DeFries, 2010).

2.1.1 Stability of reading disabilities

More recently, Astrom et al. (2011) investigated the etiology of stability of reading deficits using a novel extension of the DF method (DeFries & Fulker, 1985, 1988) which incorporates sibling data and facilitates a test for "special twin environments" (i.e., a measure of the extent to which shared environmental influences for members of twin pairs differ from those for nontwin-sibling pairs).

2.1.2 Inclusion of twins and nontwin sibling data

The sample included 33 MZ and 64 DZ twin pairs, and 44 of their nontwin siblings, who participated in the Longitudinal Twin Study of Reading Disability (LSTRD; Wadsworth, et al., 2007) approximately five years after their initial participation in the CLDRC. In order to incorporate sibling data, a simple extension of the basic DF model was employed in which the co-twin or co-sib score was predicted from the proband score and the coefficient of relationship (1.0 for MZ, 0.5 for DZ twins and siblings) and a dummy-coded variable to differentiate data

from twin pairs and twin-sibling pairs. The model was simultaneously fitted to transformed data from selected twins, their co-twins and co-sibs. MZ and DZ proband means were highly similar at both initial and follow-up assessments, suggesting that the proband deficits are highly stable. Results of fitting the basic DF model to twin data from the initial assessment yielded an h^2_g estimate of .67 (p = .004), again indicating that the proband deficit in this sample was due principally to genetic influences. When the extended DF model was fitted to both twin and sibling data, shared environmental influences for members of twin pairs were not significantly different from those for twin-sibling pairs; however the difference was not trivial (i.e., $c^2_{g(t)}$ - $c^2_{g(s)}$ = .14, p = .167) .

Other studies have also tested for special twin environments. For example, Koeppen-Schomerus, et al., (2003) investigated twin-specific effects for estimates of shared environment for measures of general cognitive ability ("g") using data from twins participating in the Twins Early Development Study (TEDS) and from their nontwin siblings. Data from 1800 MZ and 1800 same-sex DZ pairs, as well as 130 same-sex younger siblings collected at two and three years of age were subjected to structural equation modeling. Results indicated a significant effect of twin-specific environment which accounted for 20-31% of the variance. The overall effect of common environment (shared plus twin-specific) accounted for 60-70% of the variance depending on the measure, suggesting that in early childhood, twin-twin estimates of shared environmental influences on general cognitive ability may be nearly twice as large as estimates based on twin-sibling data.

Among other studies examining twin-specific environmental influences, significant effects have also been reported for substance use. Analyzing data from 3744 adolescents (645 MZ pairs, 702 DZ pairs, 429 biological sibling pairs, and 96 adoptive sibling pairs) 12 to 18 years of age

participating in the Center for Antisocial Drug Dependence (CADD), Young et al. (2006) found that special twin environmental factors were significant for tobacco use, tobacco problem use, and alcohol use.

However, in a study of 1162 twin pairs (570 MZ pairs, 370 same-sex DZ pairs, and 222 DZ opposite sex pairs) and 426 siblings of these twins participating in the CADD, Ehringer, Rhee, Young, Corley, & Hewitt et al. (2006) found little evidence for special twin environment when examining self-report data pertaining to six common adolescent psychopathologies (attention deficit/hyperactivity disorder, conduct disorder, oppositional defiant disorder, generalized anxiety disorder, separation anxiety disorder, and major depressive disorder). Results suggested that the proportion of variance in adolescent psychopathology due to special twin-environmental influences was modest, ranging from 0.00 to 0.16. Thus, in general, the impact of shared environmental influences on twins is similar to that for ordinary siblings for these measures in the CADD.

2.2. The Current Study

The CLDRC has data from sibling of approximately half of those MZ and DZ probands. Therefore, applying the novel application of the DF multiple regression method described by Astrom et al. (2011), we have included data from the full CLDRC twin sample and their nontwin siblings in the present analysis. Thus, the primary objectives of the present study were twofold: (1) to examine the etiology of reading disability using the full sample of CLDRC twin pairs; and (2) to estimate the heritability of reading deficits using a novel extension of the DF method (DeFries-Fulker, 1985, 1988) which incorporates sibling data and thereby facilitates a test for "special twin environments" (e.g., Koeppen-Schomerus et al., 2003; Medland et al., 2003; VanGrootheest, Cath, Beekman, & Boomsma, 2007; Young et al., 2006). Based on results

obtained from previous twin studies, we hypothesize that genetic influences on reading deficits are substantial. Further, we predict that the test for special twin environments will be significant in this much larger sample.

2.3 Methods

2.3.1 Participants and Measures

Subjects in the present study were tested in the ongoing Colorado Learning Disabilities Research Center (DeFries et al. 1997). Twin pairs were systematically identified through 27 different school districts within the state of Colorado. Initially, all twin pairs in a school were identified by school personnel regardless of reading status. Permission was obtained from parents to review the school records of the twins for any evidence of reading problems (i.e., sub-standard reading test scores, recommendation to resource rooms or therapists due poor reading, reports by teachers, school psychologists, and parents). If either member of the twin pair demonstrated a history of reading problems, both twins and their siblings were invited to participate in the study at the University of Colorado, Boulder, and at the University of Denver. The subjects were administered an extensive battery of psychometric tests which included the Wechsler Intelligence Scale for Children-Revised (WISC-R; Wechsler, 1974) or the Wechsler Adult Intelligence Scale-Revised (WAIS-R; Wechsler, 1981), Peabody Individual Achievement Test (PIAT; Dunn & Markwardt, 1970), and measures of reading and language processes and executive functions. Age standard scores from the Reading Recognition, Reading Comprehension, and Spelling subtests from the PIAT were used to compute a discriminant function score (DISCR) for each subject using discriminant weights based on data from an independent sample of 140 non-twin children with reading disabilities and 140 non-twin children without reading problems (DeFries, 1985). The discriminant function scores were then age

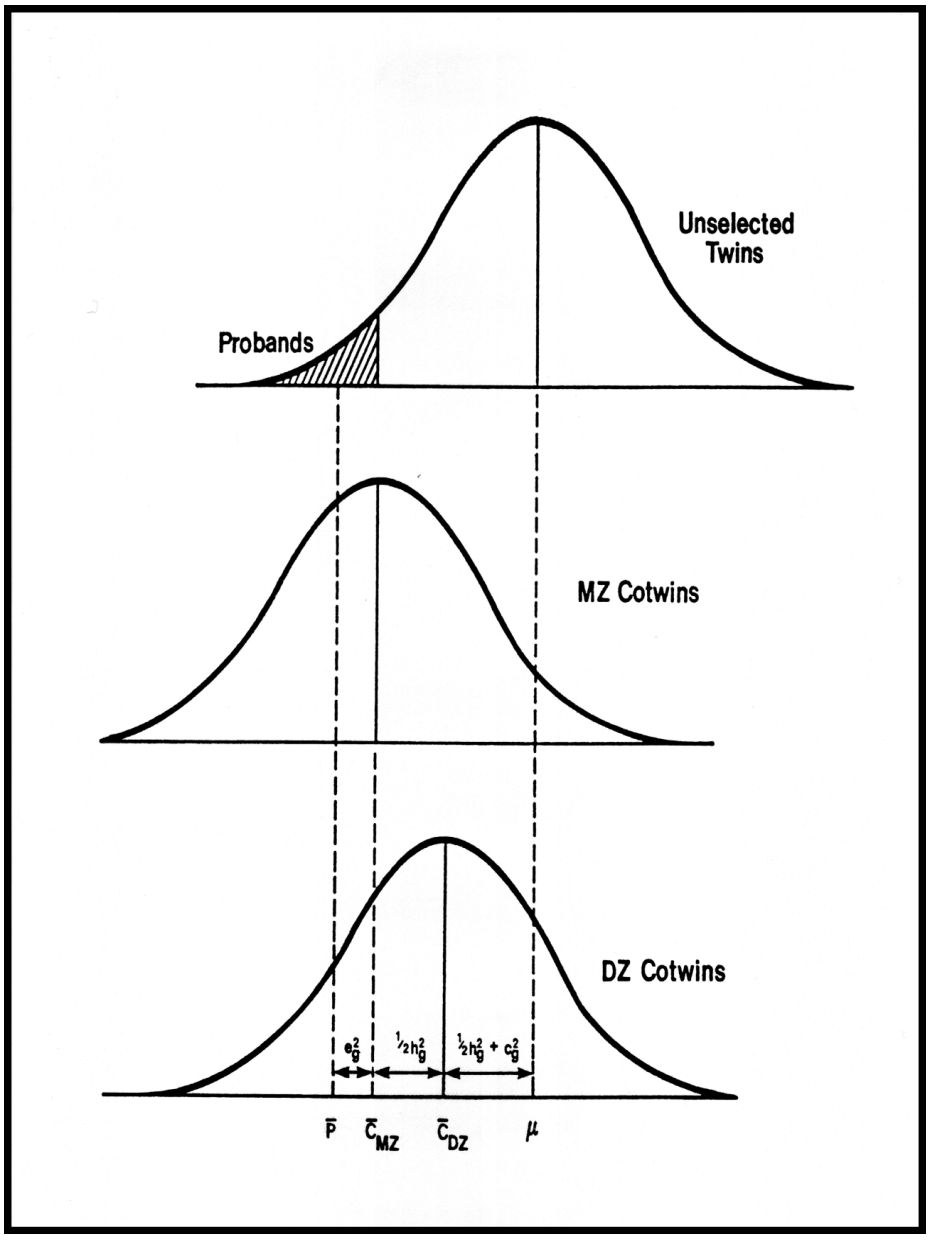
adjusted based on scores from the CLDRC control sample. Discriminant function score correlations with age were nonsignificant for both twins and siblings. In order for an individual to be included in the current proband sample, he or she must be classified as reading-disabled by the discriminant function score, have a positive history of reading problems in addition to a minimum IQ score of 85 on either the Verbal or Performance Scale of the WISC-R (Wechsler, 1974) or the WAIS-R (Wechsler, 1981), no evidence of neurological problems or serious emotional or behavioral or problems, and no uncorrected visual or auditory acuity deficit. 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 of reading problems. Zygosity of same-sex twin pairs was established using selected items from the Nichols and Bilbro (1966) questionnaire which has a reported accuracy of 95%. Where cases are undetermined from the questionnaire, zygosity is verified by analysis of blood or buccal samples.

The current sample included 254 MZ and 420 DZ twin pairs in which at least one twin met proband criteria at initial assessment in the CLDRC, and 303 of their non-twin siblings. For standardization and transformation of the variables, the control sample comprised 728 twin pairs. Informed consent and assent was obtained after characteristics of the testing session had been fully described and subjects had indicated that they understood. Informed consent was obtained from subjects 18 years of age and older and parental consent was obtained from the parents of children who were under 18. Assent was obtained from children under 18 years of age. The study protocol was approved by the Institutional Review Board of the University of Colorado, Boulder.

2.3.2 Data Analysis

When MZ and DZ probands have been ascertained because of extreme scores on a continuous measure, their co-twins are expected to regress toward the mean of the unselected population. However, to the extent that the scores of the probands are the result of heritable influences, this regression to the mean should differ for the MZ and DZ co-twins (see Figure 2.1). As MZ twins are genetically identical and DZ twins share only half of their segregating genes on average, the scores of DZ co-twins should regress more toward the mean of the unselected population if the condition is to some degree heritable. Consequently, when MZ and DZ proband means are approximately equal, a simple *t*-test of the difference between the MZ and DZ co-twin means provides a test of genetic etiology. However, the multiple regression analysis of such data facilitates a more flexible and statistically powerful test for genetic etiology (DeFries & Fulker, 1985, 1988).

Figure 2.1 Regression to the mean of the unselected population



(From DeFries, Fulker and Labuda, 1987)

2.3.2.1 Basic DF Model

The basic DF model is as follows:

$$\hat{C}_x = B_1 P_x + B_2 R + A \qquad [2.1]$$

where C is the predicted co-twin's score, P is the proband's score, R is the coefficient of relationship (R = 1.0 for MZ and 0.5 for DZ twin pairs), and A is the regression constant. When the basic model is fitted to selected twin data, B_1 provides a measure of the average MZ and DZ twin resemblance (DeFries & Fulker, 1985; 1988). Additionally, the B_2 coefficient estimates twice the difference between the means of MZ and DZ co-twins after covariance adjustment for any difference between the means of the MZ and DZ probands. Thus, B_2 provides a test for genetic etiology which is more general and statistically powerful than a comparison of concordance rates. Moreover, when the data are appropriately transformed prior to multiple-regression analysis (i.e., where each score is expressed as a deviation from the mean of the unselected population and then divided by the difference between the proband and population means), B_2 provides a direct estimate of heritability of the group deficit, h^2_g, an index of the extent to which the deficit of the probands is due to genetic factors.

To incorporate sibling data, an extension of the basic DF model can be simultaneously fitted to transformed data from selected twins, their co-twins and co-sibs. Presented in Table 2.1 are the expected transformed MZ and DZ co-twin means (see DeFries & Fulker, 1988) and the corresponding transformed co-sib means. As can be seen from these expected values, the difference between the DZ co-twin mean and the co-sib mean is a simple function of the difference between shared environmental influences in twins ($c^2_{g(t)}$) versus those in twin-sibling pairs ($c^2_{g(s)}$).

Table 2.1.

Expected transformed[1] co-twin and co-sib means

Subjects	Model
MZ Co-twins	$h^2_g + c^2_{g(t)}$
DZ Co-twins	$½\, h^2_g + c^2_{g(t)}$
Co-sibs	$½\, h^2_g + c^2_{g(s)}$

[1]Scores are expressed as a deviation from the unselected population mean and then divided by the difference between the proband and population means (after Astrom et al., 2011).

2.3.2.2 Extended Basic DF Model

Therefore, in order to test for differential c^2_g between twins and siblings, the following extended basic model can be simultaneously fitted to transformed data from probands, their co-twins and co-sibs:

$$\hat{C}_x = B_1P_x + B_2R + B_3S + A \qquad [2.2]$$

where C is now the co-twin's or co-sib's predicted score, P is the proband's score, R is the coefficient of relationship (1.0 for MZ pairs and now, 0.5 for both DZ pairs and twin/sib pairs), and S is a dummy code for pair type, i.e., twin pair versus twin-sibling pair (+.5 for MZ twins, +.5 for DZ twins and -.5 for twin-sib pairs). When this model is fitted to the data and all three partial regression coefficients are estimated simultaneously, B_3 estimates the difference between the DZ co-twin (CDZ) and co-sib (CS) means and, therefore, provides a direct test of significance for the difference between environmental influences shared by members of DZ twin pairs and those of twin-sib pairs. B_2 estimates h^2_g, derived only from the twin data as in the basic model.

Because B_3 estimates the difference between $c^2_{g(t)}$ and $c^2_{g(s)}$, its significance is relevant for obtaining an estimate of h^2_g based upon an analysis of the combined twin and co-sibling data. If B_3 is small and non-significant, S may be dropped from the extended model, and Equation 2.1 may be fitted to the combined data set of twins and siblings. In such cases, B_2 will estimate h^2_g from both the twin *and* co-sib data, and not only the twin data. Conversely, if B_3 is significant or relatively large, h^2_g should be estimated from fitting Equation 2 to the combined data set.

For the present study, data were analyzed from twin pairs in which at least one member of each pair was affected. Because truncate selection was employed (DeFries & Gillis, 1991), pairs in which both members met criteria for reading disability were double-entered for all regression analyses. This is analogous to the computation of probandwise concordance rates, in which both affected members of concordant pairs are included as probands. Standard error estimates and significance were adjusted accordingly. Models were fit using linear regression in SPSS for Windows 17.0 (SPSS, 2007).

2.4 Results

2.4.1 Descriptive Analyses

Mean reading performance scores for MZ and DZ probands, as well as those of their co-twins and co-sibs, expressed as standard deviation units from the mean of the control twins, are presented in Table 2.2. The MZ and DZ proband means are highly similar (approximately 2.5 standard deviations below the control mean). Furthermore, there is a differential regression of the MZ co-twin, DZ co-twin and co-sib means toward the mean of the control twins. The MZ co-twin mean regressed 0.21 standard deviation units toward the control mean on average, whereas that of the DZ co-twin regressed 1.02 standard deviation units. In addition, the co-sib mean regressed 1.25 standard deviation units.

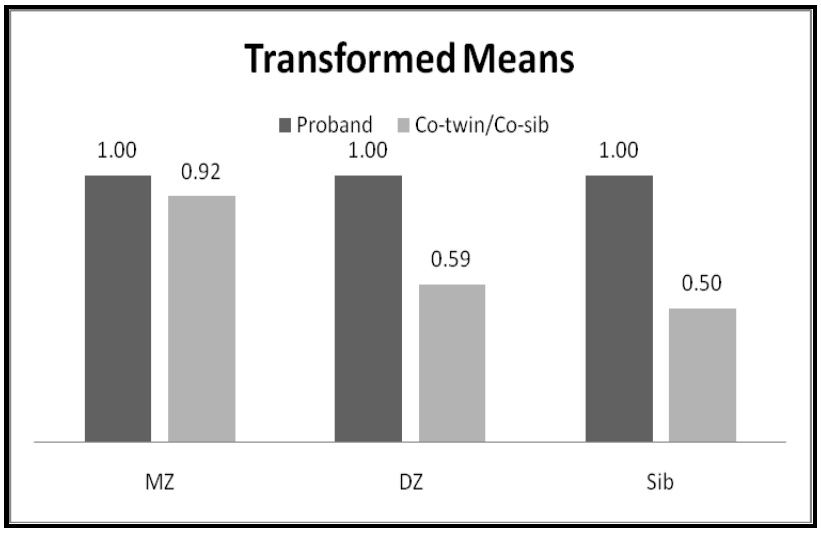
Table 2.2.

Mean standardized reading performance scores (± SD) of probands, co-twins and co-sibs

	MZ		DZ		Twin/sib	
	M	SD	M	SD	M	SD
Proband	-2.529	.766	-2.463	.854	-2.525	.811
Co-twin/co-sib	-2.324	.973	-1.442	1.336	-1.272	1.391
N pairs	254		420		303	

Corresponding transformed proband, co-twin and co-sib means are presented in Figure 2.2.

Figure 2.2.Transformed proband, co-twin and co-sib means



2.4.2 Univariate Analyses

Results of fitting Equations 1 and 2 to the twin-only data and twin-sibling data are presented in Table 2.3. When the basic model incorporating data from twins only (Equation 1) was fitted to the transformed proband and co-twin scores, the B_2 estimate was .67, confirming that the proband reading deficit in this sample is due substantially to genetic influences. Similarly, as expected, when the extended model was fitted to data from both twins and siblings (Equation 2), the B_2 estimate was also .67. However, the B_3 coefficient, while relatively small, is significant (.08 ± .04, p = .02), suggesting that shared environmental influences for members of twin pairs are greater than those of the less contemporaneous twin/nontwin sibling pairs.

Table 2.3.

