MODERATION OF COGNITIVE—ACHIEVEMENT RELATIONS FOR CHILDREN WITH SPECIFIC LEARNING DISABILITIES: A MULTI-GROUP LATENT VARIABLE ANALYSIS USING CHC THEORY

by

Christopher R. Niileksela
B. M., Arizona State University, 2004
B. A., Arizona State University, 2005
Ed. S., University of Kansas, 2009

Submitted to the graduate degree program in Psychology and Research in Education and the Faculty of the Graduate Faculty of the University of Kansas in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Dissertation Committee:

_____
Co-Chairperson: Steven Lee

_____
Co-Chairperson: Matthew Reynolds

_____
Patricia Lowe

_____
Julia Shaftel

_____
Daryl Mellard

_____
Bruce Frey

Dissertation defended: April 2nd, 2012

The Dissertation Committee for Christopher R. Niileksela

certifies that this is the approved version of the following dissertation:

MODERATION OF COGNITIVE—ACHIEVEMENT RELATIONS FOR CHILDREN WITH SPECIFIC LEARNING DISABILITIES: A MULTI-GROUP LATENT VARIABLE ANALYSIS USING CHC THEORY

_____

Co- Chairperson: Steven Lee

_____

Co- Chairperson: Matthew Reynolds

Date approved: April 2nd, 2012

ABSTRACT

Recent advances in the understanding of the relations between cognitive abilities and academic skills have helped shape a better understanding of which cognitive processes may underlie different types of SLD (Flanagan, Fiorello, & Ortiz, 2010). Similarities and differences in cognitive—achievement relations for children with and without SLDs have not been extensively studied. This study examined whether cognitive—achievement relations are similar among groups of children with SLD in reading (*n* = 181), math (*n* = 231), and writing (*n* = 149), when compared to children without SLD (*n* = 300) using the Woodcock-Johnson – Third Edition (WJ-III; Woodcock, McGrew, & Mather, 2007). Multi-group structural equation modeling (SEM) was used to examine cognitive—achievement relations. A three-stratum model of cognitive abilities based on Cattell-Horn-Carroll (CHC) theory was used in the analysis. Results showed that the factor structure and factor loadings of the CHC model were invariant among groups, and SLD group membership moderated the magnitude of several cognitive—achievement relations: (a) Knowledge (K0), Short-term Memory (Gsm), and Quantitative Reasoning (RQ) were important predictors of Basic Reading Skills (BRS) across all groups, but Perceptual Speed (PS) was also an important predictor of BRS for individuals with SLD in reading; (b) K0 and RQ were important predictors of Reading Comprehension (RC) for all groups, but RQ had a stronger relation to RC for individuals with SLD in reading; (c) PS, and Gsm were important predictors for Math Calculation Skills (MCS) in all groups, and RQ predicted MCS for all groups except those with SLD in math; (d) RQ and K0 predicted Applied Math (AM) for all groups, but visualization (VZ) was also an important predictor of AM for individuals with SLD in math and SLD in writing; and (e) Gc, RQ, VZ, Memory Span (MS), and Rapid Naming (RN) were important predictors of written expression, and the only difference between groups was those

with SLD in math relied slightly more on Gc. Results suggest that individuals with SLD in specific academic areas may rely on some different cognitive abilities as a way to compensate for deficits in academic skills.

ACKNOWLEDGEMENTS

My experience at the University of Kansas has been one of the most significant events of my life, and it is the people at the University of Kansas who have made this experience valuable. Each of the school psychology faculty members played a significant role in my education. I wish to thank Matthew Reynolds, for his patience and assistance in helping me learn about the areas and methods of cognitive ability research from the ground up, as well as his advice and guidance through the dissertation process; Patricia Lowe, for her strong guidance in nearly every area of graduate school, from the introductory classes to the internship and graduation process, as well as her insightful comments on my writing; Steven Lee, for providing a variety of research experiences and who has helped me think more critically about research, as well as considering both practical and theoretical implications in my own research and practice; and Julia Shaftel, for her supervision and expertise in assessment and how they can be best used and interpreted in the field. I also want to thank Daryl Mellard and Bruce Frey for being a part of my dissertation committee. I appreciate the time and effort they have put into helping me complete this project.

Faculty members from other departments have also been influential in the development of my current interests. Specifically, I want to thank Ed Morris and Greg Madden from the department of Applied Behavioral Science for providing me with a strong background in understanding scientific methods and philosophy. I also want to thank John Poggio for providing me with the opportunity to better understand statistics by allowing me to be a teaching assistant for his class. Without the trust put in me by these faculty members, I would not be where I am today. I also want to thank Kevin McGrew and the Woodcock-Muñoz Foundation for making the data used in this study available. Their work in compiling cognitive and academic test scores for

the Clinical Database Project will help better understand the cognitive and academic skills of children and adults with disabilities. I hope my small contribution to this research satisfactory.

Finally, I want to thank my wife, Sarah, for her unwavering support which has helped me survive the past six years of graduate school. Without her, I would not be who I am today, I would not be where I am today, and her impact on my life has been more meaningful than I can say. I am wholly indebted to her for everything.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter I: Introduction

For many years, researchers and practitioners have tried to understand the nature of learning disabilities (Fletcher, Lyon, Fuchs, & Barnes, 2007; Gregg, 2009). In addition to being difficult to define, individuals with specific learning disabilities (SLD) make up a heterogeneous group (Fletcher et al., 2007) and the identification process for SLD continues to be an area of controversy (e.g., Bradley, Danielson, & Hallahan, 2002). The idea that specific cognitive processing deficits may underlie SLD has been an explanatory hypothesis for SLD determination throughout its history (Lyon, Fletcher, & Barnes, 2003). Some researchers have recently begun to refocus their energy on understanding the underlying cognitive processes which may affect an individual's ability to learn through the use of systematic identification models based on current advances in the understanding of human cognitive abilities (e.g., Fiorello, Hale, & Snyder, 2006; Flanagan, Fiorello, & Ortiz, 2010; Flanagan, Ortiz, Alfonso, & Dynda, 2006). The assessment of cognitive abilities has been a pervasive practice for SLD identification, and with a greater emphasis on the assessment of cognitive processes it is important to have a strong theoretical and empirical foundation from which decisions about educational services must be made. In order to move in this direction, work needs to be done to better understand the relations between cognitive abilities and academic achievement in children and adolescents with and without SLD.

Cognitive Abilities and Academic Achievement

Theory and research in the area of cognitive abilities has changed dramatically in the past 20 years, especially with the advent of Cattell-Horn-Carroll (CHC) theory (McGrew, 2005, 2009). CHC theory has provided a strong theoretical and empirical foundation to the understanding of cognitive abilities, which has guided both research (Keith & Reynolds, 2010) and practice (Alfonso, Flanagan, & Radwan, 2005). CHC theory represents the conglomeration

of several important areas in cognitive ability research, including Spearman's two factor theory (Spearman, 1904), Thurstone's primary mental abilities (Thurstone, 1938), and Horn and Cattell's Gf-Gc theory (Horn & Cattell, 1966). This theory has provided a strong foundation for the development and revision of many tests, including the Woodcock-Johnson—Third Edition (WJ-III, Woodcock, McGrew, & Mather, 2001, 2007). Moreover, CHC theory provides a taxonomic structure which allows for the empirically-supported classification of cognitive ability tests.

Advances in cognitive ability theory and assessment, have helped shape a better understanding of which cognitive processes are related to academic skills has become an important research question in educational and school psychology. Researchers have examined how different cognitive abilities are related to reading (Elliot, Hale, Fiorello, Dorvil, & Moldovan, 2010; Floyd, Taub, Keith, & McGrew, 2007), mathematics (Proctor, in press; Proctor, Floyd, & Shaver, 2005; Taub, Floyd, Keith, & McGrew, 2008) and writing (Floyd, McGrew, & Evans, 2008). Some of this research has been summarized by McGrew and Wendling (2010), who suggest that a better understanding of these cognitive—achievement relations can be important for understanding SLD. Much of this research has focused on how cognitive abilities are related to achievement in normally developing populations, but few have examined these relations for individuals with learning disabilities (although see Elliot et al., 2010; Proctor, in press; Swanson & Alexander, 1997).

Examining cognitive—achievement relations in populations of individuals with SLD is important for several reasons. First, children with SLD make up the largest group of children in special education in the United States (U.S. Department of Education, 2010), making it the most pervasive disability in schools. Children with SLDs make up over 5% of the total school

population, which is nearly half of the students in special education. With the large number of students with SLD, it is essential to know whether or not the measurement instruments used for identification purposes are adequately measuring these constructs in both groups of children with and without SLD. Second, while understanding how cognitive skills are related to academic skills in normally developing children and adolescents is helpful for understanding the underlying processes of academic achievement, it is possible that identified groups of children with SLD rely on different processes. If a person has a deficit in one cognitive process which is related to achievement for normally developing individuals, it is possible that a person with a SLD is using a different process to compensate for that deficit. Thus, if SLDs are going to be adequately understood as a deficit in a psychological process, it is important to identify which processes explain performance on specific academic skills in general, as well as which processes might be important for specific academic skills in children with SLD. This information may also be important in the planning and designing instructional interventions for children with SLD. For instance, if children with SLD in writing tend to rely more on visual skills than children without SLDs, including a visual component in an intervention may be an important consideration to help improve writing skills for children with SLD since it is predictive of better performance among children with SLD. Knowing whether or not particular cognitive processes may be related to performance for children in these groups may be useful in the treatment of SLD. Additionally, clarification of measurement properties is necessary for more basic research that focuses on the cognitive and neuropsychological underpinnings of SLD.

Factorial Invariance and Structural Equation Modeling

To examine the relations between cognitive abilities and academic achievement for children with and without SLD, it is also necessary to examine whether the constructs being

measured are the same between groups. Modern methods of statistical analysis, especially the use of confirmatory factor analysis (CFA) and structural equation modeling (SEM) comprise some of the best methods available for examining the measurement properties of psychological and educational measures. One important analytic method which has been developed in this tradition is testing for *measurement invariance*, which refers to the requirement that an instrument designed to measure a particular psychological or educational variable measures the variable in the same manner across groups (Meredith, 1993; Millsap, 2011). In other words, the measurement instrument should not be affected by extraneous variables, namely group membership, and by showing that measurement parameters are nearly equivalent among groups provides evidence that the measurement instrument is measuring the same construct across groups.

In psychology and education, much of research is based on *latent constructs*, which are variables of interest to researchers and practitioners, but they cannot be measured directly (Keith, 2006; Osterlind, 2010). Rather, these constructs are measured indirectly through a variety of methods, which could include questionnaires or behavioral observation. Cognitive and academic skills are commonly measured through the use of tests. The assumption is that these measurement devices are measuring the same constructs across different groups of people. This assumption may be tested within a factor model, where the different model parameters (e.g., factor loadings, intercepts, residual variances) are tested for equality among groups. If the model parameters are similar, then the construct (or, latent variable in SEM terms) has the same effect on the measured variables among groups. When this assumption is met, differences among groups can be attributed to differences the latent construct.

For example, if two people from two groups have the same level of a latent variable (such as depression, anxiety, or mathematics ability), then the two individuals should obtain the same score on an instrument designed to measure the construct (considering measurement error). If one group is systematically higher or lower than another, even though individuals from those groups have the same amount of the construct, then the score is dependent on another factor, which might be group membership. Group membership alone should not affect someone's score, and if group membership does, the score difference suggests a possible bias in the measurement conditions. For the current study, in order to adequately determine if the relationships among cognitive and academic skills are similar or different between children with and without SLD, it is important to show that the cognitive abilities are essentially being measured in the same manner across groups of children who are developing normally, and those who have been diagnosed with SLD. The focus of the current study will be on the structural relations of cognitive abilities to achievement skills, therefore it will be important for the latent factor variances and covariances to be identified in a way that does not depend on group membership. If different groups require an alternative scaling of the variables used, the structural relations cannot be meaningfully compared. That is, corresponding relations between the latent factors and observed scores (factor loadings) should be invariant across groups in order to make the comparisons of structural relations across groups.