Comparison of twins and twin-sibling results of DF analysis

Subjects	Model	B_2 ± S.E.	*p*	B_3 ± S.E.	*p*
Twins only	$C = B_1P + B_2R + A$	.667 ± .067	≤ .001	----	----
Twins & siblings	$C = B_1P + B_2R + B_3S + A$	.667 ± .070	≤ .001	.082 ±.035	.020
Twins & siblings[1]	$C = B_1P + B_2R + A$	.736 ± .063	≤ .001	----	----

Comparison of twin and twin-sibling results of DF analysis

[1]Ignoring DZ co-twin versus co-sibling status

2.4.2.1 Calculations from transformed co-twin and co-sib means

As illustrated in Appendix 2.1, estimates of h^2_g and differential c^2_g may also be readily calculated from the transformed co-twin and co-sib means. Given obtained estimates of h^2_g, it may be seen that $c^2_{g(t)}$ = .25, $c^2_{g(s)}$ = .17 and B_3 = .25 - .17 = .08, the difference between shared environmental influences for members of twin pairs versus twin/sib pairs.

Appendix 2.1

Parameter estimates calculated from the transformed co-twin and co-sib means

	Univariate
Transformed means	
MZ Co-twin mean (CMZ)	$h^2_g + c^2_{g(t)} = .9190$
DZ Co-twin mean (CDZ)	$½\ h^2_g + c^2_{g(t)} = .5856$
Co-sib mean (CS)	$½\ h^2_g + c^2_{g(s)} = .5039$
	h^2_g = 2(CMZ-CDZ) = 2(.9190 - .5856) = .6668
Differential c^2_g	
$c^2_{g(MZ)}$	(CMZ- h^2_g) = (.9190 - .6668) = .2522
$c^2_{g(DZ)}$	(CDZ - ½ h^2_g) = (.5856 - .3334) = .2522
$c^2_{g(s)}$	(CS - ½ h^2_g) = (.5039 - .3334) = .1705
B_3	$c^2_{g\,(t)}$ - $c^2_{g\,(s)}$ = (.2522 - .1705) = .0817

2.5 Discussion

The primary goals of the present study were to assess genetic and environmental influences on reading difficulties using data from the full CLDRC twin sample, and to fit a novel extension of the DeFries-Fulker multiple regression model (Astrom et al., 2011) to reading performance data from both twins and their nontwin siblings. The present sample of reading-disabled twin pairs and siblings tested in the ongoing CLDRC is much larger than that previously analyzed by Astrom et al. (2011), providing more rigorous tests of both the etiology of reading deficits and of special twin environments.

2.6. Summary

Early twin studies of reading deficits employed a comparison of MZ and DZ concordance rates as a test for genetic etiology. Although concordance rates of MZ twin pairs typically exceeded those of DZ pairs, samples were small and results were highly variable. In the current study, concordance rates of 254 MZ pairs and 420 DZ pairs were 65% and 33%, respectively. However, the multiple regression analysis of selected twin data facilitates a more powerful and flexible test of genetic etiology than does a comparison of concordance rates. In the current study, when the basic model (Equation 1) was fitted to transformed reading performance data from MZ and DZ twin pairs with reading difficulties, h^2_g was estimated at .67. This result, similar to those of previous studies, suggests that about two-thirds of the reading deficit of the probands is due to heritable influences.

2.6.1 Extended DF model

The extended DF model for analyzing data from both twins and siblings also provides a test for "special twin environments" (e.g., Koeppen-Schomerus et al., 2003; Medland et al., 2003; Van Grootheest et al., 2007; Young et al., 2006). In Equation 2, B_3 estimates the

difference between the DZ co-twin and co-sib means, and therefore provides a test of significance of the difference between shared environmental influences for twin pairs and for twin/sib pairs, i.e., $c^2_{g(t)} - c^2_{g(s)}$. Further, the size and significance of B_3 indicates whether Equation 1 or Equation 2 should be fitted to the combined twin and co-sibling data. If B_3 is small and non-significant, *S* may be dropped from the model, and Equation 1 fitted to the combined data from twins and siblings. However, if B_3 is significant or fairly large, h^2_g should be estimated from fitting Equation 2 to the data set. In the current study a significant difference was found between twin and sibling shared environmental influences (B_3 = .08 ± .04, p = .02). Therefore B_3 should not be dropped from the model. However, for illustrative purposes, Table 3 also presents results of fitting the more parsimonious model (Equation 1) to the combined twin and sibling data. As expected, since $c^2_{g(t)}$ is significantly larger than $c^2_{g(s)}$, the estimate of h^2_g (.74 ± .06) is substantially larger than when Equation 2 was fitted to those data (.67 ± .07).

2.6.2 Limitations

These results are comparable to those of our previous analysis of reading deficits employing this extension of the DF model (Astrom et al., 2011). However, although the B_2 estimate from Equation 2 (.67) in the present study was the same as that obtained when Equation 1 was fitted to twin data only, the standard error of the B_2 term estimated from Equation 2 was slightly larger than that for Equation1. This result differs from that of Astrom et al. (2011), which had suggested that the addition of sibling data improves power. Our present results indicate that this is not always the case. In fact, inclusion of data from siblings in twin studies may result in a reduction in power. This possibility was previously suggested by Zieleniewski et al. (1987) who noted that the inclusion of data from "weaker relationships" (e.g., siblings versus twins) in analyses of the genetic and environmental etiologies of individual

differences may actually increase standard errors of some parameter estimates. In contrast to the multiple regression analysis employed by Zieleniewski et al. (1987), the basic DF twin model and the extension described here are quite powerful. For example, in our current samples, power to detect significance of the B_2 term (h^2_g) was 1.00 both with and without sibling data, even with an alpha level of .001. Given these same parameters and a sample half this size, power is still greater than .90 in both cases. Further, the addition of sibling data could potentially increase power when more complex models are fitted to twin and sibling data, such as those testing hypotheses of differential genetic etiology. Exploratory power analyses of both simulated data and our current data tentatively suggest that power may be increased when B_3 is small and Equation 1 is fitted to combined twin and sibling data. For example, in a sample of 200 twin pairs, given an h^2_g estimate of .67, such as that obtained in the current study, a B_3 estimate near zero at .01, a change in R^2 of .10, and $\alpha = .001$, power to detect significance of h^2_g is .977. Given these same parameters, when the sample size is increased with the addition of data from 100 twin-sibling pairs, power increases to .997. Such increases in power may also be greater with smaller samples. For example, for the same parameters and a sample of 100 twin pairs, power increases from .65 to .84 with the addition of data from 50 twin-sibling pairs. This is consistent with the power estimates previously reported by Astrom et al. (2011). However, more comprehensive power analyses that assess the influence of MZ and DZ twin pair and co-sibling sample sizes and the magnitudes of the B_2 and B_3 parameter estimates are clearly warranted.

2.7 Future Directions

Finally, although the results of this study indicate that reading difficulties are highly heritable this should not minimize the importance of early reading instruction and sustained involvement in reading. Reading performance is also clearly dependent on family, school and

cultural influences, and may be remediated with rigorous intervention (e.g., Wise, Ring, & Olson, 2000; Morris, Lovett, wolf, Rose, Sevcik, Steinbach, & Shapiro, 2010). Therefore, future research on reading development should employ genetically informative designs to provide a better understanding of the variety of influences on reading deficits and the types of interventions that may enhance children's reading abilities.

Chapter 3

Etiology of the stability of reading difficulties: The Longitudinal Twin Study of Reading Disabilities

3.1 Introduction

The heritable nature of reading disabilities (RD) has been well established. Heritability estimates for deficits in reading performance range from .37 - .72 for subjects aged 7 to 20 years (e.g., DeFries & Gillis, 1991; DeFries & Alarcón, 1996; Harlaar et al., 2005; Stevenson et al., 1987). Moreover, results obtained from longitudinal studies indicate that reading deficits are generally stable (e.g., Satz, Buka, Lipsitt, & Seidman,1998), with stability correlations over intervals of 1 to 8 years ranging from .23 to .96 (e.g., Bast & Reitsma, 1998; DeFries, 1988; DeFries & Baker, 1983; Shaywitz et al., 1992; Spira & Fischel, 2005; Wagner, Torgeson, Rashotte, Hecht, Barker, Burgess, Donahue, & Garon, 1997; Williams & McGee, 1996). However, because few previous longitudinal studies of RD have utilized genetically informative designs, little is known about the etiology of this stability.

The evidence that deficits in reading are both stable and heritable suggests that genetic influences may be largely continuous throughout development, i.e., the genetic factors which are important in early reading development may also be important for later reading performance. To our best knowledge, no previous studies have assessed the etiology of the stability of reading *deficits;* however, a few studies have examined the etiology of stability of individual differences in reading *performance*. As an early first step in assessing the etiology of the stability of reading performance, DeFries and Baker (1983) tested 102 RD and control probands (i.e., 51 pairs matched for sex and age) in the Colorado Family Reading Study at average ages of 9.5 and 14.9 years. Reading performance was measured using Reading Recognition, Reading Comprehension, and Spelling subtests of the Peabody Individual

Achievement Test, (PIAT; Dunn & Markwardt, 1970). Results of structural equation modeling indicated that for families of reading-disabled children, over 60% of the longitudinal stability was attributable to parental influences. However, as this was a family study, rather than a twin or adoption study, genetic effects were not distinguishable from shared family environmental effects.

A few subsequent studies have used twin and adoption data to assess the etiology of the stability of genetic and environmental influences on reading performance within the normal range. Recently, Harlaar et al. (2007) assessed the stability of genetic influences on reading achievement in participants of the Twins Early Development Study (TEDS), a longitudinal study of twins ascertained from population 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 years using a rating scale of general reading achievement based on UK National Curriculum (NC) achievement goals for literacy. In addition, at age 10, participants completed a web-based test at home, which included an adaptation of the reading comprehension subtest of the Peabody Individual Achievement Test – Revised (PIAT-R; Markwardt, 1997). Heritability estimates of .67, .65, and .57 were obtained for reading performance at 7, 9, and 10 years respectively. Results from this study confirm that individual differences in reading performance are stable (r = .59 - .63) and suggest that 68% - 77% of the phenotypic stability correlation is genetically mediated.

Similar results were recently obtained by Wadsworth, Corley, Plomin, Hewitt, & DeFries (2006) using data from participants in the Colorado Adoption Project (CAP), an ongoing longitudinal study examining genetic and environmental influences on behavioral development. Reading performance data (PIAT Reading Recognition subtest, Dunn &

Markwardt, 1970) from adoptive and non-adoptive sibling pairs who participated in the CAP at ages 7, 12, and 16 years, were subjected to Cholesky decomposition analysis. Similar to the findings of Harlaar et al. (2007), stability correlations were substantial and ranged from 0.58 between ages 7 and 16 years to 0.71 between ages 12 and 16 years. Moreover, between 53% and 86% of these stability correlations were due to genetic influences, suggesting that those genetic factors influencing reading performance at age 7 are also operating at ages 12 and 16.

Although a few previous twin and adoption studies have assessed the etiology of stability of reading performance within the normal range, we know of no other studies that have examined the etiology of stability of reading *deficits* using genetically informative designs. Thus, the primary objective of this first longitudinal twin study of reading difficulties was to assess the genetic and environmental etiologies of the stability of reading deficits using data from twin pairs tested initially in the Colorado Learning Disabilities Research Center (CLDRC; DeFries et al., 1997), and retested 5-6 years later in the Longitudinal Twin Study of Reading Disabilities (LTSRD; Wadsworth et al., 2007). Based on previous evidence of genetic influence on the stability of reading performance within the normal range of scores, we hypothesized that genetic influences on reading difficulties are stable with largely the same genes influencing reading deficits at both time points.

3.2 Methods

3.2.1 Participants and Measures

The subjects were a subset of participants in the CLDRC who also participated in follow-up testing in the LTSRD. For a complete description of subject ascertainment and measures, please see Chapter 2, section 2.3.1, *Participants and Measures*, pg. 33.

The current sample included same sex and opposite sex twin pairs participating in the CLDRC between September 1996 and August 2000 who underwent follow-up testing in the LTSRD approximately 5-6 years after their initial participation (average age of 10.3 years at initial assessment and 16.1 years at follow-up). As of May 31, 2006, 156 twin pairs and 46 siblings have participated in the follow-up study. However, data from only those twin pairs meeting criteria for inclusion in the proband sample are included in the current analyses. The analysis sample included 18 MZ and 38 DZ pairs meeting criteria for inclusion in the initial proband sample who also participated in follow-up testing.

3.2.2 Data Analyses

3.2.2.1 Multiple Regression Analysis

While a comparison of MZ and DZ concordance rates facilitates a test for the genetic etiology of a dichotomous variable (e. g., the presence or absence of a psychiatric disorder), RD is diagnosed using quantitative measures with somewhat arbitrary cut-off points. Thus, this transformation of a continuous measure into a categorical variable such as RD versus normal, results in a loss of information regarding the continuum of variation in reading performance. Consequently, DeFries and Fulker (1985, 1988) proposed fitting a multiple regression model to continuous data from twin pairs in which at least one member of each pair has a deviant score on the variable of interest.

Reading composite data at initial assessment were fitted to the following basic multiple regression model:

$$\hat{C}_x = B_1 P_x + B_2 R + A \qquad [3.1]$$

In the current study, the CLDRC control sample of 1264 subjects represents the unselected population. Because subjects were selected based on their initial reading composite scores,

and not re-selected at follow-up, only data from the initial assessment were fitted to the univariate DF model. For a complete description of the DeFries-Fulker model, please see Chapter 2, section 2.3.2.1, pg 37.

3.2.2.2 Etiology of Stability

To assess the heritable nature of the stability of reading deficits, a bivariate extension of the basic DF model was employed in which data from the initial and follow-up sessions were fitted to the following regression model:

$$\hat{C}_y = B_1 P_x + B_2 R + A \qquad [3.2]$$

where C_y is the cotwin's predicted follow-up composite reading score, P_x is the proband's initial composite reading score, R is the coefficient of relationship, and A is the regression constant. B_1 is the partial regression of the cotwin's follow-up reading score on the proband's initial reading score and is a measure of average MZ and DZ cross-variable twin resemblance. Thus, B_1 estimates the extent to which cotwin scores on the follow-up measure are related to proband scores on the initial measure across zygosity. B_2 is the partial regression of the cotwin's follow-up reading score on the coefficient of relationship. Because the data were transformed prior to multiple-regression analysis, the bivariate B_2 coefficient is a function of the square roots of the group heritabilities for reading performance at the two time points and the genetic correlation (r_G) between them (i.e., $h_{\text{initial}} \times h_{\text{follow-up}} \times r_G$; Light & DeFries, 1995). Therefore, B_2 provides an estimate of "bivariate heritability," an index of the extent to which the proband reading deficit at the initial participation is due to genetic factors which also influence the reading deficit at follow-up. Further, the proportion of the phenotypic stability correlation (r_P) attributable to genetic influences can be obtained by dividing the B_2 estimate by r_P.

For the present study, data were analyzed from twin pairs in which at least one member of the pair met the criteria for reading disability at their initial participation, and in which both members of the pair underwent follow-up testing. Because truncate selection was employed (DeFries & Gillis, 1991), pairs in which both members were diagnosed as reading-disabled were double-entered for all regression analyses. This is analogous to the computation of probandwise concordance rates, in which both affected members of concordant pairs are included as probands. Standard error estimates and tests of significance were adjusted accordingly.

3.3 Results

Table 3.1 presents the standardized MZ and DZ proband and cotwin mean reading performance score, at each time point. Scores at initial assessment were standardized against the mean of all 1264 control twins participating in the CLDRC, whereas follow-up scores were standardized against the mean of the 93 control twins tested at follow-up. The MZ and DZ proband scores are highly similar at both time points and are more than two standard deviations below the control twin mean, suggesting that the deficit of the probands is highly stable. In addition, at both time points there is a differential regression of the MZ and DZ cotwin scores towards the control mean. In the initial sample, MZ cotwins regress only 0.03 standard deviation units and DZ cotwins regress 1.08 standard deviation units toward the control mean. Similarly, for the follow-up sample, the MZ cotwins regress 0.36 standard deviation units, whereas DZ cotwins regress 1.08 standard deviation units toward the control mean.

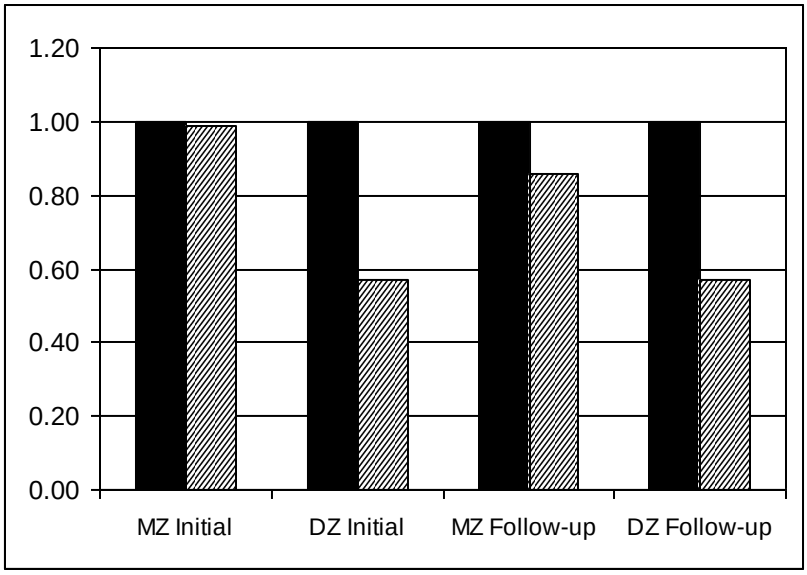
Table 3.1. Proband and cotwin standardized means (M) and standard deviations (SD) of reading composite scores at initial and follow-up testing[1]

	MZ			DZ		
	M		SD	M		SD
Initial						
Proband	-2.15	±	.836	-2.11	±	.823
Cotwin	-2.12	±	1.03	-1.03	±	1.41
Follow-up						
Proband	-2.30	±	.848	-2.24	±	.880
Cotwin	-1.94	±	1.36	-1.16	±	1.53

[1]initial scores have been standardized against the mean of 1264 control twins participating in the CLDRC; follow-up scores have been standardized against the mean of 93 control twins participating in the LTSRD.

When the transformed proband and cotwin initial scores (Figure 3.1) were fitted to the basic regression model (Equation 3.1), the resulting $B_2 = h^2_g = 0.84 \pm 0.26$ ($p \leq .002$), indicating that the proband reading deficit in this subsample is due substantially to genetic influences. Further, when the transformed proband initial scores and cotwin follow-up scores were fitted to the bivariate model (Equation 3.2), $B_2 = 0.65 \pm 0.32$ ($p < .05$), suggesting that about two-thirds of the proband deficit in reading at initial assessment is due to genetic factors that also influence reading difficulties at follow-up. Moreover, the ratio of B_2 to the observed correlation (0.84) between initial and follow-up scores suggests that common genetic influences account for approximately 75% (0.65/0.84 = 0.77) of the stability between reading difficulties at the initial and follow-up sessions. The corresponding transformed proband and cotwin means, wherein each score is expressed as a deviation from the mean of the control population and then divided by the difference between the proband and control means, are presented in Figure 3.1.

Figure 3.1 Proband and co-twin transformed means of reading composite scores at initial and follow-up testing



3.4 Discussion

Although a few previous studies have shown that reading deficits are stable and heritable, the genetic and environmental etiologies of this stability have not been previously investigated. The goal of this first longitudinal twin study of RD was to assess the etiology of the stability of reading deficits at two time points using behavioral genetic methods. Accordingly, data from twin pairs first tested in the CLDRC, and again 5 to 6 years later in the LTSRD, were subjected to bivariate DF analysis (Light & DeFries, 1995).

In the current study, the reading composite scores were highly stable over the 5-6 year interval (r_P = .84), somewhat higher than the stability correlations reported by Harlaar et al. (2007) and Wadsworth et al. (2006). In addition, the reading deficit of the probands was remarkably stable, with proband means more than two standard deviations below those of the

controls at each assessment. This is especially noteworthy given that two different test versions were administered at initial and follow-up sessions (i.e., the PIAT at initial assessment and the PIAT-R at the follow-up).

When composite reading performance data collected from twin pairs at their initial assessment were subjected to DF multiple regression analysis, an h^2_g estimate of 0.84 (± 0.26) was obtained. Although this h^2_g estimate is somewhat higher than that obtained from the full CLDRC proband sample of 283 MZ and 402 DZ pairs (h^2_g = .61 ± .06) for same-sex and opposite-sex pairs combined), the two estimates are not significantly different ($p > .30$).

When the bivariate extension of the multiple regression model (Equation 3.2) was fitted to proband scores at initial assessment and cotwin scores at follow-up, the resulting estimate of bivariate heritability was 0.65 ± 0.32 ($p < .05$), indicating that about two-thirds of the proband deficit at initial assessment was due to genetic influences which also influence reading deficits at follow-up. Further, these results suggest that approximately 75% of the observed stability correlation is due to shared genetic influences. These findings are highly consistent with those of Harlaar et al. (2007), and Wadsworth et al. (2006), who found that 58% to 77% of the stability of reading performance in the normal range between ages 7 and 16 was due to genetic influences.

3.5 Summary

The preliminary results of this first longitudinal twin study of reading difficulties suggest that reading deficits are not only stable, but that this stability is due largely to heritable influences. However, the current sample of twin pairs meeting criteria for inclusion in the proband sample and on whom we have follow-up data is still very small. Nevertheless, the results obtained in the present study are highly consistent with those of previous longitudinal

studies of reading performance within the normal range and both the univariate and bivariate h^2_g estimates are statistically significant. Moreover, follow-up testing of RD and control twin pairs continues in the LTSRD, thereby eventually facilitating more rigorous assessments of the etiology of difficulties in various reading-related cognitive processes and their stabilities.

Chapter 4

DeFries-Fulker analysis of longitudinal reading performance data from twin pairs ascertained for reading difficulties and from their nontwin siblings

4.1 Introduction

During the past few decades, much has been learned about the heritable nature of reading ability and disability. For example, twin and adoption studies have shown that individual differences in reading performance are highly heritable, although estimates of heritability have ranged from .18 to .81 for subjects 7 to 16 years of age (Alarcón, DeFries, & Fulker, 1995; Harlaar et al., 2007; Stevenson et al., 1987; Wadsworth et al., 2006; Wadsworth et al., 2007). At comparable ages, heritability estimates for reading deficits range from .37 to .72 (e.g., DeFries & Alarcón, 1996; DeFries & Gillis, 1991; Harlaar et al., 2005). Moreover, results obtained from longitudinal studies have shown that stability correlations for reading performance are considerable, ranging from .23 to .96 over intervals of 1 to 8 years (e.g., Bast & Reitsma, 1998; DeFries & Baker, 1983; Shaywitz et al., 1992; Wadsworth et al., 2006; Wadsworth et al., 2007; Wagner et al., 1997). When reading performance was estimated by Hulslander, Olson, Willcutt, & Wadsworth (2010) as a latent trait from multiple measures of word recognition, the stability correlation between age 10 and 15 years was .98.

Although studies have consistently yielded evidence that both individual differences in reading performance and reading deficits are heritable and stable, little is known about the genetic and environmental etiologies of this stability. Only a small number of studies have evaluated the etiology of stability utilizing genetically informative designs, and they have primarily assessed the etiology of individual differences in reading performance within the normal range. For example, in the Colorado Adoption Project, word recognition was examined at ages 7, 12, and 16 in a sample of adoptive and nonadoptive sibling pairs (Wadsworth et al.,

2006). Results indicated substantial stability for reading performance with stability correlations of .62 between ages 7 and 12, .71 between ages 12 and 16, and .58 between ages 7 and 16. Moreover, the proportion of observed stability attributable to shared genetic influences was 53%, 62%, and 86%, respectively.

Analyzing data from a younger sample participating in the International Longitudinal Twin Study (ITLS), Byrne, Samuelsson, Wadsworth, Hulslander, Corley, DeFries, Quain, Willcutt, & Olson (2007) found that genetic influences accounted for approximately 90% of the observed stability between word reading in kindergarten and first grade. However, Grade 1 reading was also influenced by a second genetic factor, suggesting both genetic stability and genetic change at this developmental milestone. Environmental contributions to stability were limited to nonshared environmental influences, which were also a significant source of change. In contrast, shared environment was important only for kindergarten word reading, and did not contribute to stability.

Petrill, Deater-Deckard, Thompson, Schatschneider, Dethorne, & Vandenbergh (2007) assessed the longitudinal stability of reading-related skills and reading outcomes of children who were tested on two occasions in the Western Reserve Reading Project (WRRP): in kindergarten or first grade, and a year later. Measures of reading related skills included phonological awareness (PA), expressive vocabulary (VOCAB) and rapid automatized naming (RAN); outcome was measured by performance on tests of letter knowledge (LET), word knowledge (WORD), phonological decoding (PD) and passage comprehension (COMP). Results of a series of Cholesky decomposition analyses suggested that, with the exception of RAN, genetic influences accounted for a significant proportion of the observed stability (as much as 42%, depending on the measure) for all measures of reading outcome at the two

assessments. Shared environmental influences contributed to stability of PA, VOCAB, and LET. Nonshared environment had small, but significant, effects on the stability of RAN, WORD, PD and COMP.

These results are highly similar to those of Harlaar et al. (2007) who examined the genetic and environmental influences on reading performance and reading stability from the Twins Early Development Study (TEDS). For this study, reading achievement was measured by teacher assessments of subjects at ages 7, 9, and 10 years, using a rating scale of general reading achievement that referenced U.K. National Curriculum (NC) achievement goals for literacy. Results of the study showed substantial heritabilities for reading performance ranging from .57 to .67. In addition, NC scores were found to be significantly correlated across all three age groups; .62 for ages 7 to 9, .59 for ages 7 to 10, and .63 for ages 9 to 10, with genetic influences contributing to 77%, 68%, and 77% respectively, of the observed stability.

To our knowledge, only one study has assessed the etiology of stability of reading deficits (Astrom, Wadsworth, & DeFries 2007). Subjects from this preliminary longitudinal analysis of reading disabilities of school-aged children (aged 8 – 16 years) were tested in the Colorado Learning Disabilities Research Center (CLDRC, DeFries, et al., 1997) and again approximately 5-6 years later in the Longitudinal Twin Study of Reading Disability (LTSRD, Wadsworth et al. 2007). Probands scored approximately two standard deviations below the mean of the control sample at both measurement occasions (Astrom et al. 2007), and a stability correlation of .84 was obtained between the two assessments. A bivariate extension of the DeFries-Fulker basic multiple regression model for analysis of selected twin data (DF; DeFries & Fulker, 1985; Light & DeFries, 1995) yielded an estimate of bivariate heritability of .65 (± .32), suggesting that

nearly two-thirds of the proband reading deficit at follow-up was due to genetic factors that also influenced reading at initial assessment (Astrom et al., 2007).

Although our current longitudinal sample of MZ and DZ twin pairs who met criteria for inclusion in the proband sample at their initial assessment is now nearly twice as large, it is still somewhat underpowered for more complex analyses. However, since the inception of the CLDRC, we have also collected data from siblings of approximately half of the MZ and DZ probands. Consequently, using a novel application of the DF multiple regression method, we have included data from siblings of probands in the present analysis. Thus, the primary objectives of this study were twofold: (1) to assess more rigorously the etiology of the stability of reading deficits in a larger sample of twin pairs than was previously analyzed by Astrom et al. (2007); and (2) to apply an extension of the DF method (DeFries and Fulker, 1985; 1988) to analyze both twin and sibling data and test for "special twin environments" (i.e., a measure of the extent to which shared environmental influences for members of twin pairs differ from those for nontwin-sibling pairs). Based on previous findings of genetic influences on the stability of reading performance within the normal range, as well as our preliminary findings regarding the stability of reading deficits, we hypothesize that largely the same genetic influences on reading deficits are manifested at initial and follow-up assessments. Moreover, given the results obtained from a small twin and sibling study of individual differences in reading-related measures (Zieleniewski, et al., 1987), we hypothesize that shared environmental influences for reading deficits are similar for members of twin pairs and between twins and their nontwin siblings.

4.2 Methods

4.2.1 Participants and Measures

Subjects in the present study were first tested in the ongoing CLDRC between September 1996 and March 2003, and also participated in follow-up testing approximately five years later in the LTSRD. A complete description of participants and measures can be found in Chapter 2, section 2.3.1, *Participants and Measures*, pg 33.

The current sample included 33 MZ and 64 DZ twin pairs in which at least one twin met proband criteria at initial assessment in the CLDRC, and 44 of their nontwin siblings. Selected twins and both their co-twins and siblings underwent follow-up testing approximately 5 years after their initial participation (average age of 10.6 years at initial assessment and 15.5 years at follow-up). Because concordant twin pairs were double entered, the siblings were paired with both twins in those cases. As a result, there were 58 twin/sib pairings; in 21 of those pairings the twin was selected from an MZ pair and in 37 pairings the twin was selected from a DZ pair. For standardization and transformation of the variables at initial assessment, the control sample comprised of 284 subjects tested in the CLDRC during the same time period as those subjects who participated in follow-up testing. At follow-up the control sample included 171 control twins who participated in follow-up testing.

4.3. Analyses

4.3.1 Multiple Regression Analysis

Early twin studies of reading difficulties compared concordance rates for deviant scores in identical and fraternal twin pairs to test for genetic etiology (DeFries & Alarcón, 1996). However, because reading performance is a continuous trait, a comparison of concordance rates of "affected" and "unaffected" pairs does not make optimal use of the data. Thus, DeFries and Fulker (1985, 1988) proposed a multiple regression analysis of twin data to assess the etiology of deviant scores, as well as individual differences within selected groups. This method has

become a standard of behavioral genetic analysis and the method of choice for analyzing data from selected samples. This multiple regression method is particularly appropriate when analyzing data from probands who are selected because of deviant scores on a continuous variable such as reading performance; the differential regression of MZ and DZ co-twin scores toward the mean of the unselected population provides a test of genetic etiology (DeFries & Fulker, 1985). As MZ twins are genetically identical and DZ twins share only half of their segregating genes on average, the scores of DZ co-twins should regress more toward the mean of the unselected population if the condition is heritable. Consequently, when the MZ and DZ proband means are approximately equal, a simple *t*-test of the difference between the MZ and DZ co-twin means provides a test of genetic etiology. However, DeFries and Fulker (1985, 1988) proposed that a multiple regression analysis of such data, in which a co-twin's score is regressed on both the proband's score and the coefficient of relationship, facilitates a more flexible and statistically powerful test.

The basic DF model is as follows:

$$\hat{C}_x = B_1 P_x + B_2 R + A \qquad [4.1]$$

The basic DF model is described in Chapter 2, section 2.3.1, *Participants and Measures*, pg 33.

In order to incorporate sibling data, a simple extension of the basic DF model can be simultaneously fitted to transformed data from selected twins, their co-twins and co-sibs. The expected transformed MZ and DZ co-twin means (see DeFries & Fulker, 1988) and the corresponding transformed co-sib means are presented in Table 4.1. From these expected values, it may be seen that the difference between the DZ co-twin mean and the co-sib mean is a simple function of the difference between shared environmental influences in twins versus sibling pairs.

Table 4.1

Expected transformed[1] co-twin and co-sib means

Subjects	Model
MZ Co-twins	$h^2_g + c^2_{g(t)}$
DZ Co-twins	$½\, h^2_g + c^2_{g(t)}$
Co-sibs	$½\, h^2_g + c^2_{g(s)}$

[1]Scores are expressed as a deviation from the unselected population mean and then divided by the difference between the proband and population means (see DeFries & Fulker, 1988). (after Astrom et al., 2011)

Thus, to test for differential c^2_g between twins and siblings, the following extended basic model can be simultaneously fitted to transformed data from selected twins, their co-twins and co-sibs:

$$\hat{C}_x = B_1 P_x + B_2 R + B_3 S + A \qquad [4.2]$$

where C is now the co-twin's or co-sib's predicted score, P is the proband's score, R is the coefficient of relationship (1.0 for MZ pairs and 0.5 for both DZ pairs and twin/sib pairs), and S is a dummy code for pair type, i.e., twin pair versus twin-sibling pair (+.5 for MZ twins, +.5 for DZ twins and -.5 for twin-sib pairs). When this model is fitted to the data and all three partial regression coefficients are estimated simultaneously, B_3 estimates the difference between $c^2_{g(t)}$ and $c^2_{g(s)}$ which equals the difference between the DZ co-twin (CDZ) and co-sib (CS) means, as can be observed in the Appendix. As a result, B_3 provides a direct test of significance for the difference between environmental influences shared by DZ twin pairs versus those shared by twin-sib pairs, i.e., a test for "special twin environment". As in the basic model, B_2 estimates h^2_g from twice the difference between the MZ and DZ co-twin means.

Because B_3 estimates the difference between $c^2_{g(t)}$ and $c^2_{g(s)}$, its significance is relevant for obtaining an estimate of h^2_g based upon an analysis of the combined twin and co-sibling data. If B_3 is small and not significant, S may be dropped from the extended model, and Equation 4.1 fitted to the combined data set of twins and siblings. In that case, B_2 will estimate h^2_g from both the twin *and* co-sib data, rather than from only the twin data. However, if B_3 is significant or relatively large, h^2_g should be estimated from fitting Equation 2 to the combined data set.

Because subjects were not reselected at follow-up, univariate DF models were fitted only to data from the initial assessment. Then, to assess the heritable nature of the *stability* of reading deficits a bivariate extension of the basic DF model was fitted to data from the initial and follow-up assessments as follows:

$$\hat{C}_y = B_1P_x + B_2R + A \qquad [4.3]$$

where C_y is the co-twin's or co-sib's predicted composite reading score at follow-up, P_x is the proband's initial composite reading score, R is the coefficient of relationship, and A is the regression constant. B_1 is the partial regression of the co-twin's or co-sib's follow-up reading score on the proband's initial reading score and is a measure of average MZ and DZ cross-variable twin resemblance. Thus, B_1 estimates the extent to which co-twin and co-sib scores on the follow-up measure are related to proband scores on the initial measure across zygosity. B_2 is the partial regression of the co-twin's or co-sib's follow-up reading score on the coefficient of relationship. When the data are transformed prior to multiple-regression analysis, the bivariate B_2 coefficient is a function of the square roots of the group heritabilities for reading performance at the two time points and the genetic correlation (r_G) between them (i.e., $h_{initial} \times h_{follow\text{-}up} \times r_G$; Light & DeFries, 1995). Therefore, B_2 provides an estimate of "bivariate heritability" ($h^2_{g(Biv)}$), an index of the extent to which the proband reading deficit at follow-up is due to heritable factors

which also influenced the reading deficit at the initial assessment. Further, the proportion of the phenotypic stability correlation (r_p) attributable to genetic influences can be obtained by dividing the B_2 estimate by r_p. Finally, when the following extended bivariate model is fitted to both twin and sibling data (Equation 4), B_3 provides a test of significance for the difference between bivariate shared environmental influences in twins and siblings:

$$\hat{C}_y = B_1P_x + B_2R + B_3S + A \quad [4.4]$$

Because truncate selection was employed (DeFries & Gillis, 1991), pairs in which both members met criteria for RD were double-entered for all regression analyses. This is analogous to the computation of probandwise concordance rates, in which both affected members of concordant pairs are included as probands. Standard error estimates and tests of significance were adjusted accordingly. Models were fit using linear regression in SPSS for UNIX server.

4.4 Results

Table 4.2 presents the standardized mean reading performance scores for MZ and DZ probands, as well as those of their co-twins and co-sibs at each assessment. The MZ and DZ proband means are highly similar at both initial and follow-up assessments (averaging about 2 standard deviations below the respective control means) suggesting that the deficit of the probands is highly stable. In addition, at each assessment there is a differential regression of the MZ co-twin, DZ co-twin and co-sib means toward the mean of the control twins. At the initial assessment, the MZ co-twin mean regressed 0.23 standard deviation units toward the control mean on average, whereas that of the DZ co-twin regressed 0.92 standard deviation units. In addition, the co-sib mean regressed 1.30 standard deviation units. Similarly, at follow-up the MZ co-twin mean regressed 0.26 standard deviation units toward the control mean on average,

whereas the DZ co-twin and co-sib means regressed 0.75 and 1.00 standard deviation units suggesting that reading deficits are both stable and substantially due to heritable influences.

Table 4.2

Mean standardized reading performance scores (± SD) of probands, co-twins and co-sibs at initial[1] and follow-up test sessions[2]

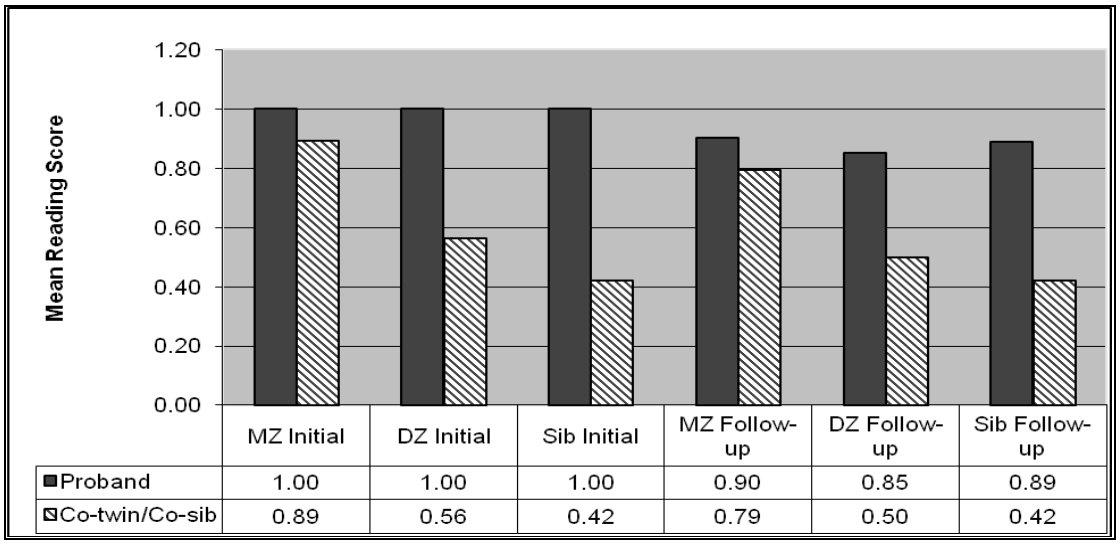
	Initial					
	MZ pairs		**DZ pairs**		**Twin/sib pairs**	
	M	**SD**	**M**	**SD**	**M**	**SD**
Proband	-2.15	± .95	-2.11	± .87	-2.24	± .95
Co-twin/sib	-1.92	± 1.17	-1.19	± 1.39	-0.94	± 1.22
	Follow-up					
Proband	-2.00	± .67	-1.93	± .64	-2.14	.70
Co-twin/sib	-1.74	± .99	-1.18	± 1.16	-1.14	1.08

[1]Standardized against the mean of 284 Control twins participating in the initial assessment in the CLDRC.