Much of the work on measurement invariance has examined differences between well-defined groups, such as gender, ethnicity, cultural status, or age (Chen, Keith, Weiss, Zhu, & Li, 2010; Keith, Reynolds, Patel, & Ridley, 2008; Taub & McGrew, 2004). Fewer studies have examined factorial invariance for individuals with SLD (although see Bowden et al., 2008). Examining factorial invariance and cognitive—achievement relations in groups of children with

and without SLDs may shed some light on children with SLDs use cognitive processes when engaging in academic tasks, but factorial invariance must be established before such comparisons are considered valid.

Purpose of Current Study

This study is designed to answer two major questions. First, the equality of covariance structure of a major test of cognitive abilities, the Woodcock-Johnson—Third Edition (WJ-III; Woodcock et al., 2001, 2007), will be evaluated between children with and without SLD. This research is important because it will provide a basis for construct validity of the cognitive and academic skills measured by the WJ-III for children with SLD, which is the necessary first step to make comparisons between cognitive skills and academic achievement. At a minimum, it is necessary to show that the factor structure and loadings from the latent variables on subtests are similar between groups of children with and without SLD. Showing that the factor structures and loadings are similar between groups of children with and without SLD will provide a basis for the cognitive-achievement relations comparisons which will be made in the second part of the study.

The second purpose of this study is to examine whether the structural relations between different cognitive abilities and academic skills are the same between groups of children with different types of SLD and a group of normally developing children. In other words, are the relations between cognitive and academic skills the same for children without SLD as they are for children with SLD, or, do children with SLD rely on the same cognitive processes to complete academic tasks as children without SLD? This research question is important because it will provide a better understanding of which cognitive skills are used by children with SLD, which is relevant with the important advances in theory and research in the diagnosis of SLD that

has occurred in the past 10 years (Flanagan et al. 2010). This study is focusing on the WJ-III because it is one of the most flexible instruments available to practitioners, and it is a theoretically and empirically driven test of cognitive abilities (Woodcock et al., 2001, 2007). The complete battery of cognitive and achievement tests offer over 50 possible subtests which measure a large number of narrow and broad cognitive and academic skills. This test battery was designed based on Cattell-Horn-Carroll (CHC) theory, which is one of the most well-supported and empirically driven theories of intelligence (Carroll, 1993; Keith & Reynolds, 2010; McGrew, 2005).

Summary

An improved understanding of the structural relations between cognitive abilities and academic skills in groups of children with SLD is important to better understand the cognitive processes used by children with SLD. Such research is important from both theoretical and applied perspectives, thus findings will inform basic research related to SLD (e.g., do children with and without SLDs use the same cognitive abilities in reading comprehension) and should help inform practitioners for SLD identification (e.g., are there specific cognitive—achievement relations which are only present in SLD groups). Additionally, understanding the measurement properties, specifically that the constructs are measured in the same manner, of these tests with different groups individuals is also important. This study will examine whether the relations between broad and narrow cognitive abilities assessed by the WJ-III show the same relations to academic skills, and whether the cognitive abilities and achievement skills on the WJ-III are measured in the same manner between children with and without SLD.

Chapter II: Review of the Literature

The purpose of this chapter is to provide a comprehensive review of the literature relevant to this study. The topics included in this review include (a) definitions, prevalence rates, and identification procedures for children with specific learning disabilities (SLD), (b) a history of intelligence theory leading up to current understanding and practice, (c) an overview of the Cattell-Horn-Carroll theory (CHC theory, McGrew, 2005, 2009) of intelligence, (d) a review of cognitive—achievement relations from a CHC perspective, (e) an overview of the common factor model and confirmatory factor analysis, (f) a review of the principles and practice of testing for measurement invariance, and (g) an overview of the current study.

Defining SLD

The history of understanding SLD comes from a variety of areas, including early medical practice, educational and clinical psychology, as well as social advocacy (Fletcher et al., 2007; Hallahan & Mercer, 2002). The roots of SLD can be traced to early medical practice, where some research supported the idea that specific areas of the brain were important for specific cognitive and behavioral functions. In the middle of the 19th century, both Broca and Wernicke discovered that particular areas of the brain were responsible for specific language functioning (Hallahan & Mercer, 2002), and different areas of the brain and specific language difficulties (Broca's aphasia, Wernicke's aphasia) continue to be named after these individuals. Researchers in the early 20th century provided important advances in both identifying that learning difficulties can be categorized by different content areas (e.g., Samuel Orton, Fletcher et al., 2007) and that there did not appear to be any directly visible or outward signs of brain damage in individuals with learning difficulties (Strauss & Werner, 1943). These researchers also found that individuals

with difficulties in specific learning domains showed strengths and weaknesses in other cognitive domains and that they appeared to have difficulties despite normal intellectual skills.

Historically, Samuel Kirk was also instrumental in providing a modern definition of SLD, one which continues to influence the definition used in current laws and organizations (Fletcher et al., 2007; Hallahan & Mercer, 2002). The consensus building in the field was based on the idea that SLDs occur in children who have learning characteristics which are different from normally developing children. These differences in learning were biological in nature and produced specific academic difficulties despite strengths in other cognitive areas. Moreover, children with these characteristics required some type of special education to achieve (Fletcher et al., 2007). Currently, the definition of learning disabilities according to Individuals with Disabilities Education Improvement Act of 2004 (IDEIA) is:

> The term ‘specific learning disability’ means a disorder in one of more of the basic psychological processes involved in understanding or in using language, spoken or written, that may manifest itself in the imperfect ability to listen, think, speak, read, write, spell, or to do mathematical calculations. Such term includes such condition such as perceptual disabilities, brain injury, minimal brain dysfunction, dyslexia, and developmental aphasia. Such term does not include a learning problem that is primarily the result of visual, hearing, or motor disabilities, of mental retardation, of emotional disturbance, or of environmental, cultural, or economic disadvantage.

This definition is important to the understanding what constitutes a SLD. As the definition explicitly states, SLD should evidence “a disorder in one or more of the basic psychological processes” which affects the individual’s ability to learn, and this disorder should

not occur because of other factors, including intellectual disability, economic reasons, or sensory issues. The National Joint Committee on Learning Disabilities (NJCLD) provided a separate definition of learning disabilities in 1991:

> Learning disabilities is a general term that refers to a heterogeneous group of disorders manifested by significant difficulties in the acquisition and use of listening, speaking, reading, writing, reasoning, or mathematical abilities. These disorders are intrinsic to the individual, presumed to be due to central nervous system dysfunction, and may occur across the life span. Problems in self-regulatory behaviors, social perception, and social interaction may exist with learning disabilities but do not by themselves constitute a learning disability. Although learning disabilities may occur concomitantly with other handicapping conditions (for example, sensory impairment, mental retardation, serious emotional disturbance), or with extrinsic influences (such as cultural differences, insufficient or inappropriate instruction), they are not the result of those conditions or influences (NJCLD, 1991).

This definition is similar to the federal definition provided by IDEA because it includes differences in specific skills related to learning and is presumed to be caused by a deficit in the central nervous system.

The Diagnostic and Statistical Manual for Mental Disorders—Fourth Edition—Text Revision (DSM-IV-TR, American Psychiatric Association, 2000) also includes diagnostic criteria for learning disorders. The DSM-IV-TR focuses on reading, mathematics, and writing (although it does include language disorders in another section), and the criteria suggest that an individual may have a learning disorder if their performance on standardized tests is substantially lower that

one would expect based on their age, intellectual skills, and education. Additionally, it mentions that this difficulty must have a significant impact on daily living skills and should not be caused primarily by a sensory impairment. This definition is similar to the previous definitions in that a person must be performing substantially below expectations, although this definition specifically mentions that this must also be causing significant impairment in a person's ability to be successful in their environment.

In summary, a child with a SLD is presumed to have a disorder in the central nervous system which is specific to a psychological process that affects their ability to learn or demonstrate their learning. This difficulty can be in a variety of domains, and it cannot be caused by either environmental influences (lack of instruction or opportunity) or to other disorders (sensory difficulties, intellectual disabilities). Several sources provide more extensive historical information on the definitions and advances in learning disability research and practice (Fletcher et al., 2007; Hallahan & Mock, 2003; Hallahan & Mercer, 2002).

Prevalence of SLD

Learning disabilities are currently the most common disability category to be served by special education in the United States. Based on data from 2005, the United States Department of Education (U.S. Department of Education, 2010) indicated that children with learning disabilities make up 5.29% of the school-aged population in the U. S. and outlying areas, whereas all children with disabilities make up 11.63% of the school-aged population. Numbers vary greatly by state as well. In Kentucky, only 2.11% of their students are served under the category of learning disability, and the total student population of Kentucky receiving services was 12.34%. In contrast, in Rhode Island 7.78% of students were served under the learning disability category, and 17.22% of students in Rhode Island were served under special education in 2005. These

differences likely show diverse identification practices across states and school districts. Although there is variation in the identification of SLD across states, research suggests that SLD appears more consistent in terms of identification practices over time when compared to other disabilities. Hallahan, Keller, Martinez, Byrd, Gelman, and Fan (2007) showed that over a 16 year period the identification rates for SLD have been less variable across states than other disability categories, including Emotional Disturbance, Multiple Disabilities, and even Intellectual Disability, which is a fairly well-defined disability category when considering psychological identification criteria.

Children with SLD often have comorbid diagnoses. In a study which focused on children with SLD in reading, Willcutt and Pennington (2000) found that both internalizing and externalizing difficulties were more prevalent in children with SLD when compared to children without SLD. Additionally, there were interactions by gender, with males more likely to exhibit comorbid externalizing problems (e.g., conduct disorder, oppositional defiant disorder), while females had higher rates of internalizing problems (e.g., depression). Difficulties in learning are also associated with other psychological disorders. Mayes, Calhoun, and Crowell (2000) found that nearly 70% of a group of children with Attention-Deficit Hyperactivity Disorder (ADHD) also had a comorbid SLD diagnosis. Additionally, they found that children who had a diagnosis of both ADHD and SLD tended to have more problems with academics than those with only SLD. Other disorders, such as depression, have also been found to have higher prevalence rates in groups of children with SLD than would be expected in the population (Wright-Strawderman & Watson, 1992). Clearly, SLD does not occur in isolation, but it can also be related to several other disorders in children and adolescents.

Identification of Children with SLD

When considering the U.S. as a whole, 5.29% is a substantial portion of the student population, which makes the identification and treatment of learning disabilities a major issue in education. The identification process has undergone much debate in recent years, especially with regard to the use of cognitive tests or response to intervention (RTI) for identification purposes. In fact, one of the major journals in school psychology, *Psychology in the Schools*, recently devoted two issues to the topic of cognitive assessment and RTI as methods for SLD identification (Mather & Kaufman, 2006a, 2006b).

Fletcher et al. (2007) noted that learning disabilities are difficult to identify for two major reasons: SLD is an unobservable construct, and SLD exists on a continuum or dimension. Regarding the first point, Fletcher et al. (2007) indicated that learning disabilities are defined partially by showing a child is not achieving at expected levels, and this should not be caused by other physical, medical, or psychological issues which may cause low academic skills (e.g., intellectual disability or sensory disability). This definition makes SLD difficult to observe because there are criteria which identify SLD, but other criteria of what SLD are *not* also need to be considered. The use of exclusionary criteria is also consistent with the definitions provided above, and this an important component of other identification practices promoted by researchers (e.g., Flanagan et al., 2010; Flanagan & Mascolo, 2005). Children may struggle academically despite the appearance of average to above average intellectual skills, but these other rule-outs need to be seriously considered first (e.g., intellectual disability, sensory issues, economic opportunity) in order to correctly identify an individual with a SLD. However, even though there may be specific definitions for SLD, these are not always followed and some level of misclassification does occur. For instance, Payette and Clarizio (1994) found that children who were older, from a Caucasian background, and had overall higher academic skills were less

likely to be identified even though a large discrepancy existed between their academic and cognitive skills, while being female and having lower overall academic skills made children more likely to be identified even though a large discrepancy did not exist.