[2] Standardized against the mean of 171 Control twins participating in the LTSRD at follow-up.

Corresponding transformed proband, co-twin and co-sib means are presented in Figure 4.1.

Figure 4.1.Mean proband, co-twin and co-sib transformed reading scores at initial and follow-up assessments



	MZ Initial	DZ Initial	Sib Initial	MZ Follow-up	DZ Follow-up	Sib Follow-up
Proband	1.00	1.00	1.00	0.90	0.85	0.89
Co-twin/Co-sib	0.89	0.56	0.42	0.79	0.50	0.42

Table 4.3 presents results of the basic univariate regression analysis of twin-only data and twin-sibling data. When the basic model incorporating data from twins only (Equation 4.1) was fitted to the transformed proband and co-twin initial scores, the B_2 estimate was .67, confirming that the proband reading deficit in this sample is due substantially to genetic influences. Similarly, as expected, when the extended model was fitted to data from both twins and siblings (Equation 4.2), the B_2 estimate was again .67, but slightly more significant. In contrast, B_3 was nonsignificant, but not trivial (.14). Therefore, although the more parsimonious model (Equation 4.1) was also fitted to the combined twin and sibling data for illustrative purposes, as shown in Table 4.3 the resulting estimate of h^2_g was biased upward (.79). However, if B_3 were nonsignificant and relatively small, fitting the more parsimonious model to the combined data set would be appropriate.

When the bivariate model (Equation 4.3) was fitted to the transformed proband initial scores and co-twin follow-up scores, the estimate of bivariate h^2_g was .59, suggesting that approximately 60% of the proband deficit in reading at follow-up is due to genetic factors that also influenced reading difficulties at the initial assessment. As expected, when the bivariate model was extended to include siblings (Equation 4.4), the resulting B_2 estimate was also .59, but more highly significant. Moreover, the ratio of B_2 to the observed correlation between initial and follow-up scores (.86) suggests that common genetic influences account for nearly 70% (.59/.86 = .69) of the stability of reading difficulties in this sample. However, the estimate of bivariate B_3 was relatively small (.08) and non-significant (p = .415). Table 4.3 also presents results of fitting the more parsimonious model to the longitudinal twin and sibling data, combining the data from DZ co-twins and co-siblings.

Table 4.3

Comparison of twin and twin-sibling results of univariate and bivariate analysis from initial and follow-up assessments

Model	Subjects	Model	B_2 ± S.E.	*p*	B_3 ± S.E.	*p*
Univariate	*Twins only*	$C = B_1P + B_2R + A$	.67 ± .23	= .004	----	----
	Twins & siblings	$C = B_1P + B_2R + B_3S + A$	.67 ± .22	= .003	.14 ± .10	.167
	Twins & siblings[1]	$C = B_1P + B_2R + A$	.79 ± .20	≤ .001	----	----
Bivariate	*Twins only*	$C_y = B_1P_x + B_2R + A$	.59 ± .22	= .008	----	----
	Twins & siblings	$C_y = B_1P_x + B_2R + B_3S + A$	.59 ± .21	= .006	.08 ± .10	.415
	Twins & siblings[1]	$C_y = B_1P_x + B_2R + A$	.66 ± .20	= .001	----	----

[1]Ignoring DZ co-twin versus co-sib status

As shown in Appendix 4.1, estimates of both univariate and bivariate h^2_g and differential c^2_g may also be readily calculated from the transformed co-twin and co-sib means.

Appendix 4.1

Parameter estimates calculated from the expectations in Table 1 and the transformed co-twin and co-sib means

Univariate	
Transformed	
MZ Co-	$h^2_g + c^2_{g(t)} = .8947$
DZ Co-	$½\ h^2_g + c^2_{g(t)} = .5620$
Co-sib	$½\ h^2_g + c^2_{g(s)} = .4199$
	h^2_g = 2(CMZ-CDZ) = 2(.8947 -
Differential	
$c^2_{g(MZ)}$	(CMZ- h^2_g) = (.8947 - .6654) = .2293
$c^2_{g(DZ)}$	(CDZ - ½ h^2_g) = (.5620 - .3327) =
$c^2_{g(s)}$	(CS - ½ h^2_g) = (.4199 - .3327) =
B_3	$c^2_{g\,(t)} - c^2_{g\,(s)}$ = (.2293 - .0872) = .1421
Bivariate[1]	
Transformed	
MZ Co-twin	$h^2_{g(biv)} + c^2_{g(biv)(t)}$ = .7934
DZ Co-twin	$½\ h^2_{g(biv)} + c^2_{g(biv)(t)}$ = .4989
Co-sib mean	$½\ h^2_{g(biv)} + c^2_{g(biv)(s)}$ = .4203
	$h^2_{g(biv)}$ = 2(CMZ-CDZ) = 2(.7934 -
Differential	
$c^2_{g(biv)(t)}$	(CMZ- $h^2_{g(biv)}$) = (.7934 - .5890) =
$c^2_{g(biv)(t)}$	(CDZ - ½ $h^2_{g(biv)}$) = (.4989 - .2945) =
$c^2_{g(biv)(s)}$	(CS - ½ $h^2_{g\ (biv)}$) = (.4203 - .2945) =
B_3	$c^2_{g(biv)(t)} - c^2_{g(biv)(s)}$ = (.2044 - .1258) =

[1]Co-twin and co-sib mean reading scores at follow-up are transformed based on the proband mean at initial assessment.

4.5 Discussion

Results obtained from previous studies have shown that reading deficits are stable and heritable; however, the genetic and environmental etiologies of the stability of reading deficits have not been well characterized. Therefore, the primary objectives of the present study were (1) to assess more rigorously genetic and environmental influences on the stability of reading deficits using data from a larger sample of twin pairs and their siblings tested in the CLDRC and re-tested approximately 5 years later in the LTSRD and (2) to fit a novel extension of the DeFries-Fulker multiple regression model (DeFries & Fulker, 1985) to reading performance data from both twins and siblings, potentially increasing power and facilitating a test for "special twin environments."

In the current study, the average reading performance of the probands at their initial assessment was approximately two standard deviations below those of the controls, and this deficit persisted at follow-up. The stability correlation of probands was .72 ± .04 and that for controls was .75 ± .03. This result is consistent with those of previous studies that have found both reading deficits and reading performance within the normal range to be highly stable (e.g., Bast & Reitsma, 1998; DeFries, 1988; DeFries & Baker, 1983; Hulslander, et al., 2010; Shaywitz et al., 1992; Wadsworth et al., 2006; Wadsworth et al., 2007; Wagner et al., 1997).

Results of univariate DF analysis of data from twins and siblings at their initial assessment suggested that about two-thirds of the proband deficit in this sample was due to genetic influences. Because subjects were not reselected at follow-up, univariate DF analyses were not conducted for follow-up data. When the bivariate extension of the multiple regression model was fitted to proband scores at initial assessment and co-twin/co-sib scores at follow-up, combining DZ co-twin and co-sibling status, the resulting estimate of bivariate heritability was

0.66 (± .20, p = .001), suggesting that about two-thirds of the proband deficit at follow-up was due to genetic influences which also influenced reading deficits at the initial assessment. Further, common genetic influences accounted for approximately 70% of the observed stability between reading scores at the initial and follow-up assessments.

These results are highly consistent with our previous longitudinal analysis of the stability of reading deficits which obtained an estimate of bivariate heritability of .65 (± .32) and found that about 75% of the stability of reading difficulties between initial and follow-up assessments was due to common genetic influences (Astrom et al., 2007). However, our previous study included only twins. Results of the current study suggest that including sibling data increases power at least to some extent. For example, given the effect size estimated from the basic univariate model and sample of 97 twin pairs, power to detect significance of the B_2 term, i.e., h^2_g, is about .85. However, when the 44 siblings are also included and the basic model is fitted to the combined data sets, power improves to .97. Similarly, power for the bivariate model using only twin data is .79; including the sibs, assuming the same parameters, power is increased to about .95. Because the basic DF twin model is quite powerful, even with relatively small samples, the addition of sibling data may be especially beneficial when fitting more complex models, such as those testing hypotheses of differential genetic etiology.

The similarity of results from Astrom et al. (2007) and the current study may be expected as there is considerable overlap between the two samples. However, previous studies of individual differences in reading performance also support these findings. For example, Byrne et al. (2007) found that 90% of the observed stability of word reading in twin pairs between kindergarten and grade one were due to common genetic influences. Results from the TEDS study (Harlaar et al., 2007) were also highly similar with more than two-thirds of the phenotypic

stability being genetically mediated. Although Petrill et al. (2007) reported somewhat lower estimates (at most, 42% of the phenotypic stability was due to genetic influences), their study examined the genetic variance of individual subtests of several different measures utilizing Cholesky decomposition models. Our findings are also highly similar to results from the Colorado Adoption Project (Wadsworth et al., 2006) in which data from related and unrelated sibling pairs were analyzed, and 86% of the phenotypic correlation between reading performance at ages 7 and 16 was accounted for by genetic influences. These remarkably similar results across different studies with different types of subjects, ages, measures and analytical methods suggest that this is a highly robust finding.

The extended DF model for analyzing data from both twins and siblings also provides a test for "special twin environments" (e.g., Koeppen-Schomerus et al., 2003; Medland et al., 2003; Van Grootheest et al., 2007; Young et al., 2006). In fact, B_3 provides a direct estimate of the difference between shared environmental influences for members of twin pairs versus those for less contemporaneous twin/co-sibling pairs. Moreover, its magnitude and significance indicate whether the estimate of h^2_g should be based upon an analysis of the combined twin and co-sibling data. If B_3 is small and not significant, S may be dropped from the extended model, and Equation 4.1 fitted to the combined data from twins and siblings. However, if B_3 is significant or relatively large, h^2_g should be estimated from fitting Equation 4.2 to the data set. In the current study, no significant differences were found between twin and sibling shared environmental influences. However, the parameter estimate of .14 for differential c^2_g in the univariate model is not trivial. Consequently, when the B_3 term was dropped from the extended twin/sib model, the resulting estimate of h^2_g was inflated. Thus, results of fitting the more parsimonious model (Equation 4.1) to the combined twin and sibling data are presented for

illustrative purposes only. It's also important to note that B_3 for this study (.14) was non-significant whereas findings for B_3 for our earlier study (.08) were significant; there is differential power to detect these results as the present study consists of 44 siblings compared to the previous study of 303 siblings.

4.6 Summary

Finally, although our findings suggest that reading deficits are highly stable and that this stability is due principally to genetic influences, this should not deter our best efforts with regard to environmental intervention and remediation. There are multiple pathways to poor reading (Gilger & Kaplan, 2001), and the relative contributions of genetic and environmental influences may differ depending on the measure (Gayán & Olson, 2001). Reading disabilities do not generally segregate in families in a simple Mendelian fashion; thus, multiple genetic and environmental factors almost certainly influence reading abilities and may also interact (e.g., Friend, DeFries, & Olson, 2008). Nevertheless, our results suggest that intensive remediation efforts may be needed to compensate for genetic and other biological constraints on learning rates (Olson, Hulslander, Christopher, Keenan, Wadsworth, Willcutt, Pennington, & DeFries, 2011). Therefore, in order to identify and treat at risk children more effectively and account for their individual abilities and disabilities, additional longitudinal studies of the genetic and environmental etiologies of reading disability are clearly warranted (Olson, Keenan, Byrne, Samuelsson, Coventry, Corley, Wadsworth, DeFries, Pennington, & Hulslander, 2011).

Chapter 5

Gender Differences in the heritability and stability of reading difficulties: A twin study from the Colorado Learning Disabilities Research Center

5.1 Introduction

Reading difficulties account for a significant proportion of learning disabilities (LD) with estimates ranging from 80-90% of all diagnosed cases of LD (Beitchman, Wilson, Brownlie, Walters, Inglis, & Lancee, 1996; Lyon, 1996). An arduous task over the past decades has been to disentangle factors contributing to reading difficulties with multiple studies examining both genetic and environmental effects. The issue of gender differences has remained a question of interest and continues to be a controversial element for studies of reading disability in which both mean differences in reading-related measures and prevalence rates are compared (DeFries 1989; Finucci & Childs, 1981; Hawke et al., 2006; Knopik, DeFries, & Alarcón, 1996; Shaywitz et al., 1990; Stevenson 1992; Vogel 1990; Wadsworth et al., 1992; Wadsworth & DeFries, 2005). Results from multiple studies examining sex differences provide little or no evidence for a differential genetic etiology of reading difficulties for males and females, although prevalence studies typically identify male subjects as reading disabled at a much higher rate than female subjects. Further, additional studies have suggested that different developmental processes may underlie the various results based on gender (Dirks, Spyer, van Lieshout, & de Sonneville, 2008; Feldman, Levin, Fleischmann, Jallad, Kushch, Gross-Glenn, Rabin, & Lubs, 1995; Kempe, Gustafson, & Samuelsson, 2011; Kush & Watkins, 1996).

In an early study by Stevenson (1992), data from a twin sample (102 MZ pairs, 111 same-sex DZ pairs, and 72 opposite-sex DZ pairs) from hospitals in London examined whether or not reading disability for males and females varied as a function of reading disability

definition. In order to examine this possibility, twelve alternative definitions for reading disability were used to examine sex differences. These included age-discrepancy, discrepancy between IQ and reading, and discrepancy between actual reading scores and that based on expected IQ. Each of these methods was applied to males and females together and males and females separately. Results indicated a higher prevalence rate for males in general. However, it was noted that the sex ratio was reduced when separate alternative definitions took into account differences in the mean scores for males and females, suggesting that characteristics of subjects identified as reading disabled depended on the definition used. Although males showed a higher prevalence rate, evidence from multiple regression analysis found no consistent gender differences in group heritability estimates for the various indicators of reading disability.

In order to examine genetic etiology as a function of gender, Wadsworth et al. (2000) employed data from 206 MZ twin pairs, 159 same-sex DZ twin pairs, and 117 opposite-sex pairs. A discriminant function score was computed for each child, calculated from subtests of the PIAT; Reading recognition, Reading comprehension, and Spelling. The difference in concordance rates between MZ and same-sex DZ pairs was somewhat greater for females than for males, however this difference was not significant ($p \geq 0.70$). Group heritability estimates for males and females were nearly identical ($h^2_g = 0.58$ and 0.59, respectively, $p \geq 0.90$), thereby providing little or no evidence for a differential etiology as a function of gender.

A more recent study by Harlaar et al. (2005) analyzed data from subjects in the Twins Early Development Study (TEDS), a longitudinal, population-based study of twin pairs born in England and Wales. Subjects were administered the Test of Word Reading Efficiency (TOWRE; Torgesen, Wagner, & Rashotte, 1999). In order to be classified as reading disabled, subjects were selected from the 10^{th} and 5^{th} percentiles of the distribution of the TOWRE. Data from MZ

(n = 308 at the 10th percentile and 153 at the 5th percentile) and same-sex DZ (n = 246 at the 10th percentile and 127 at the 5th percentile) were fitted to a DeFries-Fulker multiple regression model. Larger group heritability estimates were obtained for males (h^2_g = 0.68 at the 10th percentile and 0.60 at the 5th percentile) than females (h^2_g = 0.50 at the 10th percentile and 0.40 at the 5th percentile). These results, suggesting that the heritability of reading difficulties is higher for males than for females, is in contrast to previous results obtained by earlier studies.

Further, a study from the Colorado Learning Disabilities Research Center (CLDRC, DeFries et al. 1997) examined data from both same-sex and opposite-sex twin pairs (Wadsworth et al., 2005). Discriminant function reading scores (DISCR) from 634 twin pairs (264 MZ pairs, 214 same-sex DZ pairs, and 156 opposite-sex pairs) were subjected to DeFries-Fulker sex-limitation analysis using the model-fitting approach of Purcell and Sham (2003). Findings from the combined sample of male and female twins suggested that more than half of the proband reading deficit could be accounted for by genetic factors (h^2_g = .58, $p \leq .001$). An extension of the model tested for gender differences in the magnitude of genetic influences for reading deficits, and although h^2_g estimates for females were somewhat higher than males (.63 vs .53, respectively), the difference was nonsignificant ($p > .3$). In addition to examining the etiology of reading difficulties as a function of gender, Wadsworth et al., (2005) examined the effect of age on differential heritability. Given the wide age range for the subjects (8.0 to 20.0 years), the sample was divided into two groups, an older cohort and a younger cohort based on the mean age of 11.5 years. The mean age for the younger group was 9.6 years and for the older group, 14.1 years. Data were again subjected to DF sex-limitation analysis (Purcell & Sham, 2003). Results for the younger group indicated that heritability for females was higher than for males (.67 vs. .53, respectively) whereas, differences for females and males in the older group were negligible

(.53 vs. .55, respectively). In addition, there were no significant differences between h^2_g estimates for young males and young females. Further, the biometrical parameter estimates were equated without a significant loss of model fit ($p > .3$). Results from this study noted a trend toward higher heritability for females with regard to reading deficits when compared to males at young ages. However, results overall found little evidence for gender differences in the etiology of reading disabilities.

Hawke et al. (2006) analyzed data from subjects in either the Colorado Reading Project (DeFries, 1985) or the CLDRC (DeFries et al., 1997). The sample consisted of 264 MZ pairs (129 male pairs and 135 female pairs) and 214 same-sex DZ pairs (121 male pairs and 93 female pairs). The mean age of the participants at time of testing was 11.5 years. All of the participants were administered an extensive battery of psychometric tests including the Wechsler Intelligence Scale for Children – Revised (WISC – R; Wechsler, 1974) or the Wechsler Adult Intelligence Scale – Revised (WAIS – R; Wechsler, 1981), the Peabody Individual Achievement Test (PIAT; Dunn & Markwardt, 1970), in addition to other tests of reading and language skills. When an extended DF multiple regression model was fitted to twin data, group heritability estimates (h^2_g) were significant (0.58, $\leq .001$), suggesting that more than 50% of the group's reading deficit was attributable to genetic influences. When data from twins were fitted to an extended DF model, the differential genetic etiology was non-significant ($p \geq .35$) suggesting little or no difference for males and females in the magnitude of genetic etiology of reading disability. Subsequent analyses failed to find differential heritability as a function of gender even in more severely impaired samples (Hawke J.L., Wadsworth, S.J., Olson, R.K., & DeFries, J.C. 2007).