The second point by Fletcher et al. (2007), which states that SLD are dimensional and not categorical, is also important to consider. However, this quality is important for most psychological disorders. People's symptoms do not always fit neatly into the definitions used for various disabilities. Rather people exhibit symptoms of psychological disorders on a continuum, and even the concept of what constitutes a disorder is unclear, even to professionals (e.g., Wakefield, 1992). The dimensional nature of disabilities is a major issue to consider in psychological research, although strict diagnostic criteria are not always followed in research, and the use of a single criterion is commonly used in research on SLDs (e.g., scores below the 20th percentile, Fletcher et al., 2007).

Because SLD can exist on a dimension, it is important to acknowledge that a single score does not define this heterogeneous group, but rather it is a difficulty in learning related to a specific achievement area. Several methods were designed to elucidate what SLD look like on cognitive and academic achievement tests, including the use of simple discrepancy score procedures and profile analysis. However, many of these methods were generally unreliable in discriminating between children with and without learning disabilities (Vellutino, Scanlon, & Lyon, 2000; Watkins, Glutting, & Youngstrom, 2005). The identification of children with learning disabilities has been problematic in the field of school psychology, and several of the methods, criticisms, and future directions will be summarized below.

Cognitive ability tests have gone through a range of interpretive phases, from the initial investigation of global scores, to subtest profile analyses, to a more modern psychometric

approach which has integrated theory and research (Kamphaus, Winsor, Rowe, & Kim, 2005). For SLD, the use of subtest profile analysis has been a commonly used method used by practitioners (Watkins et al., 2005). For instance, one method of profile analysis was to examine the “scatter” of subtest scores, where it was thought that a substantial amount of variability in subtest scores would indicate the presence of an underlying psychological processing deficit. Watkins (2005) showed that different types of subtest scatter (e.g., range, variance, number of subtests which are greater than three points difference from composite score) were not able to reliably discriminate between children with and without SLD. A second method of profile analysis which was also used was based on specific patterns of scores across subtests. One profile in particular, the ACID profile, was used with the Wechsler Intelligence Scale for Children – Third Edition (WISC-III, Wechsler, 1991). This profile was defined as a child having lower scores on Arithmetic, Coding, Information, and Digit Span subtests. Research on this profile has not shown accurate discrimination between groups of children with and without SLD either (Kavale & Forness, 1984; Watkins, Kush, & Glutting, 1997).

The aptitude-achievement discrepancy model is another method which has been used to quantify unexpected underachievement. This model is based on the idea that a child with a SLD will have academic achievement scores substantially below their scores on IQ or aptitude tests. Historically, several methods were used to quantify this discrepancy, from simple differences between scores to regression based methods designed to correct for different correlations between tests as well as measurement error (Kavale, 2002). This method makes sense with the definition of SLD, where a child who appears to have adequate cognitive skills is not achieving as expected, making the discrepancy between cognitive and academic skills a possible method for identifying these children. One issue with this definition of discrepancy is that it does not

account for the part of the SLD definition which states that a person should have a deficit in a psychological process. If a deficit in a psychological process exists, and this process is measured by the test and used in the overall score, then this psychological deficit would lower the overall score of the IQ test. Thus, making full scale IQ—achievement comparisons may not be the most appropriate method for discrepancy models. Based on such difficulties, IQ-achievement discrepancy methods have come under scrutiny and several studies show that a simple IQ-achievement discrepancy does not appear to be adequate for identification purposes (Fletcher et al., 1994; Stanovich & Siegel, 1994; Vellutino et al. 2000).

One of the most important changes in recent years has been the inclusion of response-to-intervention (RTI) models in the schools (Brown-Chidsey & Steege, 2010). This model is important to the evaluation process because it allows practitioners to rule out the possibility that a lack of instruction may be the cause of academic difficulties. RTI has become an important component of the identification process in some models (Flanagan et al., 2010; Flanagan et al., 2006). However, several difficulties arise when RTI is used as the sole method for identification, including the reliability and validity of scores used for progress monitoring (Hale, Kaufman, Naglieri, & Kavale, 2006) and the difficulty with discriminating between individuals who are SLD, low achieving, or may have mild intellectual disabilities (Wodrich, Spencer, & Daley, 2006). The issue of differentiating between those who are low-achieving and those who have SLD is important because research has suggested that setting a low achievement criterion as a definition of SLD actually overidentifies students from disadvantaged backgrounds because it does not control for differences cognitive abilities (McDermott, Goldberg, Watkins, Stanley, & Glutting, 2006). The low achievement criterion is problematic because it does not adequately consider the definition for SLD, which specifies that there should be a deficit in a psychological

process. On the other hand, the RTI process provides an important step in the identification process because it not only shows that the individual is performing below their peers, it also demonstrates that their difficulties in an academic area is not corrected with more intensive instruction.

There has been some extensive debate on the usefulness of both cognitive assessment, RTI, and how they should be used in the identification of SLD. Reschly (2005) provides an overview of the problems with the use of cognitive assessment in the identification of SLDs for both aptitude-achievement discrepancy and intraindividual differences (e.g., pattern of strengths and weaknesses) approaches. He suggests that the use of an RTI process is a separate method which can be (and currently is) used as a method for identifying children with learning difficulties. Hale et al. (2006) explain that the basic principles of RTI are valued in terms of the delivery of services in the schools (e.g., focus on need for all children, continued progress monitoring, single-subject experimental designs, individualized intervention) and this approach is a positive step in the direction of preventing academic difficulties. On the other hand, implementing an RTI approach also introduces many problems. For measurement, the reliability and validity of curriculum-based measurement scores are needed for accurate placement of students in a tiered system like RTI, and for identification, determining what constitutes nonresponse to an intervention is also important to consider. Reynolds and Shaywitz (2009) also discuss many of these issues, and suggest that RTI is a good preventative model, but not adequate for diagnostic decision making. The major problem with this approach is that an RTI process is a service delivery model based on educational need and is not currently based on a strong measurement model designed to differentiate children with and without SLDs based on current definitions (e.g., deficit in a specific psychological process). Data collected during the

RTI process can provide direct evidence of learning difficulties (i.e., continued difficulty despite increasingly intensive and individualized interventions), but this does not identify a deficit in a specific psychological process. RTI may be a good policy for providing services in schools to children who require extra help in academic domains, but it may not be a sufficient process for identifying a child with a learning disability because there is no identification of a deficit in a psychological process, which is necessary based on the definitions of SLD above. The use of RTI may be useful in showing that a child has more difficulty learning than peers, but this difficulty could be due to a host of different problems unrelated to SLD.

Some researchers suggest that the use of RTI and cognitive testing should be used in conjunction with one another in order to best understand the child (Fiorello et al., 2006; Flanagan et al., 2006; Hale et al., 2006; Wodrich et al., 2006). The use of RTI allows practitioners to determine if either lack of instruction has been the cause of academic difficulties. Flanagan and colleagues (Flanagan et al., 2010; Flanagan et al., 2006) have promoted the development and use of a new model of SLD diagnosis based on an operational definition of SLD which includes an examination of academic and cognitive abilities, recurring examination of exclusionary factors (e.g., intellectual disability, second language acquisition, economic disadvantage), and an analysis of strengths and weaknesses (Fiorello et al., 2006). The child must show a normative academic deficit, generally average cognitive functioning, and the difficulties must not be able to be better explained by the exclusionary factors. An important component with this model is the identification of a deficit in one of the psychological processes which is important (e.g., empirically related) to the academic domain in which the discrepancy is occurring. Rather than looking only for a discrepancy between IQ and achievement, this model is more detailed and is designed to pinpoint the cognitive abilities which may be causing the difficulty in achievement.

This model provides a series of logical steps in which different criteria must be met in order to best understand if a child has a SLD. But, as with all models, this method is not likely to be infallible. It is one proposed model, and there may be difficulties with this method yet to be identified. More research will be needed to provide evidence that this method does adequately discriminate between children with and without SLD, but this does provide a structure for identification which is based on both theory and research.

Summary

Children with SLD comprise the largest group of children with disabilities in U. S. public schools, making the understanding of SLD very important to school practitioners and special education. Historically, SLD definitions suggest that a SLD is an unexpected difficulty in an area of academic achievement which is caused by a deficit in a psychological process. Different models of identification have been proposed and tested, and controversy continues on models used for SLD identification. Overall, the joint use of RTI (to help rule out lack of instruction) and cognitive and academic testing (to identify academic and cognitive deficits and strengths) are becoming more important as part of the identification process. The use of cognitive ability tests continues to be important to SLD identification, so a better understanding of their measurement properties and how individuals with SLD perform on cognitive tests is important.

History of Intelligence Theory

The definitions of SLD provided above suggest that a deficit in a psychological process should exist which causes the difficulty in a specific academic area. Cognitive ability tests measure a variety of psychological constructs which may be important to SLD identification, and current theories of cognitive abilities can assist in the understanding of psychological processes which may affect learning. A background on the history of how modern intelligence tests have

been constructed is relevant to the current study, especially the theoretical advances which have affected the development of the WJ-III (Woodcock et al. 2001, 2007), the test being used in the current analysis.

Although important events occurred in the development of intelligence tests from a social perspective (Ciancilo & Sternberg, 2004; Wasserman & Tulskey, 2005), the primary focus of the current review is to examine the psychometric history of intelligence tests. This section provides an overview of the development of psychometric models of intelligence and how they have influenced both the use of intelligence tests and the understanding of the structure of human cognitive abilities.

Spearman and the Two-Factor Theory of Intelligence

Undoubtedly, Spearman's work in the development of factor analysis (Spearman, 1904) is the most important event in the psychometric history of intelligence testing. He noticed that cognitive ability tests tend to be positively correlated, and posited that a general ability factor affects performance on all tests. Additionally, because test scores were not perfectly correlated, then specific processes must also affect test scores. His theory was based on these two independent factors, a general factor (*g*) which affects performance on all tests, and specific factors (*s*) which were unique to each test. He developed a method of factor analysis which extracted a general, or common, factor out of a correlation matrix of different test scores. This common factor was *g*, while any variation left after the general factor was extracted was the variance specific to the test or measurement error (Jensen, 1998). The advancement of factor analysis was important because it provided an analytic approach in which researchers could better understand how cognitive ability tests were related to each other. As Wasserman and Tulskey (2005) note, the idea of partitioning variation in test scores is still used today, only it has

been expanded to include more components. Rather than only examining the general and specific factors, modern tests are often analyzed in terms of the general factor, less general group factors (broad abilities), variation specific to the test itself (not related to the general or broad factors), and measurement error.

The idea of a single general factor was challenged soon after Spearman published his 1904 paper. As Jensen (1998) states, Cyril Burt had reservations about the single general factor and had noticed that, while test scores correlate, tests which seem to measure similar things tend to have larger correlations with one another than other tests. Once other researchers, notably L. L. Thurstone, were able to improve factor analysis to include a wider variety of tests in the analyses, it was apparent that there appeared to intercorrelated, yet separate “general” factors in addition to a higher-order general factor. Rather than a single general factor and many specific factors which accounted for test scores, there appeared to be several different general factors which accounted for performance on groups of particular tests. Spearman eventually accepted the possibility of different group factors as a possibility (Jensen, 1998; Wasserman & Tulskey, 2005).