Results from etiological studies are summarized in Table 5.1

Table 5.1

Summary of etiological studies of reading disability

Research ascertained studies of etiology			
Reference	**Sample**	**Measures**	**Findings for h^2_g**
Stevenson, 1992	102 MZ, 111 ssDZ, 72 osDZ pairs	WISC-R[1], Reading composite, Spelling composite	Males showed a higher prevalence rate No significant gender differences
Wadsworth et al., 2000	206 MZ, 159 ssDZ, 117 osDZ pairs	WISC-R[1], WAIS-R[2], PIAT[3]	No significant gender differences
Harlaar et al., 2005	392 twin pairs at 5th percentile 782 twin pairs at 10th percentile	TOWRE[4] Sight Word Efficiency Phonemic Decoding	Significant gender differences at or below the 5th percentile (higher h^2_g for males). Non-significant gender differences at the 10th percentile
Wadsworth et al., 2005	264 MZ, 214 ssDZ, 156 osDZ pairs	WISC-R[1], WAIS-R[2], PIAT[3]	No significant gender differences
Hawke et al., 2006	264 MZ, 214 ssDZ	WISC-R[1], WAIS-R[2], PIAT[3]	No significant gender differences

[1]Wechsler intelligence scale for children - revised
[2]Wechsler adult intelligence scale - revised
[3] Peabody individual achievement test
[4] Test of word reading efficiency

Whereas results from etiological studies of reading disability obtain similar ratios for both males and females, a number of studies suggest otherwise. Several possibilities have been suggested as to why males have a higher prevalence rate for reading disabilities. Subject ascertainment (i.e., research-identified subjects versus school-identified subjects), is one of the more common disparities in the range of findings between etiological studies versus prevalence studies. Results from studies suggest a potential gender ratio imbalance when subjects are school-identified because teachers are more likely to refer boys for having special problems as boys are perceived to be more disruptive than girls (Chan, Suk-han Ho, Suk-man, Suk-man, & Chung, 2007; Shaywitz et al., 1990). There is an assumption of more males than females affected with reading deficits with ratios ranging from 4:1 or 3:1. However, this gender imbalance has been called into question in some studies which suggest that the gender ratio is much closer and more likely to be 2:1 or even 1:1 (Shaywitz et al., 1990; Wadsworth et al., 1992).

In a case-control study by St. Sauver et al. (2001) data were obtained from children born between 1976 and 1982 in order to examine gender differences for reading disabilities as a function of risk factors (i.e., low birth weight, young maternal age, low maternal education, illegitimacy, and minority race or ethnicity). 303 reading disabled subjects were identified using IQ and achievement test scores collected from school and medical records. Controls consisted of the remainder of the subjects who did not meet criteria for RD (n = 4,529). Results from logistic regression models suggested that girls of low birth weight (< 2500 gms), were more than twice as likely to be identified as RD (Odds ratio, OR = 2.94, 95% CI: 1.09, 6.25). In addition, girls whose mothers obtained less than or equal to 12 years of education were also twice as likely to be identified as reading disabled (OR = 2.14, 95% CI: 1.24, 3.72). Fewer years of paternal education was found to be an indicator for reading disability in boys (OR = 2.28, 95% CI: 1.59,

3.27. The authors concluded that there was a differential susceptibility to RD risk factors and that biological processes leading to RD may differ between boys and girls.

Siegal and Smythe (2005), analyzed data from a sample of children (n = 984) seen longitudinally from kindergarten through 5th Grade. There were approximately equal numbers of girls and boys who were tested annually in order to compare ability longitudinally. In kindergarten, subjects were administered two reading tests; the Wide Range Achievement Test (WRAT; Wilkinson, 1993) and an experimental letter identification test that measured children's ability to name lowercase letters from the English alphabet. Grades 1 to 5 tests included two subtests from the Woodcock-Johnson achievement test (Woodcock & Johnson, 1977); Word and Letter Identification and Word Attack. In addition, in Grades 2 to 5, an experimental test of reading fluency which measured the number of words or pseudowords that could be read in one minute. There were significant mean differences in letter identification, word and pseudoword reading scores in kindergarten and Grade 1 with girls scoring higher overall. Except for word identification scores in Grade 4, with significantly higher mean scores for girls, gender differences disappeared following Grade 1. In addition to mean differences, the authors examined prevalence ratios from subjects identified as being reading disabled. Reading disability was identified as those subjects scoring either 1 or 2 standard deviations below the mean for each of the standardized reading tests described above. Results indicated that there were significantly more boys identified as reading disabled. However, findings did not suggest significant gender differences at the other grade levels. The authors suggested that these results may be a factor of developmental differences between boys and girls.

A study by Share and Silva (2003) examined data from a sample of children from the Dunedin Multidisciplinary Health and Development Research Unit (Silva & Stanton, 1996).

Data from this research-identified cohort were collected at ages 5 (n = 991), 7 (n = 954), 9 (n = 955), 11 (n = 925), and 13 (n = 850). Measures of IQ from the WISC-R and reading from the Burt Word Reading Test (Gilmore, Croft, & Reid, 1981) were assessed at age 11 years. Reading disability was identified using Rutter and Yule's (1975) regression method. Regression equations were computed for each gender separately (girls, n = 443, boys, n = 471) as well as a combined sample (n = 914). Children who scored more than 1.5 standard deviations below their predicted score were designated as having specific reading disabilities. Whereas the mean of the distribution of reading scores was higher for girls than for boys, there were no significant gender differences in the slopes t (910) < 1.0, or for the standard errors of estimates, t (910) = 1.34, p > .05. Interestingly, there was a gender difference for the intercepts, t (910) = 5.47, $p < .05$, with intercepts for boys being nearly 9 points lower. The authors suggested that this discrepancy in intercepts over-predicts male reading scores and under-predicts female reading scores by nearly 3 points. Moreover, this disparity may inflate the magnitude of the differences between the actual and the predicted reading scores for boys, as well as the opposite for girls. Findings from this study indicated that IQ distributions for means and variances were very similar for both boys and girls, suggesting that gender differences are a result of reading distributions. Share and Silva (2003) further suggest that children with reading disability may benefit from the use of separate definitions for boys and girls.

Rutter, Caspi, Fergusson, Horwood, Goodman, Maughan, Moffitt, Meltzer, & Carrol (2004) presented results from four epidemiological studies from English speaking countries, all of which showed a greater prevalence of reading disability in males than females. For each of the four studies, both IQ-achievement discrepancy and low achievement definitions of reading disability were utilized. The first study examined data from 989 subjects aged 7, 9, and 11 years

from the Dunedin Multidisciplinary Health and Development Study (Silva, 1990). The second study, from the Christchurch Health and Development Study (Fergusson, Horwood, Shannon, & Lawton, 1989) examined reading and IQ data from 895 individuals for ages 8, 9, and 10. The third study comprised data from 5752 subjects in a cross-sectional study in the United Kingdom, The Office for National Statistics Study (Meltzer et al., 2000). Tests included reading and IQ assessments at ages 9 to 15 years. The fourth study from the Environmental Risk Longitudinal Twin Study (Moffitt, 2002) included 2232 twins from England and Wales, who were given IQ tests at age 5 and reading tests at age 7. None of the four studies assessed writing or oral language skills. In addition, none of the studies used a measure of phonological decoding skills, which is affected in both children and adults with dyslexia (Berninger, 2006). Three of the studies classified subjects as being reading disabled if their reading scores were less than 15% of the distribution or if scores were 1 standard deviation or more below that expected based on their IQ. The fourth study from the Office of National Statistics Study, examined gender ratios in two severity groups. Subjects were classified as being in the lower 15% or those in the lower 5% of the distribution of scores utilizing both a non-IQ-referenced reading score, and an IQ-discrepant score. Results from the four studies of research-identified probands found a significantly higher number of reading disabled males than females. Odds ratios ranged from 1.39 to 3.19 for non-IQ-referenced RD and from 1.74 to 3.29 when an IQ-discrepant score was used to define RD. Based on their results, the authors suggested that reading disabilities are clearly more prevalent in boys than in girls.

In a recent study by Chan et al. (2007), data from 99 children between the ages of 6 and 10.5 years, were examined to study gender ratio and gender differences in reading-related cognitive abilities. The subjects were identified as dyslexic based on diagnostic criteria as

specified in the Manual of Hong Kong Test of Specific Learning Difficulties in Reading and Writing (HKT-SpLD: Ho et al., 2000a). Children were arbitrarily divided into three groups (mild, moderate and severe) using their literacy composite scores with cutoff scores at gradations of 0.5 SD below – 1 SD. There were 43 males and 22 females resulting in a gender ratio of 2:1 which was a significant difference (x^2 = 8.02, p < .05). However, there was not a significant difference between gender and severity, suggesting that there was no evidence for a differential gender bias in mild, moderate, or severe cases. Results from this study indicated an overabundance of males compared to females. The imbalance could not readily be explained by referral or selection bias as an equal number of males and females were initially tested.

Dirks et al. (2008), presented data from 4th and 5th Graders. The sample included 799 children with data from both arithmetic and reading (word recognition), 622 of these children had reading comprehension data available and 689 of these children had spelling measures completed. The full sample was assigned to four different groups; 1) RAD – combined reading and arithmetic disability, those who performed at or below the 25th percentile for both word recognition and mathematical computation, 2) RD – reading-disability only, participants who performed below or at the 25th percentile on word recognition and above the 25th percentile for mathematical computation, 3) AD – arithmetic-only disability, those subjects who performed at or below the 25th percentile on mathematical computation and above the 25th percentile on word recognition, and 4) NA – normal achievement, those who performed above the 25th percentile on both word recognition and mathematical computation. Ratios for each group, boy-to-girl was as follows: 0.6:1 for the RAD group, 1:1 for the RD group, 0.5:1 for the AD group, and 1.1:1 for the NA group. The pairwise comparison between the AD group and the NA group was

significant ($p = .007$). Findings indicated that arithmetic difficulties occur more frequently in girls than in boys, however, there were no significant gender differences for reading disabilities.

Results from prevalence and mean difference studies with varying findings are summarized in Table 5.2.

Table 5.2

Prevalence and mean difference results for RD in School and Research identified studies

School ascertained subjects			
Reference	**Sample**	**Measures**	**Findings**
St. Sauver et al., 2001	303RD/4529 Control	IQ discrepancy/achievement test scores	Females of low BW showed an increased risk for RD over females of normal BW. Males were twice as likely to be diagnosed w/RD than females of normal BW
Siegal & Smythe, 2005	984 children	WRAT[2]/Word & Letter Identification/Word Attack[3]	More boys than girls with lower reading achievement in Kindergarten
Research identified subjects			
Share & Silva, 2003	443 girls/471boys	WISC-R[4]/BWRT[1]	No prevalence differences
Rutter et al., 2004	4 epidemiological studies	IQ discrepancy/BWRT[1]	RD more prevalent in males
Chan et al., 2007	43 boys/22 girls	HKT-SpLD[5]	A significant ratio difference (2:1), but, no significant gender differences as a function of severity of RD
Dirks et al., 2008	418 girls/381 boys	Cito Leerlingvolgsysteem[6]	No difference in prevalence for boys and girls for RD Gender differences in math with more girls having arithmetic difficulties than boys

[1] Burt Word Reading Test (Scottish Council for Research in Education, 1976)
[2] Wide Range Achievement Test
[3] Subtests from the Woodcock-Johnson test
[4] Wechsler intelligence scale for children - Revised
[5] Hong Kong Test of Specific Learning Difficulties in Reading and Writing
[6] A monitoring system that consists of national norm-referenced achievement tests

As discussed, previous studies have suggested that there is a preponderance of males affected with reading disabilities. Results from some studies suggest that differences in development for males and for females may be a factor for this discrepancy. Therefore, reading disability and stability may be confounded by developmental processes specific to gender.

In an early study by Feldman et al. (1995), reading disability was examined with a three generation family history. Thirty-seven adults (average age 40 years) were administered a battery of psychometric tests including the WAIS-R (Wechsler 1981), the Woodcock-Johnson Psycho-educational Battery (Woodcock & Johnson, 1977) and additional tests of reading comprehension and spelling. Subjects were compared to a control sample of forty-eight adults (average age 44 years). Results indicated those adults with a strong history of RD continued to display reading and spelling problems into their adult years. These deficits were evident despite average to above average IQ. In addition, the data did not provide evidence for gender differences. However, females exhibited less marked reading impairment than their male counterparts. This study suggested that female dyslexics develop more compensatory skills to assist with their reading disability. Although results did not find significant gender differences over time, female group means trended towards a less impaired severity score for RD suggesting different patterns of reading skills for males and females.

A longitudinal study by Kush et al. (1996), examined the stability of children's attitudes towards reading. Participants comprised 190 students in Grades 1 through 4 (83 males, 107 females) over a three year period. Students were given the Elementary Reading Attitude Survey (ERAS; McKenna & Kear, 1990), a 20-item questionnaire which presents a short simple worded statement about reading based on two component subscales: recreational reading attitude and academic reading attitude. Recreational items focus on reading for fun outside of school,

whereas Academic items focus on workbooks, worksheets and schoolbooks. Results overall showed a downward trend towards reading attitudes in general. This decline occurred for attitudes regardless of recreational or academic reading variables across the 3 years. However, females consistently demonstrated a more positive reading attitude than did males. The long-term stability coefficient for the ERAS scale, for the three-year period, was .26, $p > .01$, indicating that females had consistently greater stability than males for reading attitudes. Results from this study suggest that males have more negative attitudes about reading and as a result are more frequently identified as problem readers when compared to their female counterparts. In addition, males as a group demonstrate lower reading achievement and are more frequently placed into learning disability programs when compared to females.

Badian (1999) examined whether longitudinal stability, gender differences, and prevalence is a result of disparities between tests of listening and reading comprehension rather than IQ-achievement test discrepancy definitions. Participants comprised 1008 students (506 boys and 502 girls) followed from prekindergarten through Grade 7-8, approximately 13 years. All subjects were tested with a preschool screening battery prior to kindergarten (Badian, 1990) in addition to two subtests of the Wechsler Preschool and Primary Scale of Intelligence (WPPSI; Wechsler, 1990). Also, measures of reading and verbal skills were obtained from the Reading Comprehension subtest of the Stanford Achievement Test (SAT; Gardner, Rudman, Karlsen, & Merwin, 1982). Results suggested that for students with consistent reading disability, the boy-to-girl ratio was 3.2:1. This result was higher when compared to the nondiscrepant poor readers in both the lower and upper grades, 1.3:1. However, Badian (1990) noted that this higher gender ratio for reading disabled subjects may be a reflection of the fact that girls were significantly better in reading comprehension at each grade level. Additionally, relying on the listening

comprehension/reading comprehension discrepancy resulted in a decrease in the overall stability of reading disabilities. This contrasts studies which rely on IQ and reading tests to define reading disability.

A study by Bornstein, Hahn, & Haynes (2004) examined issues of stability and gender for general language performance across early childhood. The sample comprised 329 children ranging in ages of 1 year to 7 years, who participated in four independent longitudinal studies of specific and general language performance. Data was based on age-appropriate maternal questionnaires and interviews, teacher reports, and transcripts of children's own spontaneous speech. Formal assessments included the MacArthur Communicative Development Inventory: Words and Sentences (CDI-WS; Fenson, Dale, Reznick, Thal, Bates, Hartung, Pethick, & Reilly, 1993), two scales from the Reynell Developmental Language Scales, Verbal Comprehension Scale and the Expressive Language Scale (RDLS; Reynell & Gruber, 1990), and the Wechsler Preschool and Primary Scale of Intelligence – Revised (WPPSI-R; Wechsler, 1990). Results for early years, prior to attending school, indicate moderate to strong stability for individual differences for both boys and girls. Additionally, for the second through fifth years, girls consistently out-performed boys in multiple specific and general measures of language. Findings for stability provide information regarding the developmental progression of a given psychological function. Language development is thought to be an important predictor of developmental status at a given age (Lerner & Busch-Rossnagal, 1981), and because girls typically mature faster than boys, we would expect some language abilities to occur earlier in girls than in boys.

In a study by Kazelskis, Thames, Reeves, Flynn, Taylor, Beard, & Turnbo (2005), the stability of reading attitude was examined across gender. Participants comprised 718 students

from grades four through six. There were 374 boys and 344 girls representing 52% and 48% of the total sample, respectively. Each subject completed the Elementary Reading Attitude Survey (McKenna & Kear, 1990). The survey was repeated with a 7-day interval between administrations. Stability results were similar to that of Kush et al. (1996) although somewhat higher. However, these results are not surprising given the length of assessment intervals between the two studies (3 years versus 7 days). The mean attitude scores trended higher for girls than for boys, although the difference was not significant. The authors argued that while stable attitudes for reading were evident, more frequent assessments by classroom teachers and reading specialists would help avoid short-term fluctuations in reading attitude and in effect help to avoid long-term reading issues (i.e., poor reading related to negative attitudes towards reading).

A study by Hulslander et al. (2010), examined the prediction of reading development based on longitudinal stability of reading-related skills from an older cohort (average age at follow-up testing was 15.8 years). Subjects (n = 329) participated in the LTSRD (Wadsworth et al., 2007) and were originally tested in the CLDRC (DeFries et al., 1997). Participants were administered a series of reading and related cognitive tests, including the PIAT-R, WRAT-3, and the WISC-III/WAIS-III. Results indicated significant stability for individual differences in word comprehension and spelling across later grades with stability correlations of .98 and .95 respectively. These findings are consistent with strong genetic influences on individual differences for reading and related skills that are evident as early as the end of first grade (Bryne et al., 2007; Harlaar et al., 2005). However, in contrast, results did not indicate longitudinal prediction for reading disabilities from children's reading-related skills.

Results from stability studies are summarized in table 5.3.

Table 5.3

Summary of results from Stability Studies

Reference	Sample	Measures	Findings
Feldman et al., 1995	37 adult dyslexics w/ childhood history of RD 48 controls	WAIS-R[1], Woodcock-Johnson Psycho-educational Battery	No prevalence differences, reading disabilities were highly stable. Females developed more compensatory skills over time
Kush et al., 1996	83 males/107 females	ERAS[2]	Females consistently had more positive attitudes towards reading over time
Badian 1999	506 males/502 females	Reading /listeninComprehension –SAT[3] WPPSI[4]	Gender differences and stability for RD was low
Bornstein, et al., 2004	Population study	Questionnaires and language comprehension tests	No differences for stability with regard to gender, but girls were significantly higher in language performance over boys
Kazelskis, et al., 2005	718 students	ERAS[2]	Females scored higher with more positive attitudes. Stability decreased at higher grade levels
Hulslander et al., 2010	324 children (114 with RD)	PIAT[5]/WRAT[6]/WISC[7]/WAIS[1] and phonological tests	Substantial stability. Reading related skills that are independent from word-reading & spelling skills do not influence development of these skills over time.