Thurstone and Primary Mental Abilities

L. L. Thurstone was another important figure in the development of the psychometric methods to analyze intelligence tests. He was instrumental in the development of multiple factor analysis, which was different from Spearman’s method in that it was able to extract several general factors from a correlation matrix (Wasserman & Tulskey, 2005). In a comprehensive examination of different cognitive tests, Thurstone (1938) examined the performance of students on 56 different tests and used this new method of multiple factor analysis to extract group factors, or groups of tests which tended to load on the same factors. He identified seven major

factors, which were remarkably similar to many of the group, or broad, factors recognized by researchers today. These factors included spatial visualization, perceptual speed, numerical facility, verbal comprehension, associative memory, word fluency, and reasoning (Wasserman & Tulskey, 2005). Thurstone initially argued against a general factor based on this analysis (Thurstone, 1938), but his method of factor extraction was designed to create factors which were independent of each other. This method implicitly did not allow a general factor to be extracted because each of the group factors was uncorrelated with one another. Keeping the factors independent of each other, or orthogonal, hid a general factor (Jensen, 1998). Thurstone's method of multiple factor analysis also allows factors to be correlated (oblique rotation), and if substantial correlations exist among the factors, a general factor can often be extracted as well (Carroll, 1993; Jensen, 1998). The ideas of Spearman and Thurstone are essential to the rest of the history of psychometrics and intelligence. Their work brought forth analytic methods as well as a theoretical structure of tests for future researchers to use as a basis of clarifying the factors which emerged through analyses.

Horn and Cattell's Gf-Gc Theory

Another major figure in the development of psychometric and intelligence theory was Raymond B. Cattell, who developed the idea of two different general factors (Cattell, 1943). These factors were crystallized intelligence (Gc) and fluid intelligence (Gf), where Gc was primarily measured by tests based on what has been learned from culture and schooling, and Gf was reasoning ability which was independent of cultural learning. Cattell also hypothesized that differences in Gf also drove differences in Gc, and this investment of Gf in Gc is what produces the correlation between these two factors, from which a general factor emerges. Cattell (1943) thought that fluid intelligence was basically general intelligence, an idea which continues to be

debated today, since Gf is usually highly or perfectly correlated with *g* (Carroll, 1993; Gustafsson, 1984; Kan, Keivit, Dolan, van der Maas, in press).

Collaborating with his student John Horn, Cattell's theory became known as the Gf-Gc theory, and was subsequently extended to include several different broad intellectual abilities (Horn & Cattell, 1966). Abilities most recently outlined by the Gf-Gc theory include visualization (Gv), abilities in listening and hearing (Ga), cultural knowledge (Gc), reasoning abilities in novel situations (Gf), short-term apprehension and retrieval (SAR), long-term storage and retrieval (TSR), speed of thinking abilities (Gs), and quantitative mathematical abilities (Gq; Horn & Blankson, 2005). Horn and Blankson (2005) do not subscribe to the idea that *g* subsumes the broad abilities in Gf-Gc theory. They believe that these abilities may better be described developmentally, or how they change over time. Specifically, they categorize abilities which are vulnerable to decline with age (Gf, SAR, and Gs), those that do not decline with age (Gc and TSR), and sensory related abilities (Gv and Ga).

Other Influential Theories

Vernon (1950) also contributed an interesting model of intelligence. This model was hierarchical with narrow skills at the bottom (called specific factors), minor group factors which subsumed the specific factors, two major factors in the middle, and a general factor at the top. The difference between this model and Gf-Gc is that Vernon categorized the major group factors as Verbal:Educational (*v:ed*) and Spatial, Practical, and Mechanical abilities (*k:m*). The *v:ed* factor is similar to Gc, or abilities which are culture specific, and *k:m* are more similar to fluid abilities, or Gf, which depend less on culture. Although this theory did not have widespread impact, it has regained some recent support through research by Johnson and Bouchard (2005a, 2005b), where a reanalysis of several large datasets provide some evidence for their Verbal-

Perceptual-Image Rotation (VPR) model of intelligence, which was influenced by Vernon's model.

With the empirical evidence mounting that cognitive skills appear to be grouped into broad ability factors which form a hierarchical structure, further research showed that a general factor could be included creating a hierarchical structure of cognitive abilities. In an early examination of the hierarchical structure of intelligence, Gustafsson (1984) examined this possibility, and he found evidence that a set of tests included not only the broad ability factors of verbal, visual, and reasoning. He also found that the reasoning factor was almost perfectly correlated with the general factor, so he argued that the Gf factor is essentially *g*. Carroll's (1993) extensive work helped show that this structure is, in fact, supported by many of the studies on cognitive abilities previously carried out by researchers.

Cattell-Horn-Carroll Theory

A more extensive background of Cattell-Horn-Carroll (CHC) theory (Carroll, 1993, 2005; Keith & Reynolds, 2010; McGrew, 2005, 2009) is provided here because it has become an important taxonomy of cognitive abilities in recent years. Additionally, CHC theory has also been influential in the examination of cognitive—achievement relations (Floyd, et al., 2007; McGrew & Wendling, 2010; Taub et al., 2008), which ultimately has also influenced the understanding of SLD as a deficit in psychological processing (Flanagan, et al., 2010). Because of its impact on current intelligence theory, a more extensive background of CHC theory is provided here.

John Carroll's (1993) book, *Human Cognitive Abilities: A survey of factor-analytic studies*, compiled an extensive reanalysis of over 460 datasets which included a broad range of cognitive ability tests. The objective of this analysis was to compile what is currently known

about human cognitive abilities, reanalyze datasets using a systematic method of exploratory factor analysis, provide a structure to categorize the results, and relate these findings to current findings in cognitive and developmental psychology (Carroll, 1993, p. 74). Carroll's reanalysis was based on a systematic method of exploratory factor analyses (EFAs) which were designed to examine the broader dimensions underlying the variables in datasets form previous studies. The major purpose of reanalysis was to use a single method for all datasets to obtain consistent results across studies, since a variety of extraction and rotational methods were used in original studies. Carroll employed a number of specific decision rules regarding the number of factors and rotation methods across all datasets (see Carroll, 1993, pp.80-90). Through the use of this method, he found many similarities across datasets, which he was able to synthesize into a comprehensive taxonomy of cognitive abilities.

Based on this reanalysis, Carroll (1993) created a model of intelligence which was called "three-stratum" theory, named this way because he found three levels of abilities through his reanalysis. The first stratum included *narrow abilities*, which are specific skills typical of tasks used on psychological tests. For example, the subtests from the WJ-III are designed to measure specific skills representing narrow abilities, which can then be combined to provide composite scores for broader abilities. Dozens of narrow abilities were recognized in this first model (over 60), and these continue to be extended and clarified with further research (Keith & Reynolds, 2010; Schneider & McGrew, 2012). The second stratum consisted of *broad abilities*. At this stratum, several narrow abilities fall under the domain of broad abilities, and Carroll identified approximately 10 of these abilities (Carroll, 2005), most of which are similar to those identified by Gf-Gc theory. It is important to note that each of the broad abilities can influence a large number of tasks, and a variety of these tasks are typically used to measure each of the broad

abilities. Finally, at stratum three was a *general factor*. The general factor, called *g*, subsumes the broad abilities, and it is derived from the covariances among the broad ability factors. Controversy continues over whether this factor is an actual ability or if it is only an artifact of factor analysis (Horn & Blankson, 2005), but convincing evidence supports that a *g* factor is a useful construct for summarizing cognitive abilities, and it is highly related to many academic, social, and physiological measures (Jensen, 1998).

Overall, Carroll's three stratum theory can be viewed as a synthesis of previous theories into a comprehensive taxonomy of cognitive abilities based on extensive empirical analysis. The three-stratum theory brings together aspects of Horn and Cattell's Gf-Gc theory (Cattell, 1943, 1963; Horn & Blankson, 2005; Horn & Cattell, 1966), in which several broad abilities explain the correlations among more narrow abilities in a consistent manner, Thurstone's (1938) primary mental abilities, and Spearman's (1904) two-factor theory which included a general factor (*g*) and specific factors (unique to each test).

Development of Modern CHC Theory

Carroll's analysis provided an impressive summary of research on cognitive abilities. It should be noted, however, that this theory appeared to be in the "wings," and Carroll provided the analysis needed to solidify the thoughts which were developing regarding the structure of human cognitive abilities (McGrew, 2005). As stated above, Horn and Cattell's Gf-Gc theory had identified several abilities which were identified in Carroll's second stratum, including visual intelligence, auditory intelligence, and cognitive processing speed (Horn & Cattell, 1966), and these were similar to Thurstone's (1936) primary mental abilities. Similarly, Gustafsson (1984) had begun analyzing cognitive abilities in a hierarchical manner using the fairly new methods of confirmatory factor analysis (CFA) and structural equation modeling (SEM). As McGrew (2005)

recalls, when he was working on the revision of the Woodcock-Johnson Psycho-Educational Battery (Woodcock & Johnson, 1977), both Horn and Carroll were asked to be consultants on this project and for the first time brought together a theoretical structure to the revision of a major test battery. Based on the direction of the field, it appeared as though a synthesis such as Carroll's was inevitable, providing a comprehensive synthesis which has been a major influence on cognitive ability tests ever since (Alfonso et al., 2005).

McGrew's (1997, 2005, Schneider & McGrew, 2012) taxonomy and nomenclature has become the common framework when discussing CHC theory. Schneider and McGrew (2012) recently provided revised definitions based on current research to help clarify the definitions of the CHC abilities. The seven most commonly examined broad abilities are described below:

1. *Fluid Intelligence/Reasoning (Gf)* – Gf is comprised of the ability to solve novel problems. Gf includes inductive and deductive (sequential) reasoning, as well as quantitative reasoning (use of induction or deduction with numbers or mathematical symbols). Schneider and McGrew suggest that inductive reasoning is at the core of Gf.
2. *Short-Term Memory (Gsm)* – Gsm is the ability to hold and manipulate information held in memory. Gsm includes two primary narrow skills, memory span (the amount of information which can be held) and working memory (ability to both hold information and make simple manipulations or transformations of the information in memory).
3. *Long-Term Storage and Retrieval (Glr)* – Glr is a person's capacity to retrieve information which has been learned previously, where the period must be longer than that required for the information to be stored in Gsm. Two narrow skills make up Glr: learning efficiency (associative memory, meaningful memory, free recall memory) and retrieval fluency (includes a number of skills requiring the fluent recall of information in

long-term storage). Schneider and McGrew (2012) indicate that retrieval fluency appears to be distinct from processing speed, but it is necessary to include enough measures of retrieval fluency in order for a unique factor to emerge which is distinct from processing speed. The retrieval fluency factor, however, may also depend on processing speed, making it a complex ability.

4. *Processing Speed (Gs)* – Gs is the ability to complete simple cognitive tasks quickly. Perceptual speed, which is based on quickly identifying similarity and difference between stimuli, is thought to be the foundation of Gs (Schneider & McGrew, 2012). These tasks are usually simple, which may include recognizing patterns or visually scanning information for similarities and differences.
5. *Crystallized Intelligence/Knowledge (Gc)* – Gc represents an individual's knowledge and skills which have been acquired by culture. This knowledge includes a broad range of skills including general knowledge, language development, listening ability, and oral language. Some recent research has questioned the distinctiveness of Gc as a separate broad ability, rather, it is based primarily on Verbal Comprehension (Kan et al., in press). Most intelligence batteries include tests which measure Gc, and more work needs to be done to clarify what Gc is (Keith & Reynolds, 2010).
6. *Visual Spatial Abilities (Gv)* – Gv ability includes skills related to the mental visualization and transformation of visual patterns or objects. This includes visualization (transforming or imagining how objects look when they are changed), visual memory, and visual closure (being able to identify objects when it is partially covered or obstructed) are Gv narrow abilities

7. *Auditory Processing (Ga)* – Ga represents an individual's ability to process auditory stimuli. This ability is most important for understanding speech sounds in language development, and includes skills such as phonetic coding, speech sound discrimination, and resistance to auditory distortion.

These seven broad abilities are most commonly measured by major tests of cognitive abilities, including the WJ-III. Some other broad abilities are more sensory-based (e.g., Olfactory abilities (Go), Kinesthetic abilities (Gk)), but these are not often assessed by cognitive batteries. The seven cognitive abilities defined above will be the focus of the current study.

Impact of CHC Theory on Current Tests of Intelligence

CHC theory pulled together much of the previous theoretical work on intelligence into a comprehensive taxonomy of cognitive abilities which has strong empirical support. CHC theory has greatly affected the measurement of intelligence, both at the theoretical and practical levels. First, it has provided a common language for researchers and practitioners. Researchers may have been studying the same types of cognitive processes, but may have used different terms to refer to the same processes. Practitioners previously were tied to using and interpreting different tests separately, even though many of these tests were found to be measuring the same processes. CHC theory has provided a universal language to talk about the constructs measured by the wide variety of cognitive ability tests available for research and practice.

Secondly, CHC theory has directly influenced the development and revision of modern intelligence batteries (Alfonso, et al., 2005; Keith & Reynolds, 2010). The Woodcock-Johnson – Revised (Woodcock & Johnson, 1989) was the first test to be designed specifically using Gf-Gc theory (McGrew, 2005), and this influence has continued with the most current revision of the test (Woodcock et al., 2001, 2007). Some tests, such as the Reynolds Intellectual Assessment

Scales (RIAS, Reynolds & Kamphaus, 2003) and the Wide Range Intelligence Test (WRIT, Glutting, Adams, & Shelslow, 2002) were based partially on CHC theory in their initial development (Alfonso, et al., 2005). Many major intelligence batteries have used CHC theory explicitly in their revisions, including the Kaufman Assessment Battery for Children—Second Edition (KABC-II, Kaufman & Kaufman, 2004) the Stanford-Binet—Fifth Edition (SB-5, Roid, 2003), and the Differential Ability Scales—Second Edition (DAS-II, Elliot, 2007). As mentioned by Keith and Reynolds (2010) and Alfonso et al. (2005), the Wechsler scales (e.g., Wechsler Intelligence Scale for Children—Fourth Edition, Wechsler, 2003) have not been influenced as highly by CHC theory, although CHC theory is recognized by the current authors as an important theoretical structure, and the Wechsler scales have been interpreted from a CHC perspective (Benson, Hulac, & Kranzler, 2010; Keith, Fine, Taub, Reynolds, & Kranzler, 2006; Ward, Bergman, & Hebert, 2011). It should be noted that while the impact of CHC theory has been relatively widespread, it is still a theory in the works and continued work and clarification of broad and narrow abilities are needed (Carroll, 2005; Keith & Reynolds, 2010; McGrew, 2005, 2009; Schneider & McGrew, 2012).

Broad abilities have been the focus of a large amount of research concerning CHC theory and modern intelligence batteries (Benson, 2007; Benson, et al., 2010; Evans, Floyd, McGrew, & Leforgee, 2001; Floyd, McGrew, Barry, Rafael, & Rogers, 2009; Keith, et al., 2006; Keith, Low, Reynolds, Patel, & Ridley, 2010; Keith, et al., 2008; Kranzler, Keith, & Flanagan, 2000; Reynolds & Keith, 2007; Reynolds, Keith, Ridley, & Patel, 2008; Reynolds, Keith, Fine, Fisher, & Low, 2007; Taub, et al., 2008). However, broad abilities subsume more narrow cognitive abilities. This distinction is important to consider for SLD, since part of the definition for SLD is having a deficit in a psychological process. Differences in a psychological process may be at the

broad ability level or narrow ability level, and understanding the measurement of broad and narrow abilities is important if cognitive tests are going to be useful in the identification of SLD. In their review of the research on relations between CHC abilities and academic achievement, McGrew and Wendling (2010) suggest that the narrow ability level is the place at which the "primary action" between cognitive abilities and academic achievement takes place (p. 669). Their analysis shows that broad cognitive abilities can best predict broad academic skills, but when examining more narrow academic skills (e.g., reading decoding skills), more narrow cognitive abilities are better predictors.

In terms of SLD assessment, this suggests that narrow abilities may need to be a major focus during the evaluation process. Some studies have begun to identify which cognitive ability deficits are present in SLD populations (Compton et al., in press), and these are instructive in determining which processes may need to be assessed most carefully in the identification process. O'Shaughnessy and Swanson (1998) completed a meta-analysis to examine the role of memory and recall in SLD in reading. Their analysis showed that memory deficits were a persistent difference in children with SLD across childhood, where the difference in performance on memory tasks is not reduced as children get older. The difference in memory performance is best thought of as a deficit as opposed to a developmental lag. Importantly, Hoskyn and Swanson (2000) showed that there are also differences in some cognitive areas between low-achieving and SLD students. They found that both low-achieving and SLD groups had deficits in phonological processing, but children with SLD had an advantage in verbal intelligence, specifically in syntactic knowledge. A more recent meta-analysis (Johnson, Humphery, Mellard, Woods, & Swanson, 2010) has supported the notion that several cognitive abilities show reliable differences between groups of children with and without SLD. Johnson et al. showed differences in

processes such as phonological processing, processing speed, executive functioning, and memory show consistent differences when aggregated across studies.

Most modern intelligence tests include a range of narrow skills to measure broad abilities. Few include enough tests to provide an adequate measure of narrow abilities as stand-alone composites. Most tests include a range of skills from which to created composite scores for broad abilities, but not often for narrow abilities. For instance, the WJ-III uses Numbers Reversed (a working memory task) and Memory for Words (a short-term memory task) to measure Gsm. But, it also includes Auditory Working Memory (a second working memory task) and Memory for Sentences (a second memory span task). By including two tests for these narrow abilities, it is also possible to obtain scores for the narrow abilities Memory Span and Working Memory, in addition to the broad ability Gsm.

Since many tests do not include a substantial range of tasks, the practice of cross-battery assessment, or XBA, has been promoted as a method for more comprehensive assessment of narrow abilities (McGrew & Flanagan, 1998; Woodcock, 1990). With XBA, it is possible to use subtests from different intelligence batteries to gather more information about an individual's performance on a particular narrow ability. Many intelligence batteries have been subjected to joint confirmatory factor analyses (CFA) to examine which broad and narrow abilities are included in the tests (Flanagan & McGrew, 1998; Keith, Kranzler, & Flanagan, 2001; Sanders, McIntosh, Dunham, Rothlisberg, & Finch, 2007). Such joint CFAs have provided evidence that tests from different intelligence batteries can be used in conjunction with one another to help adequately measure constructs of interest. Several XBA studies, as well as exploratory and confirmatory factor analyses with the WJ-III have supported the validity of the WJ-III constructs, as reported in the technical manual of the WJ-III (McGrew & Woodcock, 2001).

Although broad abilities have been a main focus of a large amount of research, fewer studies have focused on the narrow abilities measured by intelligence batteries. When the WJ-III was developed, a number of factor analyses were included in the technical manual showed how the tests included on the WJ-III load on both broad and narrow CHC abilities (McGrew & Woodcock, 2001, pp. 191-209). Some studies have included narrow abilities in the analysis, such as Phelps, McGrew, Knopik, and Ford (2005), who completed a joint CFA with the WJ-III and the WISC-III. Floyd et al. (2007) examined the relationships between CHC narrow and broad abilities and reading decoding, and including the narrow abilities helped identify some of the more specific processes within the broad abilities which were important for reading decoding. For example, when analyzing the data for children between the ages of 7 and 8, Gc was a significant predictor of reading decoding when only broad abilities were included in the model. When the narrow abilities were included in the model as well, Gc no longer had a direct effect on reading decoding, the narrow skill of Listening Ability had a direct effect on reading decoding while effects from Gc were indirect. Including the narrow abilities provided more precision in identifying which skills are important for reading decoding, providing evidence to the claim by McGrew and Wendling (2010) that narrow achievement skills (e.g., reading decoding) may be best explained by narrow cognitive abilities (e.g., listening ability).

Summary

The history of intelligence testing and the psychometric traditions behind it have been a cumulative effort across the 20th century. Through the work of Spearman, Thurstone, Cattell, and Carroll, modern theories of intelligence have a strong empirical background that closely resembles the theoretical structure of cognitive abilities. Improved theory has been an important component to a better understanding of SLD because it has clarified the nature of psychological

processes which may be important to learning, and this is important to consider in the current study because one of the purposes is to examine the similarities and differences in cognitive-achievement relations between groups of children with and without SLD.

Relations between Cognitive Abilities and Academic Skills

One important consequence of CHC theory is a better understanding of the relations among cognitive abilities and academic achievement. Since the advent of CHC theory, researchers have begun to examine the correlations between cognitive abilities and academic achievement skills to better understand what processes are used for different academic areas. Much research remains, but at this point a relatively strong understanding of the CHC abilities which underlie academic skills has begun to emerge. The following section will review the research which has examined cognitive processes related to academic skills with a focus on research which has used CHC theory as a basis.

The research for the effects of CHC abilities for reading and math have been summarized comprehensively by McGrew and Wendling (2010). They examined two constructs of reading, Basic Reading Skills (BRS) and Reading Comprehension (RC). For BRS, they found that Gsm, Gs, Glr (only for ages 6 through 8), Gc, and Ga all had significant relationships with BRS. A study by Floyd et al. (2007) which examined the relations between CHC broad and narrow abilities and reading decoding found relationships from all of these abilities mentioned in McGrew and Wendling's analysis, although these relations were not consistent across age ranges. For example, Gs is related to reading decoding for children ages 5 to 8, but not for older children.

A recent study by Elliot et al. (2010) used the DAS-II to examine cognitive-achievement relations for children with and without reading difficulties. They found that Gc, Gsm, Ga, and Gf

all had significant effects on basic reading skills in typically developing children, but in children with reading difficulties the significant effects from cognitive variables were from Ga, Gv, Glr, and Gs. All effects from *g* were mediated through the broad abilities for both groups, suggesting the importance of considering cognitive abilities beyond *g*. They consider that the reason for differences in processes may be related to compensatory strategies used by individuals with reading difficulties. For example, they suggest that the significant effect from Gv to reading may be due to the fact that children with reading difficulties may be using more of a whole word reading approach than a phonics approach (Weekes, Coltheart, & Gordon, 1997).

Some other influential researchers have suggested that specific cognitive processing deficits are especially important for reading SLD. Although many of these studies have not been based primarily on a CHC model, they can be interpreted in light of CHC cognitive abilities. One theory is the double-deficit hypothesis (Bowers & Wolf, 1993), which states the reading difficulties come from a deficit in both phonological processing (Ga), and rapid automatized naming (RAN, a Glr and Gs ability, Schneider & McGrew, 2012). Bowers and Ishaik (2003) summarize this research and showed that children with a deficit in phonological processing or rapid naming tend to have lower reading scores compared to those who do not have deficits in either of these areas, but the differences were much more pronounced for learners with a deficit in both phonological processing and rapid naming. Phonological processing has been repeatedly shown to be related to reading, since reading requires phonics skills for decoding. Bowers and Ishaik (2003) hypothesized that RAN abilities were related to reading through orthographic skills, which is the ability to use patterns of letters for reading. Children who have greater difficulties with recalling orthographic codes also have greater difficulties with reading fluently.