[1]Wechsler Adult Intelligence Scale – revised
[2]ERAS
[3]Stanford Achievement Test
[4]Wechsler preschool and primary scale of intelligence
[5]Peabody individual achievement test
[6]Wide range achievement test
[7]Wechsler intelligence scale for children

Given the propensity for disparate results in studies of reading disabilities, it is important to consider factors that may affect the stability of reading deficits and whether those factors vary for boys or for girls. To our knowledge no one has examined the differential etiology of stability for reading disabilities as a function of gender. Thus, the purpose of this study is to investigate whether the heritability of deficits in reading may be more or less stable depending on whether reading disabled subjects are male or female. Considering the plethora of mixed results in studies with regard to gender, stability of heritability may be a contributing factor to the diverse results.

5.2 Methods

5.2.1 Participants and measures

The subjects in the present study were first tested in the ongoing Colorado Learning Disabilities Research Center (CLDRC, DeFries et al., 1997) between September 1996 and March 2003 and also participated in follow-up testing approximately 5 years later in the Longitudinal Twin Study of Reading Disability (LTSRD; Wadsworth et al., 2006). Twin pairs were initially identified by personnel in 27 school districts throughout Colorado and in order to minimize the possibility of ascertainment bias, twins were identified without knowledge of reading status. Parental permission is then sought to review the school records of the twins for any evidence of reading problems. If either member of a twin pair displays a history of reading problems, both twins and their siblings were invited to participate in the study at the University of Colorado, Boulder, and at the University of Colorado, Denver. All subjects completed an extensive battery of psychometric tests, including measures of reading, language and perceptual processes, in addition to the PIAT (Dunn & Markwardt, 1970), and the Wechsler Intelligence Scale for Children – Revised (WISC-R; Wechsler, 1974), or the Wechsler Adult Intelligence Scale – Revised (WAIS-R; Wechsler, 1981). A comparison of sample control twin pairs was also tested

in which neither member of the pair had a school history of reading problems. Control twins were matched to the probands by age, gender, and school district.

For the present analysis a composite measure of reading performance (DISCR) was computed for each subject employing discriminant weights estimated from an analysis of data from the Reading Recognition, Reading Comprehension, and Spelling subtests of the PIAT obtained from an independent sample of 140 non-twin children with reading difficulties and 140 non-twin children without reading problems (DeFries 1985). A diagnosis of reading-disabled for the present study requires than an individual meet the following criteria; a positive history of reading problems and a score at least one standard deviation below the mean of the control sample. In addition, a minimum IQ score of 85 on either the Verbal or Performance Scale of the WISC-R (Wechsler 1974) or the WAIS-R (Wechsler 1981); no evidence of neurological problems or severe behavioral or emotional problems. In addition, subjects may have no uncorrected visual or auditory acuity deficits. Zygosity of the twin pairs was determined using selected items from the Nichols and Bilbro (1996) zygosity questionnaire, which has a reported accuracy of 95% for same-sex twin pairs. When cases were questionable, blood or buccal samples were obtained and twin pairs were genotyped using polymorphic DNA markers.

Participants from the initial time-point included 303 identical and 267 same-sex fraternal twin pairs in which at least 1 member of each pair had reading difficulties. For the second time-point analysis, the number of MZ, and same-sex DZ twin pairs on whom follow-up data were available is quite small and included 33 identical and 37 fraternal twin pairs. Subjects included 34 male pairs (14 MZ and 20 DZ) and 36 female pairs (19 MZ and 17 DZ) meeting criteria for inclusion in the initial proband sample, a ratio of 1.1:1. Participants were reared in mainly

English-speaking, middle class homes, and ages ranged between 8.0 and 20.2 years (average age 11.42 years).

5.3. Analyses

5.3.1 Multiple Regression Analysis

The differential regression of the MZ and DZ co-twin means toward the mean of the unselected population provides a powerful test of genetic etiology (DeFries & Fulker, 1985). As MZ twins are genetically identical and DZ twins share only half of their segregating genes on average, the scores of the DZ co-twins would be expected to regress more toward the mean of the unselected population if the variable of interest is heritable. Fitting the following multiple regression model to selected twin data facilitates a more powerful and versatile test (DeFries & Fulker, 1985, 1988):

$$\hat{C}_x = B_1 P_x + B_2 R + A \qquad [5.1]$$

where C is the predicted co-twin's score, P is the proband's score, R is the coefficient of relationship (1.0 for MZ pairs and 0.5 for DZ pairs), B_1 and B_2 are the partial regression coefficients, and A is the regression constant. B_1 provides a measure of twin resemblance for reading performance that is independent of zygosity. B_2 equals twice the difference between the MZ and DZ co-twin means after covariance adjustment for differences between MZ and DZ proband means. Therefore, B_2 provides a test of statistical significance for genetic etiology. Moreover, when the data have been suitably transformed prior to multiple regression analysis (i.e. each score is expressed as a deviation from the mean of the control sample and then divided by the difference between the proband and control means), B_2 directly estimates h^2_g. Equation 5.1 was initially fitted to transformed discriminant function data (DISCR) for all twin pairs to assess the genetic etiology of the entire sample. Twin pairs were then grouped by gender and

Equation 5.1 was fitted to the data for each group, providing separate estimates of h^2_g for male and female twin pairs. Then in order to test for differential genetic etiology as a function of gender, the following extended model was fitted to data from the entire sample:

$$\hat{C}_x = B_1P_x + B_2R + B_3S + B_4PS + B_5RS + A \quad [5.2]$$

where *S* is the sex of the twin pair (coded 0.5 for male pairs and -0.5 for female pairs), *PS* is the product of the proband's DISCR score and sex, and *RS* is the product of the coefficient of relationship and sex. When sex is coded this way, B_5 equals the difference between the h^2_g estimate obtained for male twin pairs and female twin pairs separately. Therefore, B_5 provides a relatively powerful test of statistical significance for the differential etiology as a function of gender.

Because subjects were not reselected at the follow-up time-point, univariate DF models were fitted only to data from the initial time-point. In order to assess the heritable characteristics of the *stability* of reading deficits, data from the initial and follow-up time-points were fitted to a bivariate extension of the basic DF model. The bivariate model is as follows:

$$\hat{C}_y = B_1P_x + B_2R + A \quad [5.3]$$

where C_y is the co-twin's composite reading score at follow-up, P_x is the proband's initial composite reading discriminant score, *R* is the coefficient of relationship, and *A* is the regression constant. B_1 is the partial regression of the co-twin's follow-up reading score on the proband's initial reading score and is a measure of average MZ and DZ cross-variable twin resemblance. Therefore, B_1 estimates the extent to which co-twin scores on the follow-up measure are related to proband scores on the initial measure across zygosity. B_2 is the partial regression of the co-twin's follow-up reading score on the coefficient of relationship. When the data are transformed prior to multiple regression analysis, the bivariate B_2 coefficient is a function of the square roots

of the group heritabilities for reading performance at the two time-points and the genetic correlation (r_G) between them (i.e., $h_{initial}$ x $h_{follow\text{-}up}$ x r_G; Light and DeFries, 1995). As follows, B_2 provides an estimate of "bivariate heritability" ($h^2_{g(Biv)}$), a measure of the extent to which the proband reading deficit at follow-up is due to heritable factors which also influence the reading deficit at the initial time point. Moreover, the proportion of the phenotypic stability correlation (r_p) attributable to genetic influences can be obtained by dividing the B_2 estimate by r_p. Finally, when the following extended bivariate model is fitted to twin data, Equation 5.4, B_5 provides a test of the significance for differential etiology of the *stability* of reading differences as a function of gender:

$$\hat{C}_y = B_1P_x + B_2R + B_3S + B_4P_xS + B_5R_xS + A \qquad [5.4]$$

Because truncate selection was employed (DeFries and Gillis 1991), pairs in which both members met criteria for RD were double-entered for all regression analyses. This is analogous to the computation of probandwise concordance rates, in which both affected members of concordant pairs are included as probands. Standard error estimates and tests of significance were adjusted accordingly. Models were fitted using linear regression in SPSS version 17.0 (SPSS 2007).

5.4 Results

The ratio of males to females in the total sample was 1.1:1, a non-significant deviation from 1:1 (x^2 = 0.333, df = 1, p = 0.56).Table 1 presents the numbers of concordant and discordant pairs of MZ and DZ twins and their probandwise concordance rates assessed at the initial time-point. The sizeable differences between the MZ and DZ concordance rates for both male and female twin pairs indicates that reading difficulties are due substantially to genetic

influences in both sexes. Concordance rates for males and females regardless of zygosity are similar (see Table 5.4).

Table 5.4

Concordance for reading disability in MZ and same-sex DZ twin pairs in the initial time-point study

	Male		Female	
	Identical	Fraternal	Identical	Fraternal
Initial time-point				
Pairs concordant	67	28	65	18
Pairs discordant	67	106	74	88
Probandwise concordance (%)	67	35	64	29

Table 5.5 presents the standardized mean reading performance scores for MZ and DZ probands for each gender. The MZ and DZ proband means are highly similar at both initial and follow-up time-points and are approximately 2 standard deviations below the control mean for each assessment which indicates that the deficit for probands is substantially stable.

Table 5.5

Mean standardized reading performance scores (± SD) of probands and co-twins at intial[a] and follow-up[b] test sessions.

	MZ pairs				DZ pairs			
	Male		Female		Male		Female	
	M	SD	M	SD	M	SD	M	SD
Initial time-point								
Proband	-2.406	0.848	-2.120	0.754	-2.288	0.896	-2.181	0.848
Co-twin	-2.213	1.030	-1.952	0.897	-1.414	1.328	-1.229	1.121
Follow-up time-point								
Proband	-2.075	0.852	-1.664	0.644	-2.027	0.792	-1.728	0.845
Co-twin	-1.712	1.267	-1.542	0.638	-1.277	1.525	-.997	1.061

[a] Standardized against the mean of 1456 Control twins participating in the initial assessment of the CLDRC (not shown)

[b] Standardized against the mean of 171 Control twins participating in the LTSRD at follow-up (not shown).

As shown, the MZ co-twin mean for males in the initial assessment, regressed .19 standard deviations units toward the control mean on average, whereas, the DZ co-twin for males regressed 0.87 standard deviation units. Similarly for females, the MZ co-twin mean regressed .17 standard deviation units and for the DZ co-twin, 0.95 standard deviation units on average. The assessment from the follow-up study had comparable results. The MZ co-twin mean for males regressed .36 standard deviation units and for the DZ co-twin, results were .75 standard deviation units. Females followed with similar results, as female MZ co-twins regressed .12 standard deviation units and the female DZ co-twins, .73 standard deviation units. Figure 5.1 presents the corresponding transformed proband and co-twin means.

Figure 5.1.Transformed proband and co-twin means

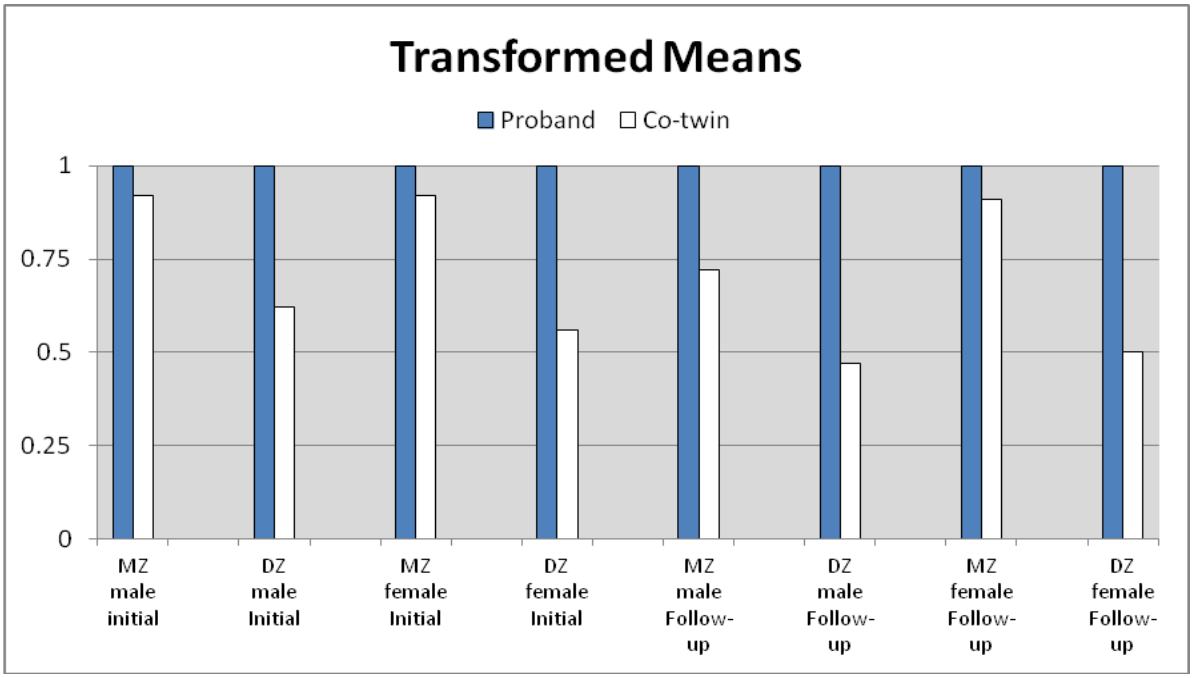


Table 5.6 presents stability correlations between males and females at initial assessment and follow-up assessment based on the discriminant function score (DISCR) and suggest that stability between the two time-points for both genders is substantial.

Table 5.6

Stability correlations

Measure	Males	Females
DISCR	0.89, $p \leq .001$	0.84, $p \leq .001$

Basic univariate regression analyses of twin data are presented in Table 5.7. Equation 5.1 was fitted to transformed discriminant function data of all twin pairs simultaneously. The resulting estimate for B_2 was .65 (±.07, $p \leq .001$), suggesting that two-thirds of the proband's reading deficit is attributable to genetic influences. Equation 1 was then fitted to the DISCR data of the twin pairs separately by gender. Although the estimate of B_2 for females is somewhat larger than for males (.71 vs. .60 respectively), when the entire twin pair data were fitted simultaneously to Equation 5.2, the B_5 test for differential etiology was not significant ($p = .424$).

Table 5.7

Results of univariate and bivariate DF analysis

Model	Subjects	Model	B_2 ± S.E.	p	B_5 ± S.E.	p
Univariate	Twins-all	$C = B_1P + B_2R + A$	.65 ± .07	$p \leq .001$		
	Male	$C = B_1P + B_2R + A$	.60 ± .10	$p \leq .001$		
	Female	$C = B_1P + B_2R + A$	.71 ± .10	$p \leq .001$		
Extended	Twins-all	$C = B_1P + B_2R + B_3S + B_4PS + B_5RS + A$	------	------	.11 ± .14	$p = .424$
Bivariate	Twins-all	$C_y = B_1P_x + B_2R + A$	.64 ± .22	$p = .006$		
	Male	$C_y = B_1P_x + B_2R + A$	.55 ± .34	$p = .113$		
	Female	$C_y = B_1P_x + B_2R + A$	.83 ± .27	$p = .005$		
Extended	Twins-all	$C_y = B_1P_x + B_2R + B_3S + B_4P_xS + B_5R_xS + A$	------	------	.28 ± .44	$p = .532$

When the bivariate model, Equation 4.3, was fitted to the transformed proband initial scores and the co-twin follow-up scores for males and females, the estimate of bivariate h^2_g was .64 (± .22, p = .006), suggesting that more than 60% of the proband deficit in reading at follow-up is due to genetic factors that also influenced reading difficulties at the initial assessment. When the extended model, Equation 4.4, was fitted to the bivariate data, the resulting B_5 estimate was .28 (± .44) and non-significant (p = .532).

5.5 Discussion

The primary goals of the present study were to assess the genetic and environmental influences on reading disability for males and females. Further, to employ an extension of the DeFries-Fulker multiple regression model, to examine whether or not the stability of heritability for reading deficits varies as function of gender.

5.6. Summary

Reading difficulties account for a significant proportion of learning disabilities, with as much as 90% of all diagnosed cases of LD. A difficult job has been to untangle the genetic and environmental factors that contribute to reading disability. One aspect has been the issue of gender. The issue of gender has remained controversial particularly due to subject ascertainment and reading disability definition and varying assessments. For example, school-identified subjects are more likely to be boys as boys are more often perceived as more disruptive in a school when compared to girls. However, most etiological studies have found no significant mean differences in reading deficits between boys and girls.

5.6.1 Basic and extended DF model

In the current study, results of fitting the DF basic model to reading data from twins in the CLDRC, heritability estimates were similar to those of previous studies, (.67 ± .07, $p \leq .001$),

with more than two-thirds of the proband deficit accounted for by genetic influences. Although results for female heritability were higher than for males (.71 vs. .60), this difference was not significant (p = .022).

5.6.2 Bivariate DF models

When data are transformed prior to multiple regression analysis, B_2 is a function of the square roots of the group heritabilities between the two time points and the genetic correlation between them. Therefore, B_2 provides an estimate of bivariate heritability. When the data were fitted to a bivariate extension of the DF model, resulting estimates for bivariate heritability of the entire twin sample were significant (.64 ± .22, p = .006), indicating that deficits for the probands at the follow-up assessment were likely influenced by the same genes at the initial assessment, thereby suggesting that reading deficits are highly stable. When equation 4 was fitted to twin data, B_5, a test of the significance for differential etiology of the stability of reading deficits as a function of gender, the results were non-significant (p = .532).

Whereas these results are similar to our first study examining the genetic etiology of the stability of reading disabilities (Astrom et al., 2007), our current results do not provide evidence for a differential heritability for the stability of reading deficits as a function of gender. There are several limitations to this study however. First, our sample size is quite small. However, even with the small sample we obtained significant results for the heritability of stability in females in the bivariate model (.83 ± .27, p = .005) and although our results for males were non-significant for equation 5.3 when applied to data from only males, the p value was relatively low (p = .113). Therefore, with an increased sample size we may find significant results for males as well as females. Further, examining a differential heritability for the stability of reading difficulties may be reflected in varying age groups such that a younger cohort may be more genetically

influenced than an older cohort. This difference was suggested by Harlaar et al. (2005). Further affecting our results is the issue of attrition. Because we do not re-select our subjects for follow-up analyses, close scrutiny of our current data revealed that only a small proportion of our original subjects from the initial assessment were included in this analysis. It is possible that if we had the same subjects at both time-points we may see an increased affect for a differential heritability of stability as a function of gender. Whereas DF analysis is very powerful for obtaining results in selected samples, establishing whether the affects of gender are being mediated developmentally warrants additional testing with more subjects.

Chapter 6

Summary and Conclusions

6.1 Introduction

Reading disability is a complex and serious disorder affecting both adults and children. Most often reading disability is diagnosed in part by difficulties with single-word reading and spelling (Lyon et al., 2003; Pennington, 2009). It is often described as an unexpected problem with learning to recognize printed words that is not due to general intellectual impairments, sensory impairments, or lack of an opportunity to learn from appropriate instruction. The current definition assumes that reading difficulties are biological in origin. Originally thought to be a deficit in the visual system, the present thought is that reading difficulties are a result of a deficit in a specific language skill that is responsible for the processing of phonological information. Most estimates of prevalence for reading disability in school-age children range between 7% and 15%. In addition, negative correlates of the disorder extend well into adulthood.

Reading disability was first recognized in the late 19th century and was termed "congenital word-blindness". Often thought to be a result of low intelligence, physical disabilities or less than adequate education, it became readily apparent that congenital word-blindness had a heritable component (Thomas, 1905).

Multiple studies using multiple methods have examined the genetic and environmental etiologies of reading disabilities. Behavioral Genetics is one such approach that investigates the relative contributions that genes and environment make in explaining differences in behavior. An especially powerful method in behavioral genetics is the study of twins. A comparison of scores of MZ twins who share all their alleles IBD to those of DZ twins who share half their alleles on average provides an indication of the extent to which a disorder is due to genetic or

environmental influences. A straightforward test for genetic influences on a particular disorder is a comparison of concordance rates. However, DeFries and Fulker (1985, 1988) proposed that a differential regression of MZ and DZ co-twin means toward the unselected population provides a more appropriate test of genetic etiology.

The purpose of the four projects presented here was to more fully understand the genetic and environmental etiologies of reading disabilities. Furthermore, to examine the effects of "special twin environments" and to investigate potential confounds which contribute to the varying results specific to gender studies of reading disabilities. Results and plausible implications for the four studies are discussed below.

6.2 Summary of Results

6.2.1 Chapter 2 – Genetic and environmental etiologies of reading difficulties: DeFries-Fulker analysis of reading performance data from twin pairs and their nontwin siblings

In chapter 2, reading performance data from 254 pairs of identical (MZ) and 420 pairs of fraternal (DZ) twins, 8.0 to 20.0 years of age, were subjected to multiple regression analyses. An extension of the DeFries-Fulker (DF) analysis (DeFries & Fulker, 1985, 1988) that facilitated inclusion of data from 303 of their nontwin siblings was employed. In addition to providing estimates of heritability, this analysis yields a test of "special twin environments", a measure of the difference between the regression of the DZ co-twin and co-sib means (Astrom et al., 2011). Our results indicate that proband reading deficits are largely due to genetic factors (.67 ± .07, $p <$ .001), and shared environmental influences are higher for members of twin pairs than for those of twins and their nontwin siblings (viz., .25 versus .17, $p = .02$).

Similar to results of previous studies, group heritability (h^2_g) was estimated at .67 and suggests that approximately two-thirds of the proband reading deficit is due to heritable

influences. In addition, our measure of special twin environment was significant (p = .02). It's important to note that while results of this study suggest that reading deficits are substantially heritable, it should not negate the importance of early diagnoses of reading problems and early intervention where warranted.

6.2.2 Chapter 3 – Etiology of the stability of reading difficulties: The Longitudinal Twin Study of Reading Disabilities

In chapter 3 – we introduced the first longitudinal twin study of reading difficulties to provide an initial assessment of genetic and environmental influences on the stability of reading deficits. Data were analyzed from a small selected sample of 56 twin pairs, 18 MZ and 38 DZ, in which at least one member of each pair was classified as reading-disabled in the Colorado Learning Disabilities Research Center (DeFries et al., 1997). In addition to data from the initial assessment, follow-up data were available. Twins were tested at two time points (average age of 10.3 years at initial assessment and 16.1 years at follow-up). A composite discriminant function score (DISCR) was highly stable, with a stability correlation of .84.

Data from the initial time point were first subjected to univariate DeFries-Fulker multiple regression analysis and the resulting estimate of the heritability of the group deficit (h^2_g) was .84 (± .26). When the initial and follow-up data were then fitted to a bivariate extension of the basic DF model, bivariate heritability was estimated at .65, indicating that common genetic influences account for approximately 75% of the stability between reading measures at the two time points.

The purpose of this first longitudinal twin study of reading disability was to investigate the etiology of stability using data from two time points. The reading deficit of the probands was highly stable as indicated by the means of the probands in that they were more than two standard deviations below the control means at both initial and follow-up time-points. Importantly, results

indicate that this deficit is substantially due to genetic influences. It should be noted that our subjects who met criteria for inclusion in the proband sample and on whom follow-up data was available is quite small. However, results of this analysis were similar to those obtained by other previous longitudinal studies (Harlaar et al., 2005; Wadsworth et al., 2006). Although reading disability is shown to be highly heritable and stable, early intervention remains an important component to reading success for disabled readers. Thus, additional longitudinal studies of heritability and stability of reading deficits with larger samples are clearly warranted.

6.2.3 Chapter 4 – DeFries-Fulker analysis of longitudinal reading performance data from twin pairs ascertained for reading difficulties and their nontwin siblings

Our previous study examined the etiology of the heritability and stability of reading deficits. The current study examined data from a larger sample of twins and their co-twins. In addition, in an effort to increase power, and to test for special twin environment, data from co-sibs were included. Data were subjected to an extension of DeFries-Fulker analysis (DeFries & Fulker, 1985; 1988). As well as providing estimates of univariate and bivariate heritability, this analysis facilitates a test of the difference between shared environmental influences for twins versus siblings. Discriminant function scores (DISCR) at 10.6 and 15.5 years of age, on average, were analyzed from 33 MZ and 64 DZ twin pairs in which at least one member of each pair had reading difficulties, and from 44 siblings of the probands. Scores were highly stable (.86 ± .03, across probands, cotwins and siblings) and heritability of the group deficit at initial assessment was .67 ± .22. Longitudinal bivariate heritability was .59 ± .21, suggesting that nearly 60% of the proband reading deficit at follow-up is due to genetic factors that influenced reading difficulties at the initial assessment. However, tests for special twin environmental influences were nonsignificant.

These results our highly similar to our previous results (Astrom et al., 2007) which should be expected as there is considerable overlap between the two samples. However, previous studies of individual differences in reading performance also support these findings. The extended DF model analyzes data from both twins and their siblings and provides a test for "special twin environments". For these results, B_3 was .14 and although it was non-significant it was not trivial. While these findings further indicate that reading deficits are highly heritable and stable, our best efforts with regard to intervention and remediation nonetheless, are extremely important. There are multiple genetic and environmental factors influencing reading abilities and these areas are not well understood. Therefore, in order to identify and treat at risk children more effectively and account for their individual abilities and disabilities, additional studies of genetic and environmental etiologies are necessary.

6.2.4 Chapter 5 – Gender Differences in the heritability and stability of reading difficulties: A twin study from the Colorado Learning Disabilities Research Center

A significant proportion of all learning disabilities are attributable to reading deficits. Although reading disability has been recognized as heritable since the early 1900's, factors which influence this disorder are complex. Multiple studies have examined genetic and environmental influences with widely different results. The issue of gender differences for reading disability is controversial. Therefore, chapter 5 took a different approach to the issue of gender; is it possible that given the various results for gender studies, that differential heritability of stability is a factor? In other words, do genes impact males and females differently at different stages in their development? And if so, would that contribute to varying results in reading abilities?

Data were subjected to a bivariate extension of the DF multiple regression model. Results were similar to our previous studies and reading disability was found to be both heritable and stable (h^2_g = .58, p = .013). When the twin pair data were fitted to an extension of the DF model to test for differential heritability and stability, the results were non-significant, indicating that for these data there were no gender differences. However, it's important to note that our sample is quite small. Nevertheless, given outcomes of developmental studies for reading disability it remains plausible that genes affect boys and girls differently at during various stages of their development. As such, more rigorous studies with larger samples to investigate these possibilities are needed.

6.3 Conclusions

Taken together, these studies have addressed various issues that are important for understanding the etiology of reading disabilities. First, findings indicate that reading disabilities are substantially heritable. Second, that reading disabilities are both heritable and highly stable. In addition, reading disability is a complex disorder to study in part due to complicated ascertainment issues. Contemporary definitions of RD have evolved via research and presently there is a good understanding of how to define dyslexia, in addition to the neuropsychological and behavioral correlates. However, much of this understanding has not reached the public and children continue to be misdiagnosed and receive treatments that have little efficacy and deviate from the scientific understanding of the disorder. Nonetheless there are models of prevention and remediation through multitiered approaches which include 1) universal screening for reading, 2) monitoring progress of at-risk children and 3) providing intense intervention based on the child's progress. Therefore, it's extremely important that there is continued research in the field of

behavior genetics to continue to investigate the multitude of variables which affect such a fundamental aspect of our everyday lives.

References

Alarcón, M., DeFries J.C., & Fulker, D.W. (1995). Etiology of individual difference in reading performance: A test of sex limitation, *Behavior Genetics, 25,* 17-23.

Astrom, R., Wadsworth, S.J., & DeFries, J.C. (2007). Etiology of the stability of reading difficulties: The Longitudinal Twin Study of Reading Disabilities. *Twin Res. & Hum. Gen., 10,* 434-439.

Astrom, R.L., Wadsworth, S.J., Olson, R.K., Willcutt, E.G., & DeFries, J.C. (2011). DeFries-Fulker analysis of longitudinal reading performance data from twin pairs ascertained for reading difficulties and from their nontwin siblings. *Behavior Genetics, 41,* 660-667.

Astrom, R.L., Wadsworth, S.J., Olson, R.K., Willcutt, E.G., & DeFries, J.C. (manuscript submitted for review). Genetic and environmental etiologies of reading difficulties: DeFries-Fulker analysis of reading performance data from twin pairs and their nontwin siblings.

Badian, N.A. (1990). Background factors and preschool test scores as predictors of reading: A nine-year longitudinal study. *Reading and Writing: An Interdisciplinary Journal, 2,* 11-30.

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.

Bakwin, H. (1973). Reading disability in twins. *Dev. Med. Child Neurol.15,* 184-187.

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.

Beitchman, J.H., Wilson, B., Brownlie, E.B., Walters, H., Inglis, A., & Lancee, W. (1996). Long-term consistency in speech/language profiles: II. Behavioral, emotional, and social outcomes. *Journal of the American Academy of Child & Adolescent Psychiatry, 35,* 815-825.

Benton, A.L., & Pearl, D. (Eds.). (1978). *Dyslexia: An appraisal of current knowledge*. New York: Oxford.

Berninger, V. (2006). Defining and differentiating dyslexia, dysgraphia, and language learning disability within a working memory model. In E. Silliman & M. Mody (Eds.), *Language impairment and reading disability – Interactions among brain, behavior, and experience:* Guilford Press.

Bornstein, M.H., Hahn, C., & Haynes, O.M. (2004). Specific and general language performance across early childhood: Stability and gender considerations. *First Language, 24,* 267-303.

Bouchard, T.J., Lykken, D.T., McGue, M., Segal, N.L., and Tellegen, A. (1990). Sources of human psychological differences: The Minnesota study of twins reared apart. *Science 12,* 223-228.

Byrne, B., Samuelsson, S., Wadsworth, S.J., Hulslander, J., Corley, R., DeFries, J.C., Quain, P., Willcutt, E.G., & Olson, R.K. (2007). Longitudinal twin study of early literacy development: Preschool through Grade 1. *Reading and Writing, 20,* 77– 102.

Chall, J.S., (1983). Learning to read: The great debate. New York: McGraw-Hill.

Chan, D.W., Suk-han Ho, C., Suk-man, T., Suk-han, L., & Chung, K.K.H. (2007). Prevalence, gender ratio and gender differences in reading-related cognitive abilities among Chinese children with dyslexia in Hong Kong. *Education Studies, 33,* 249-265.

Cole, P.M., Teti, L.O., and Zahn-Waxler, C. (2003). Mutual emotion regulation and the stability of conduct problems between preschool and early school age. *Development and Psychopathology, 15,* 1-18.

Coles, M., & Hall, C. (2002). Gendered readings: Learning from children's reading choices. *Journal of Research in Reading, 25,* 96-108.

Coulson, A.J., (1996). Schooling and literacy over time: The rising cost of stagnation and decline. *Research in the Teaching of English, 30,* 311-327.

Critchley, M. (1970). *The dyslexic child.* London: William Heinemann Medical Books.

Decker, S.N., and Vandenberg, S.G. (1985). Colorado twin study of reading disability. In D.B.Gray and J.F. Kavanagh (eds), *Biobehavioral measures of dyslexia* Parkton, MD: York Press. (pp. 123-135).

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. (1988). Colorado Reading Project: Longitudinal analysis. *Annals of Dyslexia, 38,* 120-130.

DeFries, J.C. (1989). Gender ratios in children with reading disability and their affected relatives: A commentary, *Journal of Learning Disabilities, 22,* 39-47.

DeFries, J.C., & Alarcón, M. (1996). Genetics of specific reading disability. *Mental Retardation and Developmental Disabilities Research Reviews, 2,* 39-47.

DeFries, J.C., & Baker L.A. (1983). Parental contributions to longitudinal stability of cognitive measure in the Colorado Family Reading Study. *Child Development, 54*, 388-395.

DeFries, J.C., Filipek, P.A., Fulker, D.W., Olson, R.K., Pennington, B.F., Smith, S.D., & Wise, B.W. (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: Twin Research, 37,* 205-216.

DeFries, J.C., & Gillis, J.J. (1991). Etiology of reading deficits in learning disabilities: Quantitative genetic analysis. In J.E. Obrzut & G.W. Hynd (Eds.), Neuropsychological foundations of learning disabilities: *A handbook ofissues, methods and practice* (pp. 29-47). Orlando, Fl: Academic Press.

DeFries, J.C. & Light, J.G. (1996). Twin studies of reading disability. In: J.H. Beitchman, N.J. Cohen, M.M. Konstantareas & R. Tannock (eds.) Language, learning and behavior disorders: Developmental, biological, and clinical perspectives (pp. 272-292). New York: Cambridge University Press.

DeFries, J.C., Singer, S.M., Foch, T.T., & Lewitter, F.L. (1978). Familial nature of reading disability. *British Journal of Psychiatry, 132,* 361-367.

DeFries, J.C., Vogler, G.P., & LaBuda, M.C. (1986). Colorado family reading study: An overview. In J.L. Fuller & E.C. Simmel (Eds.), *Perspectives in behavior genetics* (pp. 29-56). Hillsdale, H.J: Erlbaum.

Dirks, E., Spyer, G., van Lieshout, E. C., & de Sonneville, L. (2008). Prevalence of combined reading and arithmetic disabilities. *Journal of Learning Disabilities, 41,* 460-474.

Drew, A. (1956). A neurological appraisal of familial congenital word-blindness. *Brain, 79,* 440-460.

Dunn, L.M., & Markwardt, F.C. (1970). *Examiner's manual: Peabody Individual Achievement Test*. Circle Pines, MN: American Guidance Service.

Ehringer, M.A., Rhee, S.H., Young, S., Corely, R. & Hewitt, J.K. (2006). Genetic and environmental contributions to common psychopathologies of childhood and adolescence: A study of twins and their siblings. *J. of Abnormal Child Psychology, 34,* 1-17.

Feldman, E., Levin, B.E., Fleischmann, J., Jallad, B., Kushch, A., Gross-Glenn, K., Rabin, M., & Lubs, H.A. (1995). Gender differences in the severity of adult familial dyslexia. *Reading and writing: an Interdisciplinary Journal, 7,* 155-161.

Fenson, L., Dale, P., Reznick, J.S, Thal, D., Bates, E., Hartung, J., Pethick, S. & Reilly, J. (1993). *The MacArthur Communicative Development Inventories: User's guide and technical manual.* San Diego: Singular Publishing Group.

Fergusson, D.M., Horwood, L.J., Shannon, F.T., & Lawton, J.M. (1989). The Christchurch Child Development Study: a review of epidemiological findings. *Paediatr Perinat Epidemiol, 3,* 278-301.

Finucci, J.M. (1978). Genetic considerations in dyslexia. In Myklebust H.R. (ed): Progress in Learning Disabilities, vol 4, New York: Grune and Stratton, (pp 41-63).

Finucci J.M. & Childs, B. (1981). Are there really more dyslexic boys than girls? In Ansara A, Geschwind N, Galaburda A et al. (eds.): Sex Differences in Dyslexia. Towson, MD: Orton Dyslexia Society, (pp 1-9).

Finucci J.M. & Childs, B. (1983). Dyslexia: Family studies. In Ludlow CA, Cooper JA (Eds.). Genetic Aspects of Speech and Language Disorders. New York: Academic Press, pp 157-167.

Finucci, J.M., Guthrie, J.T., Childs, A.L., Abbey, H., & Childs, B. (1976). The genetics of specific reading disability. *Annals of Human Genetics, 40,* 1-23.

Fisher, J.H. (1905). Case of congenital word-blindness (Inability to learn to read). *Ophthalmic Review,* 315.

Fisher, S.E. & DeFries, J.C. (2002). Developmental dyslexia: genetic dissection of a complex cognitive trait. *Nature Reviews Neuroscience, 3,* 767-780.

Fletcher, J.M., Coulter, W.A., Reschly, D.J., & Vaughn, S. (2004). Alternative approaches to the definition and identification of learning disabilities: Some questions and answers. *Annals of Dyslexia, 54,* 304-331.

Fletcher, J.M. & Morris, R. (1986). Classification of disabled learners: Beyond exclusionary definitions. In *Handbook of Cognitive, Social, and Neuropsychological Aspects of Learning Disabilties,* Vol. I, ed. S.J. Cicci. Hillsdate, NJ: Lawrence Erlbaum Associates.

Foch, T.T., DeFries, J.C., McClearn, G.E., & Singer, S.M. (1977). Familial patterns of impairment in reading disability. *Journal of Educational Psychology, 69,* 316-329.

Friend, A., DeFries, J.C., & Olson, R.K. (2008). Parental education moderates genetic influences on reading disabilities. *Psychological Science, 11*, 1124-1130.

Gardner, E.F., Rudman, H.C., Karlsen, B., & Merwin, J.C. (1982). *Stanford achievement test*. San Antonio, Tx: Psychological Corp.

Gayán, J. & Olson, R.K. (2001). Genetic and environmental influences on orthographic and phonological skill in children with reading disabilities. *Developmental Neuropsychology, 20,* 483-507.

Gilger, J.W., & Kaplan, B.J. (2001). Atypical brain development: A conceptual framework for understanding learning disabilities: *Developmental Neuropsychology, 20,* 465-481.

Gilger, J.W., Pennington, B.F., & DeFries, J.C. (1991). Risk for reading disability as a function of parental history in three family studies. *Reading Writing: Interdisciplinary J,* 3, 205-217.