Swanson and colleagues (e.g., Swanson, 1993; Swanson & Alexander, 1997; Swanson, Kehler, & Jerman, 2010) have focused on working memory deficits as a causal mechanism for SLD in general. Working memory is consistent with Gsm from CHC theory, which consists of Working Memory (ability to hold and manipulate information in memory) and Memory Span (amount of information which can be held in memory). In fact, research has suggested that both memory span and working memory are uniquely related to reading difficulties, each contributing to deficits in reading (e.g., Swanson & Ashbaker, 2000). The differences in memory have also been replicated in more recent research on cognitive profiles for children with SLD in reading (Compton et al., in press).

Overall, processes related to Gsm, Gs, Gc, and Ga appear to be especially important for reading decoding. However, little work has examined where differences may exist when examining reading in groups of children with SLD in reading. Elliot et al. (2010) suggested that poor readers may rely on different processes, and other research has demonstrated differential predictive power of some processes over others. For example, Swanson and Alexander (1997) found that phonological awareness was the best predictor of decoding skills for good readers, but general intelligence was the best predictor of this skill for poor readers. Clearly, more work needs to be completed to better understand where these differences lie.

In their review of the literature, McGrew and Wendling (2010) found that math calculation skills were related to Gsm (specifically working memory), Gs, Glr (specifically RAN skills), Gc, and Gf. Math Reasoning was related to all the same processes except for RAN. Taub et al. (2008) found that *g* was not a direct predictor of mathematics skills, rather the broad abilities were better predictors and any effects of *g* were indirect. This finding is consistent with the proposition that narrow cognitive skills are better predictors of narrow academic skills (e.g.,

Floyd et al., 2007; McGrew & Wendling, 2010). Proctor and colleagues (Proctor, in press; Proctor et al., 2005) have examined cognitive ability differences in mathematics from a CHC perspective. Proctor and colleagues (2005) found that the cognitive profile of low achievers in math calculation was similar to those without deficits in math calculation, although deficits in math reasoning were concurrent with deficits in Gf and Gc. Proctor (in press) examined the relations between CHC abilities and math calculation and math reasoning with college students who have been diagnosed with a SLD in math. For math calculation, Gs was the only significant predictor, whereas for math reasoning the significant predictors were Gc, Gf, and the narrow ability of Working Memory. Again, these results suggest that important differences in cognitive—achievement relations may apply within mathematics difficulties. Overall, studies have shown that math abilities tend to be related to Gf, Gc, Gsm (specifically working memory), Glr, and Gs. More work is needed to clarify these findings, but overall consensus across studies is that the processes mentioned previously are important to math achievement.

Little substantial work has examined the relationship between cognitive abilities and writing. Writing was not included in McGrew and Wendling's (2010) review of cognitive—achievement relations, and it appears as though Floyd and colleagues (2008) is the only study in which the relations between CHC abilities and writing were examined. They specifically examined basic writing skills (which includes Spelling, Punctuation and Capitalization) and written expression (which includes Writing Fluency and Writing Samples). Gc was the best predictor of basic writing skills, and Gs and Gsm were also important across childhood. For younger children, Glr and Ga were important, but this dropped off quickly with age. For adolescents, Gf became more important with age as well. For written expression, again Gc, Gs,

and Gsm all were important, and Glr and Ga were important for younger children, but this relation again decreased in importance quickly with age.

Written expression appears to be more complex than other academic skills because of the inclusion of the visual-motor domain. Fletcher et al. (2007) suggest that this is especially important for handwriting and spelling, which require the use of both motor skills and orthographic processing, which Stanovich and West (1989) define as "the ability to form, store, and access orthographic representations" (p. 402). They showed that orthographic processing was significantly related to reading skills, but this is not currently a part of CHC theory. Orthographic processing may be a visual skill primarily necessary for reading, although the integration of orthographic and motor skills would be necessary for writing. In some recent research with the DAS-II, Niileksela and Reynolds (manuscript submitted for publication) found that children with SLD in reading and writing showed differences in *g*, Gs, and specifically on the subtest Recall of Designs, where the examinees are required to look at a design and recreate it on their own. This study's findings suggested that orthographic processing, possibly in the form of visual memory, may play a role in writing difficulties. The finding was of particular interest, because the DAS-II was not used as part of the diagnostic process, it was administered to those who were already identified as having a SLD. There has not been a substantial amount of work examining CHC abilities and writing, making the current investigation an important contribution to the understanding of cognitive—achievement relations and writing.

Summary

Researchers continue to identify areas related to achievement in normally developing children and areas which may potentially cause academic deficits in children with SLD. Thus, it is important to continue to examine both the relations between cognitive skills and academic

achievement, and it is similarly important to continue to examine whether these relations are the same between groups of individuals with and without SLD. Using strong research methods and adequate measurement instruments is essential. The next section includes an overview of the methods to be used in the current study, namely confirmatory factor analysis (CFA) and structural equation modeling (SEM). Also included is a review of the basic assumptions of this model, as well as a review of the methods which can be used for multiple group comparisons.

The Common Factor Model

Dozens of techniques are available for data analysis. One of the most prominent methods available in social sciences for understanding the underlying structure of a set of variables is factor analysis, although it has often been misused and misunderstood (Preacher & MacCallum, 2003). Factor analysis has been one of the most commonly used methods for evaluating the structure of cognitive test scores (Carroll, 1993). Thus, it is important to provide an overview of the logic and assumptions of the common factor model, how it is used to analyze the structure of cognitive tests, and how it can be used to examine the relations between cognitive abilities and academic achievement.

Factor analysis assumes that the correlations among observed variables, such as a battery of cognitive tests, can be explained by a smaller set of latent, or unobserved, variables. For intelligence testing, this assumes that the latent broad and narrow cognitive abilities (as well as *g*) discussed above affect an individual's performance on specific cognitive tests. Two major types of factor analysis are available for researchers, which are *exploratory factor analysis* (EFA) and *confirmatory factor analysis* (CFA). Although a variety of techniques can be used for EFA (Mulaik, 2010), the primary focus here will be on CFA. EFA is, as suggested by its name, an exploratory technique which is designed to uncover latent factors which may reflect common

causes of the relations among observed variables (Mulaik, 2010). One major difference between EFA and CFA is that EFA extracts factors based on the correlations among the observed variables, and the pattern of factor loadings for the variables is determined completely by the data. In contrast, CFA analyzes covariances and uses a prior theoretical orientation or hypothetical structure which guides the decision regarding the number of factors which underlie the data and which variables load on different factors (Keith, 2006). This distinction is important because covariances should be analyzed whenever studying groups who may be subsets of a population (e.g., individuals with SLD) because the correlational structure may be altered via the selection of the subgroups. Hence, covariance matrices, not correlation matrices should be analyzed in this method.

CFA is based on the common factor model, which assumes that a number of latent variables (unobserved variables) explain the covariances between the observed variables (Krane & Slaney, 2005). In this technique, the number of common factors is less than the number of observed variables. For example, if a group of students take three tests which measure reading skills and three tests which measure math skills, one may hypothesize that two common factors underlie the scores on these tests, a “Reading” factor and a “Math” factor. Even though six test scores were included, in the common factor model assumes the covariation between these tests is explained by two common factors.

Mathematically, the common factor model is described by the following equation:

Equation 1: $$\mathbf{y} = \boldsymbol{\tau} + \boldsymbol{\Lambda}\boldsymbol{\eta} + \boldsymbol{\varepsilon}$$

Assuming six variables and two factors in this model, $\mathbf{y}$ is a vector of n x 6 observed scores for each variable, $\boldsymbol{\tau}$ is a vector of 6 x 2 intercepts for the observed variables, $\boldsymbol{\Lambda}$ is a 6 x 2 matrix

factor loadings for the latent variables, $\boldsymbol{\eta}$ is a 2 x 2 matrix of latent variable means, and $\boldsymbol{\varepsilon}$ is a 6 x 6 matrix of residual variances not accounted for by the latent variables in the model.

In the CFA framework, the goal is to specify a theoretical model which adequately recreates the observed covariance matrix. Then, model fit is evaluated by examining the differences between the covariance matrix derived from the model and the observed covariance matrix. The model implied covariance matrix is obtained through the following equation:

Equation 2: $$\boldsymbol{\Sigma} = \boldsymbol{\Lambda\Phi\Lambda}^{-1} + \boldsymbol{\Theta}$$

In this equation, $\boldsymbol{\Sigma}$ is the model implied covariance matrix, $\boldsymbol{\Lambda}$ is the matrix of factor loadings, $\boldsymbol{\Phi}$ is the variance-covariance matrix of the latent variables, $\boldsymbol{\Lambda}^{-1}$ is the inverse of the matrix of factor loadings, and $\boldsymbol{\Theta}$ is the matrix of residual variances of the observed variables (Brown, 2006). Once a model is hypothesized and created, the model implied covariance matrix, $\boldsymbol{\Sigma}$**,** is compared to the observed variance-covariance matrix, **S**. Therefore, the main test is $\boldsymbol{\Sigma} = \mathbf{S}$, where covariances obtained through the model parameters is compared to the covariances obtained from the observed data.

Different indices can be used to determine how well the two matrices approximate each other, such as χ^2, the comparative fit index (CFI), root mean square error of approximation (RMSEA), and standardized root mean square residual (SRMR; Brown, 2006; Kline, 2011). If these two matrices are similar to one another, the fit indices will provide a quantitative index of how similar they are. If they are substantially different from one another, the fit indices will indicate worse fit. It is important to note that fit indices are affected by many extraneous factors, such as sample size, the number of factors, number of variables, and complexity of the model (Brown, 2006; Hu & Bentler, 1998, 1999; Kenny & McCoach, 2003; Kline, 2011), thus the

reporting of several fit indices is the standard practice in CFA (Schrieber, Stage, King, Nora, & Barlow, 2006; Jackson, Gillaspy, & Purc-Stephenson, 2009).

In addition to specifying which variables load on specific factors, one can structure the nature of the relations among latent variables, creating a structural equation model (SEM). With CFA, all of the latent variables are correlated with each other. In a SEM, the researcher specifies which latent variables are related to each other and how they are related to each other (Keith, 2006; Kline, 2011). For instance, using the example above where students completed the math and reading tests, one could also give the students an academic motivation scale. He or she could specify that the reading tests load on the reading factor, the math tests load on the math factor, and the items from the academic motivation scale load on an academic motivation factor. With typical CFA, covariances would be included between the latent factors. With SEM, regression paths are used in lieu of covariances, where some latent variables may be specified as the causes or predictors of other latent variables. For the above example, we might draw regression paths from the academic motivation factor to the reading and math factors, creating a model which implies that academic motivation is the "cause" of reading and math scores. Likewise, researchers can test competing models, such as allowing reading and math to act as causes for academic motivation, or showing whether the effects of academic motivation on math and reading are equal or if they are statistically significantly different from each other. Based on the assumption that a model which better fits the data is likely to be a better explanation of real-world phenomena, models can be evaluated based on the adequacy of fit indices, theoretical relationships, and prior research. However, conclusions must be made based on strong previous research, theoretical bases, and adequate analysis of other possible rival explanations.

Overall, CFA and SEM are powerful and flexible techniques which can be used to evaluate the structure of a large number of variables. CFA is especially powerful because it uses theory and prior research and a hypothetical structure to guide the underlying measurement of latent variables, and the step from CFA to SEM allows one to create causal models based on the data as well.

Multiple Group Analysis in SEM

A large proportion of psychological and educational research is based on statistics designed to compare group means to each other (e.g., ANOVA, *t*-tests), or to examine the relations between a number of observed variables (e.g., correlation, regression). In a typical experiment, participants are randomly assigned to groups (e.g., experimental and control groups), they experience some type of environmental manipulation (e.g., no training vs. extensive training), measurements are taken on their performance or on some psychological construct, and groups are compared using traditional statistical methods. Similarly, when experimental manipulations are not possible, data may be collected on a number of variables and naturally occurring groups (e.g., gender, ethnicity, or cultural status) can be compared to determine if differences exist between groups on the constructs of interest. If differences exist, then the assumption is that individuals from different groups, on average, tend to have different levels of the psychological construct of interest. The analysis of multiple groups is also available in SEM, and this powerful method can be used to not only compare group means, but also compare the similarities and differences in factor loadings, factor structure, and relationships between latent variables between different groups.