Gilmore, A., Croft, C., & Reid, N. (1981). *Burt Word Reading Test,* New Zealand Revision: Teachers Manual. Wellington: New Zealand Council for Education Research.

Hallgren, B. (1950). Specific dyslexia: A clinical and genetic study. *Acta Psychiatr Neurolog Scand, 65, (Suppl),* 1-287.

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.

Harris, A.S., & Sipay, E.R. (1990). *How to improve reading ability*. White Plains, NY: Longman.

Hawke, J.L., Olson, R.K., Willcutt, E.G., Wadsworth, S.J., & DeFries, J.C. (2009). Gender ratios for reading difficulties. *Dyslexia, 15,* 239-242.

Hawke, J.L., Wadsworth, S.J., & DeFries, J.C. (2006). Genetic influences on reading difficulties in boys and girls: The Colorado twin study. *Dyslexia, 12,* 21-29.

Hawke, J.L., Wadsworth, S.J., Olson, R.K., & DeFries, J.C. (2007). Etiology of reading difficulties as a function of gender and severity. *Reading and Writing, 20,* 13-25.

Hinshelwood, J. (1907). Four cases of congenital word-blindness occurring in the same family. *British Medical Journal, 2,* 1229-1232.

Hinshelwood, J. (1911). Two cases of hereditary word-blindness. *British Medical Journal, 1,* 608-609.

Ho, C.S.-H., Chan, D.W., Tsang, S. & Lee, S. (2000a). *The Hong Kong Test of Specific Learning Difficulties in Reading and Writing (HKT-SpLD) Manual* (Hong Kong, Hong Kong Specific Learning Difficulties Research Team).

Hulslander, J., Olson, R.K., Willcutt, E.G., & Wadsworth, S.J. (2010). Longitudinal stability of reading-related skills and their prediction of reading development. *Scientific Studies of Reading, 14,* 111-136.

Juel, C. (1988). Learning to read and write: A longitudinal study of 54 children from first through fourth grades. *Journal of Educational Psychology, 80,* 437-447.

Kazelskis, R., Thames, D., Reeves, C., Flynn, R., Taylor, L., Beard, L.A., & Turnbo, D. (2005). Reliability and stability of Elementary Reading Attitude Survey (ERAS) scores across gender, race, and grade level. *Elementary Reading Attitude Survey Scores, 27,* 29-37.

Kempe, C., Gustafson, S., & Samuelsson, S. (2003). A longitudinal study of early reading difficulties and subsequent problem behaviors. *Scandinavian Journal of Psychology, 52,* 242-250.

Kirk, S.A., & Bateman, B. (1962). Diagnosis and remediation of learning disabilities. *Exceptional Child, 29,* 73-78.

Knopik, V.S., DeFries, J.C., & Alarcon, M. (1996). Gender differences in cognitive abilities of opposite-sex and same-sex twin pairs with reading disability. *Annals of Dyslexia, 46,* 241-257.

Koeppen-Schomerus, G., Spinath, F.M., & Plomin, R. (2003). Twins and non-twin siblings: different estimates of shared environmental influence in early childhood. *Twin Research, 6,* 97-105.

Kush, J.C. & Watkins, M.W. (1996). Long-term stability of children's attitudes toward reading. *The Journal of Education Research, 89,* 315-319.

Levine, P., & Nourse, S.W. (1998). What follow-up studies say about postschool life for young men and women with learning disabilities: a critical look at the literature. *J Learn Disabil, 31,* 212-233.

Lerner, R.M. & Busch-Rossnagel, N.A. (Eds.) (1981). *Individuals as producers of their development: A lifespan perspective.* New York: Academic Press.

Light, J.G., & DeFries, J.C. (1995). Comorbidity of reading and mathematics disabilities: Genetic and environmental etiologies. *Journal of Learning Disabilities, 28,* 96-106.

Lyon, G.R. (1995). Toward a definition of dyslexia. *Annals of Dyslexia, 45,* 3-27.

Lyon, G.R. (1996). Learning Disabilities. In E. Mash & R. Barkley (Eds.), Child Psychology (pp. 418-419). New York: Guilford Press.

Lyon, G.R., & Chhabra, V. (1996). The current state of science and the future of specific reading disability. *Mental Retardation and Developmental Disabilities Research Reviews, 2,* 2-9.

Lyon, G.R., Fletcher, J.M., & Barnes, M. (2002). Learning disabilities. In E.J. Marsh & R.A. Barkley (Eds.), Child psychopathology (2nd edn, pp. 520-588). New York: Guilford.

Lyon, G.R., Shaywitz, S.E., & Shaywitz, B.A. (2003). A definition of dyslexia. *Ann Dyslyexia, 53,* 1.14.

McKenna, M.C. & Kear, D.J. (1990). Measuring attitude toward reading: Anew tool for teachers. *The reading teacher, 43,* 626-639.

Medland, S.E., Wright, M.J., Geffen, G.M., Hay, D.A., Levy, F., Martin, N.G., & Duffy, D.L. (2003). Special twin environments, genetic influences and their effects on the handedness of twins and their siblings. *Twin Research, 6,* 119-130.

Melekian, B.A. (1990). Family characteristics of children with dyslexia. *J Learning Disabilities, 23,* 386-391.

Meltzer, H., Gatward, R., Goodman, R. & Ford, T. (2000). The mental health of children and adolescents in Great Britain. London, England: Stationery Office.

Merriman, C. (1924). The intellectual resemblance of twins. *Psychological monographs, 33,* 1-58.

Moffitt, T.E., for the E-Risk Study Team. (2002). Teenaged mothers in contemporary Britain. *Journal of Child Psychology and Psychiatry, 43,* 727-742.

Morgan, W.P. (1896). A case of congenital word blindness. *The British Medical Journal, Nov*. 1378.

Morris, R.D., Lovett, M.W., Wolf, M., Rose, A., Sevcik, K.A., Steinbach, J.C., & Shapiro, B. (2010). Multiple component remediation for developmental reading disabilities: IQ, socioeconomic status, and race as factors in remedial outcome. *J. of Learning Disabilities, online, DOI: 10.1177/0022219409355472*

Nettleship, E. (1901). Cases of congenital word-blindness (inability to learn to read). *Ophthalmic Review, 20,* 61-67.

Nichols, R.C., & Bilbro Jr., W.C. (1966). The diagnosis of twin zygosity. *Acta Geneticae Medicae et Gemellologiae, 16,* 265-275.

Norrie, E. (1954). Ordblindhedens (dyslexiens) arvegang. *Laesepaedagogen,* 2:61; cited by Stromgren, E. Trends in psychiatric genetics in Scandinavia, in Kallmann, F.J. Expanding goals of genetics in psychiatry. London/New York: Grune & Stratton, 1962.

Olson, R., Hulslander, J., Christopher, M., Keenan, J., Wadsworth, S.J., Willcutt, E.G., Pennington, B., & DeFries, J.C. (2011). Genetic and environmental influences on writing and their relations of language and reading. *Annals of dyslexia.* DOI 10.1007/s11881-011-0055-z.

Olson, R., Keenan, J., Byrne, B., Samuelsson, S., Coventry, W., Corley, R., Wadsworth, S., DeFries, J., Pennington, B., & Hulslander, J. (2011). Genetic and environmental influences on vocabulary and reading development from pre-kindergarten through grade 4. *Scientific Studies of Reading.*

Pennington, B.F., & Lefly, D.L. (2001). Early reading development in children at family risk for dyslexia. *Child Development, 72,* 816-833.

Pennington, B.F., Peterson, R.L., & McGrath, L.M. (2009). Dyslexia. In: Pennington, B.F. (ed.). *Diagnosing Learning Disorders: A neuropsychological framework*. 2nd ed. New York, NY: Guilford; 2009: 45-82.

Pedersen, N.L., Plomin, R., Nesselroade, J.R., & McClearn, G. (1992). A quantitative genetic analysis of cognitive abilities during the second half of life span. *Psychological Science, 3,* 346-353.

Petrill, S.A., Deater-Deckard, K., Thompson, L.A., Schatschneider, C., Dethorne, L.S., & Vandenbergh, D.J. (2007). Longitudinal genetic analysis of early reading. The Western Reserve Reading Project. *Reading and Writing, 20,* 127-146.

Purcell, S., & Sham, P.C. (2003). A model-fitting implementation of the DeFries-Fulker model for selected twin data. *Behavior Genetics, 33,* 271-278.

Raskind, M.H., Goldberg, R.J., Higgins E.L., & Herman, K.L. (1999). Patterns of change and predictors of success in individuals with learning disabilities: results from a twenty-year longitudinal study. *Learn Disabil Res Pract, 14,* 35-49.

Resnick, L.B., & Weaver, P.A. (1979). *Theory and practice of early reading (3 vols.).* Hillsdale, N.J.: Lawrence Erlbaum.

Reynell, J.K. & Gruber, C.P. (1990). *Reynell Developmental Language Scales U.S. Edition – Second Revision.* Los Angeles, CA: Western Psychological Services.

Rutter, M., Caspi, A., Fergusson, D., Horwood, L.J., Goodman, R., Maughan, B., Moffitt, T.E., Meltzer, H., & Carrol, J. (2004). Sex differences in developmental reading disability: New findings from four epidemiological studies. *Journal of the American Medical Association, 291,* 2007-2012.

Rutter, M., & Yule, W. (1975). The concept of specific reading retardation. *J. Child Psychol Psychiatry, 16,* 181-197.

Rutter, M., Caspi, A., Fergusson, D., Horwood, L.J., Goodman, R., Maughan, B., Moffitt, T.E., Meltzer, H. & Carroll, J. (2004). Sex differences in developmental reading disability: New findings from 4 epidemiological studies, *Journal of 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*. Timonium, MD: York Press.

Share, D.L., & Silva, P.A. (2003). Gender bias in IQ-discrepancy and post-discrepancy definitions of reading disability. *Journal of Learning Disabilities, 36,* 4-14.

Shastry, B. (2007). Developmental dyslexia: an update. *Journal of Human Genetics, 52,* 104-109.

Shaywitz, S.E. (1996). Dyslexia. *Sci. Am. 275,* 98-104.

Shaywitz, S.E., Shaywitz, B.A., Pugh, K.R., Fulbright, R.K., Constable, R.T., Mencl, W.E., et al. (1998). Functional disruption in the organization of the brain for reading in dyslexia. *Proc Natl Acad Sci USA, 95,* 2636-2641.

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 ability. *New England Journal of Medicine, 326,* 144-150.

Shaywitz, S.E., Shaywitz, B.A., Fletcher, J.M., & Escobar, M.D. (1990). Prevalence of reading disability in boys and girls. *Journal of American Medical Association, 264,* 998-1002.

Shaywitz, S.E., Fletcher, J.M., Holahan, J.M., Shneider, A.E., Marchione, K.E., Stuebing, K.K., Francis, D.J., Pugh, K.R., & Shaywitz, B.A.. (1999). Persistence of dyslexia: The Connecticut longitudinal study at adolescence. *Pediatrics, 104,* 1351-1359.

Shaywitz, S.E., & Shaywitz, B.A. (2005). Dyslexia (Specific reading disability). *Biol Psychiatry, 57,* 1301-1309.

Siegal, L.S., & Smythe, I.S. (2005). Reflection on research on reading disability with special attention to gender issues. *Journal of Learning Disabilities, 38,* 473-477

Sinclair, A.H.H. (1948). Developmental aphasia. *The British Journal of Ophthalmology, 32,* 522-531.

Silva, P.A. (1990). The Dunedin Multidisciplinary Health and Development Study: A 15-year longitudinal study. *Paediatric and Perinatal Epidemiology, 4,* 76-107.

Silva, P.A., & Stanton, W.R. (1996). From child to adult: *The Dunedin Multidisciplinary Health and Development Study.* Auckland, New Zealand: Oxford University Press.

Spira, E.G., & Fischel, J.E. (2005). The impact of preschool inattention, hyperactivity, and impulsivity on social and academic development: a review. *Journal of Child Psychology and Psychiatry, 46,* 755-773.

Stanovich, K. (1986). Cognitive processes and the reading problems of learning-disabled children: evaluating the assumption of specificity. In: Torgesen JK, Wong BYL (Eds.) Psychological and education perspectives on learning disabilities. Academic Press, Orlando.

Stanovich, P.J., & Stanovich, K.E. (2003). Using research and reason in education: How teachers can use scientifically based research to make curricular & instructional decisions. Produced by RMC Research Corporation, Portsmouth, New Hampshire.

St. Stauver, J.L., Katusic, S.K., Barbaresi, W.J., Colligan, R.C., & Jacobsen, S.J. (2001). Boy/Girl differences in risk for reading disability: Potential clues? Am J of *Epidemiology, 154,* 787-794.

Stephenson, S. (1907). Six cases of congenital work-blindness affecting three generations of one family. *Opthalmascope, 5,* 482-484.

Stevenson, J. (1992). Identifying sex differences in reading disability: Lessons from a twin study, *Reading and Writing 4,* 307-326.

Stevenson, J., Graham, P., Fredman, G., & McLoughlin, V. (1987). A twin study of genetic influences on reading and spelling ability and disability. *Journal of Child Psychology and Psychiatry, 28,* 229–247.

Stevenson, J., Graham, P., Fredman, G. & McLoughlin, V. (1984). The genetics of reading disability. In Turner C.J., Miles, H.B. (eds.). *The Biology of Human Intelligence.* Nafferton: Nafferton Books Limited, pp 85-97.

Thompson, J.S. & Thompson, M.W. (1986): Genetics in Medicine, Philadelphia, PA: W.B. Saunders.

Thomas, C.J. (1905). Congenital "word-blindness" and its treatment. *Ophthalmoscope, 3,* 380-385.

Torgensen, J., Alexander, A., Wagner, R., Rashotte, C., Voeller, K., & Conway, T. (2001). Intensive remedial instruction for children with severe reading disabilities: Immediate and long-term outcomes from two instructional approaches. *Journal of Learning Disabilities, 34,* 33-58.

Torgensen, J.K., Wagner, R.K., & Rashotte, C.A. (1999). *Test of Word Reading Efficiency.* Austin, Tex: PRO-ED Inc.

Tunmer, W., & Greaney, K. (2010). Defining dyslexia. *Journal of Learning disabilities, 43,* 229-243.

Van Grootheest, D.S., Cath, D.C., Beekman, A.T., & Boomsma, D.I. (2007). Genetic and environmental influences on obsessive-compulsive symptoms in adults: a population-based twin-family study. *Psychological Medicine, 37,* 1635-1644.

Vellutino, F.R., Scanlon, D.M., & Jaccard, J. (2003). Toward distinguishing between cognitive and experiential deficits as primary sources of difficulty in learning to read: A two-year follow-up of difficult to remediate and readily remediated poor readers. In B.R. Foorman (Ed.) Preventing and remediating reading difficulties: Bringing science to scale (pp. 73-120). Baltimore: York Press.

Vellutino, F.R., Scanlon, D.M., Small, S., Fanuele, D.P. (2006). Response to intervention as a vehicle for distinguishing between children with and without reading disabilities: Evidence for the role of kindergarten and first-grade interventions. *J of Learning Disabilities, 39,* (157-169).

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.

Vogler, G.P., DeFries, J.C., & Decker, S.N. (1985). Family history as an indicator of risk for reading disability. *Journal of Learning Disabilities, 18,* 419-421.

Wadsworth, S.J., Corley, R.P., Plomin, R., Hewitt, J.K., & DeFries, J.C. (2002). Parent-offspring resemblance for reading performance at 7, 12, and 16 years of age in the Colorado Adoption Project. *Journal of Child Psychology and Psychiatry, 46,* 769-774.

Wadsworth, S.J., Corley, R.P., Plomin, R., Hewitt, J.K., & DeFries, J.C. (2006). Genetic and environment influences on continuity and change in reading achievement in the Colorado Adoption Project. In A.C. Huston & M.N. Ripke (Eds), *Developmental contexts in middle childhood.* 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.

Wadsworth, S.J., DeFries, J.C., Stevenson, J., Gilger, J.W., & Pennington, B.F. (1992). Gender ratios among reading-disabled children and their siblings as a function of parental impairment. *Journal of Child Psychology and Psychiatry, 33,* 1229-1239.

Wadsworth, S.J., DeFries, J.C., Olson, R.K., & Willcutt, E.G. (2007). The Colorado Longitudinal Twin Study of Reading Disability. *Annals of Dyslexia, 57,* 139-160.

Wadsworth, S.J., Knopik, V.S., & DeFries, J.C. (2000). Reading disability in boys and girls: No evidence for a differential genetic etiology. *Reading and Writing: An Interdisciplinary Journal, 13,* 133-145.

Wadsworth, S.J., Olson, R.K., & DeFries, J.C. (2010). Differential genetic etiology of reading difficulties as a function of IQ: an update. *Behavior Genetics,40,* 751-758.

Wagner, R.K., Torgeson, K.K., Rashotte, C.A., Hecht, S.A., Barker, T.A., Burgess, S.R., Donahue, J., & Garon, T. (1997). Changing relations between phonological processing abilities and word-level reading as children develop from beginning to skilled readers: A 5-year longitudinal study. *Developmental Psychology, 33,* 468-479.

Warburg, F. (1911). Uber die angeborene Wortblindheit und die Bedeutung ihrer Kenntnis fur den Unterricht. *Zeischrift fur Kinderforschung, 16,* 97-113.

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. (1990). *Wechsler Preschool and Primary Scale of Intelligence – Revised.* London, England: Psychological Corp, Harcourt Brace & Co.

Wilkinson, G. (1993). *Wide range achievement tests – Revised.* Inc., Wilmington, DE: Wide Range.

Williams, S., & McGee, R. (1996). Reading in childhood and mental health in early adulthood. In J.H. Beitchman et al., (Eds.). *Language, learning, and behavior disorders: Developmental, biological, and clinical perspectives,* pp. 530-554. Cambridge: Cambridge University Press.

Wise, B. W., Ring, J., & Olson, R. K. (2000). Individual differences in gains from computer-assisted remedial reading with more emphasis on phonological analysis or accurate reading in context. *Journal of Experimental Child Psychology*, *77*, 197–235.

Woodcock, R. (1987). *Woodcock reading mastery tests-Revised.* Circle Pines, MN: American Guidance Service.

Woodcock, R., & Johnson, B. (1990). *Woodcock-Johnson psycho-educational battery revised tests of achievement*. Chicago: Riverside Publishing.

Young, S.E., Rhee, S.H., Stallings, M.C., Corley, R.P., & Hewitt, J.K. (2006). Genetic and environmental vulnerabilities underlying adolescent substance use and problem use: general or specific? *Behavior Genetics, 36*, 603-615.

Zerbin-Rudin, E. (1967). Kongenitale Worblindheit oder spezifische dyslexie (congenital word-blindness). *Bulletin of the Orton Society, 17*, 47-56.

Zieleniewski, A.M., Fulker, D.W., DeFries, J.C., & LaBuda, M.C. (1987). Multiple regression analysis of twin and sibling data. *Personality of Individual Differences 8*, 787-791.