The multiple group model uses a separate covariance matrix for each group, but estimates all parameters simultaneously. First introduced by Jöreskog (1971), this method allowed

researchers to examine the same factor model with different groups, and also to test whether the constructs are the same by applying equality constraints to model parameters. This method is very flexible because it can allow groups to have completely different factor models, where different variables load onto different factors for each group, or it can require that the same factor model is used in both groups, and all parameters can be constrained to be equal across all groups. Jöreskog (1971) only included the covariance structure in his work, and Sörbom (1974) extended this work to include latent variable means as well. The major use of multiple group analysis in practice is to test for measurement invariance, where model parameters are sequentially tested for equality across groups to ensure that the constructs being measured are the same. More formally, multiple group analysis tests whether or not differences in observed scores across groups are explained by differences in the common factors. Measurement invariance is a necessary step if groups are to be validly compared on latent or observed variable means (Meredith, 1993; Meredith & Teresi, 2006). However, less strict forms of measurement invariance can allow for the comparison of other model parameters. For example, weak factorial invariance, where factor loadings are equal across groups, is necessary for valid comparisons of latent variable variances and covariances across groups, which will be the focus of this study.

Multiple group analysis is not only useful for testing parameters in the measurement model, it is also useful for testing whether or not the relations between latent variables are the same across groups (Brown, 2006). For example, one may want to know if reading self-efficacy predicts reading achievement equally for males and females. In a multiple group model, the parameters for males and females will be estimated separately to see if this relation is significant. The relation between reading achievement and reading self-efficacy may be significant in both groups, but it is then also possible to test whether this relation is equal between groups. This test

is done by requiring the unstandardized regression path from reading self-efficacy to reading achievement to be equal in both groups. If there is a statistically significant decrease in model fit, then it is an indication that this relation is moderated by sex. But, if there is not a statistically significant change in model fit, then it indicates that this relation is the same for both males and females.

In summary, the use of multiple group models can be especially useful in SEM because of the ability to examine all model parameters. In the next section, the use of multiple group models will be used to identify where groups may differ on specific parameters in the factor model.

Measurement Invariance

Psychological and educational measures (e.g., intelligence and achievement tests, personality and behavior rating scales) are based on the assumption that the measures are adequately measuring the same constructs across different groups, and adequate measurement is an important component of the research process. That is, if a researcher is interested comparing males and females in levels of depression based on responses to a depression rating scale, then the researcher is making the assumption that the scale is measuring depression in the same way across groups. If the scale is adequate, then a particular score on the depression scale for people in each of the groups should represent the same amount of depression for individuals in different groups.

Millsap (2011) provides an example of this based on physical measurement. We assume that two objects with the same mass would weigh the same on a scale regardless of their shape. Imagine that two spheres which have the same mass are shaped into a pyramid and a cube, respectively, without removing any matter from the objects. When weighed as spheres, they had

the same weight. When they are weighed a second time as a pyramid and a cube, the pyramid weighed more than the square. Because we know that the underlying mass of the two objects should be the same, and shape should not affect mass, the scale can be said to be biased because it provides different weights based on a separate factor other than actual mass (shape). In this case, the shape affects the measurement even though shape does not directly affect mass. The scale is not measurement invariant between objects of different shapes.

Relating the physical example to the depression example, the mass of the objects can be thought of as the underlying construct of depression. The scale can be thought of as the depression questionnaire, and the shapes of objects can be thought of as different groups. If two people from different groups have the same amount of an underlying construct (depression, in this case), and the depression scale provides different scores for people who belong to different groups, then the scale is not measurement invariant. *Measurement invariance* refers to the property of a test or scale where people who have the same level of a construct also obtain the same score on a measurement instrument, regardless of group membership (Meredith, 1993; Meredith & Teresi, 2006; Millsap, 2011). If a depression scale is measurement invariant, then males and females who have the same level of depression should also obtain the same score on the scale.

Measurement invariance is an important concept to understand in psychological and educational measurement. When groups systematically differ on a measure, this is referred to as *population heterogeneity* (Muthén, 1989). Often group comparisons are made based on the assumption that the groups come from the same population even though different subsets of individuals of the population may differ in how psychological or educational constructs are measured. Muthén (1989) states that *population homogeneity*, where measurement instruments

are the same between groups, is an unrealistic assumption to make when comparing groups in applied research. Thus, methods for detecting and controlling for differences in measurement are important.

Measurement invariance has become an important component of large scale psychological research (Millsap & Meredith, 2007). Several important reviews of measurement invariance principles and practice have been published in the last 20 years (Horn & McArdle, 1992; Little, 1997; Meredith, 1993; Millsap, 2011; Vandenberg & Lance, 2000; Widaman & Reise, 1997), and these reviews have outlined both the importance of measurement invariance in social science research as well as the steps which should be taken by researchers when testing for measurement invariance. Measurement invariance tests are typically carried out in a hierarchical manner, where cross-group equality constraints are added in a systematic fashion to the model. If constraints on one part of the model meet the requirements for measurement invariance (e.g., do not lead to substantial change in model fit) then those constraints are left in place as constraints are added to other parameters in the model. The steps which are typically followed in testing measurement invariance are presented below. For purposes of this study, the most important tests of invariance will be similar to weak factorial invariance, where *the factor loadings* will be tested for equality prior to comparisons of structural relations between latent variables.

Testing homogeneity of variance-covariance matrices. The first step recommended by Jöreskog (1971) and Vandenberg and Lance (2000) is to test the equality of the variance-covariance matrices across groups. This step is completed before any particular factor model is created, and it is tested by constraining the variances and covariances between observed variables in all groups to be equal. If this test is not statistically significant (according to χ^2, with excellent values on other fit indices), then it can be assumed that the covariance matrices between groups

are essentially equal. Jöreskog (1971) suggested that no other invariance tests needed to be completed if the covariance matrices are equal, since all analyses in CFA and SEM are contained within the covariance matrices.

Configural Invariance. The second step in testing for measurement invariance is referred to as configural invariance. With this model, the hypothesized factor structure is tested simultaneously for both groups, where the pattern of fixed and free factor loadings is the same within each group. None of the model parameters, other than those used for identification purposes, (e.g., factor loadings set to 1), is required to be equal across groups. This step does not make any comparisons regarding the size or equality of the factor loadings (λ) between groups, only that the pattern of factor loadings is the same between groups. If configural invariance is not tenable, then the groups cannot be compared because the underlying factor pattern is not the same across groups. If configural invariance is tenable, the groups can be said to have the same underlying structure on the variables (Vandenberg & Lance, 2000).

Weak factorial invariance. Also called metric invariance (Horn & McArdle, 1992), testing for weak factorial invariance goes beyond simply using the same pattern of fixed and free factor loadings and includes equality constraints on corresponding factor loadings across groups. By constraining the corresponding unstandardized factor loadings equal across groups, this step tests whether or not the latent variables have the same effects on the observed variables across groups. If the corresponding factor loadings from the observed variables on the latent variables are not the same across groups, then the measurement instrument is not measuring the same construct and group membership moderates the effect of the latent variable on the observed variables (Meredith & Teresi, 2006). If weak factorial invariance is tenable, then it is possible to move on to the next step. If there is a significant change in model fit by including these constraints, then it

basically means that one or more of the factor loadings are significantly different between groups.

Strong factorial invariance. Also called scalar invariance, when testing for strong factorial invariance equality constraints are added to the observed variable intercepts in addition to the factor loadings. This step is testing whether the value where the observed variable meets the intercept (mean) on the latent variable is the same across groups. Similar to regression, this step is determining if one group has a systematically higher or lower score on the *y*-intercept of the predicted variable. If they are higher or lower, then this means that the latent common factor mean between the two groups cannot be meaningfully compared because one group will show a systematic advantage over the other in a specific factor mean. A lack of strong factorial invariance is essentially a form of uniform differential item functioning (DIF, Osterlind, 2009) where a person's score on a test or item depends not only on their ability, but also on group membership. If the factor loadings and measurement intercepts are the same across groups, then mean differences on observed measures are attributed to differences in the latent variable.

Strict factorial invariance. Although strong measurement invariance is considered adequate for comparing latent or observed variable means, it is also possible to continue testing measurement parameters in the model. Strict measurement invariance is a step which places equality constraints on the residual variances of the observed variables. Residual variance is composed of two types of variance, specificity and measurement error (Meredith & Teresi, 2006; Mulaik, 2010). Specificity refers to the variance which is systematic and reliable, but not attributable to the common factor, and error refers to unsystematic variance attributed to measurement error. By holding the residual variances equal across groups, strict measurement invariance is determining whether the amount of variance not accounted for by the latent variable is equal across groups. If

the amount of variance is not the same, it may mean that the groups differ in the reliability of the measured variable (e.g., there was more measurement error for one group over another) or there is a greater amount of systematic variance which may be attributable to a different variable for one group over another. Widaman and Reise (1997) state that, if weak, strong, and strict invariance constraints can be added without a significant degradation in model fit, then all group differences in both means and variances of the observed variables are attributable to the latent variable.

Testing equivalence of other model parameters. Up to this point, all the steps have only included invariance tests for the relations between first-order latent variables and the measured variables (the Λ, τ, and Θ components of the common factor model). It is also possible to use invariance tests to examine other parameters in the model, including the latent variable variances, latent variable means, as well as the covariances between the latent variables (Vandenberg & Lance, 2000; Widaman & Reise, 1997). These tests are not related to measurement invariance, rather they address substantively important questions. For instance, the comparison of latent covariances tests whether the relationships between latent variables are equal across groups.

In addition to testing the parameters in a factor model which includes a number of first-order factors (factors for which observed variables load), it is also possible to create a hierarchical model, where the covariances between the first-order factors may be explained by one or more higher-order factors. For example, in CHC theory, the covariances between the latent broad abilities are thought to be partially explained by a single higher-order general factor, *g*. This model is well established in the literature and has been used to describe the structure of a number of intelligence tests (Carroll, 1993, 2005; Gustafsson, 1984; Jensen, 1998; Keith, et al., 2006; Keith, et al., 2010; Keith, et al., 2008; Reynolds, et al., 2008; Taub & McGrew, 2004). It is

also possible to extend factorial invariance tests to parameters in the second order factor model (Chen, Sousa, & West, 2005). These steps have been followed in previous research on hierarchical CFA models (F. F. Chen et al., 2005; H. Chen et al., 2010; Keith et al., 2010). If there are not substantial changes in model fit as these constraints are added, then this provides evidence that the structural relations (covariances between latent variables) are similar between groups.

Previous Measurement Invariance Research with Tests of Cognitive Abilities

The use of measurement invariance with tests of cognitive abilities is a fairly new area of research, mostly because the examination of measurement invariance has become more common only in recent years (Millsap & Meredith, 2007). Most examinations of measurement invariance in cognitive ability tests have been completed with well-defined groups, including gender (H. Chen & Zhu, 2008), cultural groups (H. Chen et al. 2010; Edwards & Oakland, 2008; Locke, McGrew, & Ford, 2010), or age groups (Reynolds et al., 2010; Taub & McGrew, 2006). The most prolific researchers of measurement invariance with tests of cognitive abilities in groups that are less well-defined are Bowden and colleagues. Most of their investigations have focused on the Wechsler scales, specifically the Wechsler Adult Intelligence Scales—Third Edition (WAIS-III, Wechsler, 2001). For example, they have examined measurement invariance between the U. S. standardization sample and Australian neurological clinical sample (Bowden, Cook, Bardenhagen, Shores, & Carstairs, 2004), a community sample and sample from a neurological clinic from Australia (Bowden et al., 2004), a community sample and a sample of individuals with alcohol dependency (Bowden et al., 2001), and groups of college students with ADHD and learning disabilities (Bowden et al., 2008). In general, they have found that measurement invariance held fairly well across these samples.

Of particular interest here was the investigation of measurement invariance with college students diagnosed with a learning disability (Bowden et al., 2008) on the WAIS-III. The authors found that the basic structure of the WAIS-III with an additional long-term memory factor (from the Wechsler Memory Scales, 3rd edition, Wechsler, 2001) was essentially invariant across the SLD sample and college-aged students from the normative sample. This was an important finding because it provided evidence that the same latent variables were measured across these two populations with the WAIS-III, indicating that the presence of a learning disability does not affect how latent variables assessed by the WAIS-III are measured.

Purpose of the Current Study

The purpose of this study is two-fold. First, an examination of the invariance of the covariance structure for the WJ-III between children with and without SLD will be examined. This analysis will be completed to verify that the broad and narrow constructs are the same across groups, which is a necessary step to take before examining relations between cognitive abilities and academic achievement.

The second part of this study is designed to examine whether or not the relations between cognitive abilities and academic skills are the same for children with and without SLD. In other words, do children with SLD rely on the same cognitive abilities for different achievement areas as children without SLD, or do they rely on different abilities? Additionally, if children with SLD do rely on the same cognitive skills for different achievement areas, are the magnitudes of these relations the same among groups?

Importance of the Current Study

There are several reasons why the current research is important, both from a practical and theoretical level. As it has been emphasized in the previous literature review, the identification of

SLD has been controversial and has changed over time. If SLD is to be understood, at least in part, as a deficit in a psychological process that affects a person's ability to learn, then more research in understanding the relations between cognitive abilities and academic skills is essential. This research extends this notion by examining these relations among children and adolescents with and without SLDs. Additionally, if there are differences in which cognitive abilities are related to different academic skills, this could also be informative for intervention. For instance, if it is found that cognitive skills which are malleable (e.g., working memory, see Klingberg, 2010) are related to better performance on academic tasks for children with SLDs, it is possible that interventions which target cognitive skills may be beneficial for improving academic skills as well. In fact, some research (e.g., Loosli, Buschkuehl, Perrig, & Jaeggi, 2012) has found preliminary evidence that working memory training enhances reading skills in children. School-based interventions which may focus on improving more basic cognitive skills may have residual effects and improve academic skills as well. Knowing which cognitive skills are related to difference academic skills will be important if this type of intervention is a possibility.

From a theoretical standpoint, the current research is of interest because it will provide some insight regarding the similarities and differences between cognitive—achievement relations which have been researched extensively (e.g., McGrew & Wendling, 2010). Such information would be of interest to researchers who are using neuroscience techniques, such as fMRI, to investigate SLD because of a need to explicitly link that measurement model with one obtained from test scores. If children with SLDs show differential relations between cognitive abilities and academic skills, this may provide some important insight into how these children cope with difficulties in academic areas. For instance, if children with SLD in reading tend to decode better

if they have better visualization skills, but visualization skills are not related to decoding for normally developing children, this can provide some insight into how children with SLDs use their cognitive skills differently to improve academic performance. Additionally, this study is one of the first to examine the relations between cognitive abilities and academic skills using a CHC-based model which includes both broad and narrow cognitive abilities. Narrow abilities have been emphasized as being important indicators of SLD (e.g., Compton et al., in press), although narrow abilities have not been examined with children who have SLDs as much in the cognitive—achievement relations literature.

Hypotheses

This study is primarily exploratory, but there are several outcomes which may be expected based on theory and current research.

1. **Hypothesis 1: The factor loadings between the Norm group and SLD groups will be invariant.** For the first part of the study, it is expected that the WJ-III will at least meet the requirements for weak factorial invariance, where the factor loadings from the measured variables to the latent variables are equal across groups. Bowden et al. (2008) showed that a group of college students with SLD were invariant across factor loadings on the WAIS-III, although this finding may not translate to the current study because the SLD groups for this study are likely more homogenous than the sample of college students used in Bowden et al. (2008). Such a finding will provide evidence that the latent constructs are the same across the groups.
2. **Hypothesis 2: Narrow cognitive abilities will be most strongly related to academic skills.** McGrew and Wendling (2010) indicate that the relations between narrow cognitive abilities and specific academic skills tend to occur at the narrow ability level. Therefore,

it is hypothesized that narrow abilities will be more strongly related to academic skills than broad abilities or *g*. Previous research has indicated that *g* only plays an indirect role in its influence on academic skills for normally developing children (Elliot et al., 2010; Floyd et al., 2008, Taub et al., 2007). Elliot et al. (2010) found that this was also the case for children with difficulties in reading. SLD is hypothesized to be a deficit in a cognitive process, and not a global deficit. Therefore, *g* should not impact academic skills directly in children with SLD, *g* will only impact academic skills indirectly through more specific processes.

3. **Hypothesis 3: There will be differences in cognitive—achievement relations among groups. The greatest differences in cognitive—achievement relations will be related to SLD groups for whom the academic skill is related to their disability (e.g., SLD Reading will have significant relations between different cognitive abilities and Basic Reading Skills).** Children with and without SLD may be equal on all relations among cognitive abilities and academic achievement. This is unlikely, especially based on previous research which has compared the relations between academic achievement and cognitive abilities in a CHC model (e.g., Elliot et al., 2010). Elliot et al. identified Ga, Gv, Glr, and Gs as predictors of reading skills in children with reading difficulties, while for children without reading difficulties Ga, Gs, Gf, and Gsm were significant predictors. It is possible that similar differences may emerge in the current analysis, where a very different set of cognitive skills predicts academic skills. One important difference between the current study and previous research is that unstandardized estimates will be compared, whereas standardized estimates have typically been

compared in previous research (though see Keith, 1999). Standardized estimates are those most likely influenced by selection.

Chapter III: Methods

Participants

The participants for this study were obtained from two different samples. The first sample consisted of children and adolescents from the standardization sample of the Woodcock-Johnson Tests of Cognitive Abilities and Tests of Achievement (WJ-III, Woodcock, et al., 2001, 2007), and the second sample included children and adolescents diagnosed with a SLD obtained from the clinical database compiled by the Woodcock-Muñoz Foundation.

The full WJ-III standardization sample consists of 8,782 participants between the ages of 1 (19 months) and 90 years. A subsample of 300 school-aged children between the ages of 6 and 19 years were randomly selected from the normative sample for use in this study. The WJ-III clinical sample includes 1,374 children between the ages of 3 and 19 years, but only children between the ages of 6 and 19 were included in the analysis to ensure that they were likely to have had some formal academic instruction. This sample consists of scores from evaluations completed by neuropsychologists who used the WJ-III during the assessment process. These scores were provided to the Woodcock-Muñoz Foundation (WMF) as part of a clinical database project. Scores for the clinical database were obtained from either archival clinical records from licensed clinical or neuropsychologists, or from clinical research studies (McGrew, personal communication, December 28, 2011). All scores provided to the database were reviewed by two qualified neuropsychologists before they were included, and this was completed to either validate the diagnosis provided by the clinic or to provide a more specific or different diagnosis based on the information provided. All diagnoses from the clinics were made using the DSM-IV or International Classification of Diseases criteria, and the diagnoses were then recoded to be included in the clinical database under a uniform system of classification which combined both

of these diagnostic systems. The classification for the WJ-III clinical database included categories DSM-IV categories for Reading Disorder, Mathematics Disorder, and Disorder of Written Expression.

To reduce the effects of comorbidity of SLD in more than one academic areas (e.g., Reading Disorder and Mathematics Disorder), all individuals with either a secondary or tertiary diagnosis in one of the SLD areas were removed from the analysis. Individuals with only one diagnosis in one of the SLD areas were included in the groups. This provided sample sizes of n = 180 for SLD Reading, n = 231 for SLD Math, and n = 149 for SLD Writing. Demographic information for the normative subsample, SLD Reading, SLD Math, and SLD Writing are presented in Table 1. Proportions for gender, race, ethnicity (Hispanic origin/Not Hispanic origin), mother education, and father education are presented, along with the mean and standard deviation for age. Comparisons between the SLD group and normative subsample on demographic variables were completed using the χ^2 test, and age was compared using one-way ANOVA. Because there were only a few individuals in the Asian, Native American, and Other categories for race, these groups were combined into a single group for this analysis. There were no statistically significant differences between groups for age, F (3, 846) = .540, p = .655, mother education level, χ^2 (9) = 12.29, p = .197, and father education level, χ^2 (9) = 13.84, p = .128. There were statistically significant differences between groups for gender, χ^2 (2) = 40.62, p < .001, race, χ^2 (6) = 19.73, p = .003, and ethnicity, χ^2 (3) = 9.97, p = .019. For gender, the SLD Writing group appeared to have a substantially larger proportion of males (74.50%) than was present in other groups. This difference is consistent with previous research indicating that the ratio of males to females for written expression disorders is approximately 1.5:1 (Fletcher et al., 2007). The clinical groups had slightly higher proportions of Caucasian individuals and slightly

lower proportions of individuals from a Hispanic origin, suggesting that the racial and ethnic backgrounds of the individuals in the SLD groups were slightly different than the Norm group. Previous research has shown that the constructs measured by the WJ-III are invariant across gender (Keith et al., 2008) and the structural relations between cognitive and achievement variables were invariant across ethnicity (on the WJ-R, Keith, 1999), thus this should not be an issue in the current analyses.

Even after removing comorbid academic disorders, there were a substantial number of individuals with secondary and tertiary diagnoses. These comorbid disorders included Attention Deficit-Hyperactivity Disorder (ADHD), Expressive Language Disorder, Expressive/Receptive Language Disorder, Anxiety Spectrum Disorder, and Depressive Disorder. The proportions of individuals with secondary and tertiary diagnoses are included in Table 2. The SLD Reading group had the highest proportion of individuals with a secondary (26.11%) and tertiary (6.11%) diagnoses. The SLD Math and SLD Writing groups had similar frequencies of comborbid disorders, where approximately 10% had secondary diagnoses, and 4% had for tertiary diagnoses.

Measurement Instrument

The WJ-III consists of two different test batteries, the WJ-III Tests of Cognitive Abilities (WJ-III COG, 30 tests) and the WJ-III Tests of Achievement (WJ-III ACH, 22 tests). These tests are designed to measure seven broad CHC cognitive abilities (Gf, Gc, Ga, Gv, Gs, Gsm, Glr) and four areas of academic achievement (Reading, Writing, Mathematics, and Oral Language). In addition to the broad CHC abilities, the WJ-III is also able to measure a number of narrow abilities, which are described below. The WJ-III is one of the most flexible cognitive ability tests because of the wide variety of tests and abilities it measures. It has received positive reviews

from independent sources, and it a widely-used measure of cognitive and achievement (Ciszek, 2003; Sandoval, 2003).

The technical manual for the WJ-III (McGrew & Woodcock, 2001) indicates that the subtests and composite scores are reliable and valid measures of cognitive abilities across the lifespan. Reliability estimates for all test scores are above .74, with most between .80 and .90. A number of exploratory and confirmatory factor analyses using tests from the WJ-III and other well-established cognitive batteries (e.g., WISC-III, Differential Ability Scales) provide evidence that the tests from the WJ-III measure similar abilities as other often used intelligence tests. In addition to the reliability and validity evidence for scores from the WJ-III, several aspects of the WJ-III make it especially desirable as a measurement instrument for research. It is the first test of cognitive abilities to be designed specifically based on CHC theory (McGrew, 2005; McGrew & Woodcock, 2001). Initially, Gf-Gc theory (Horn & Cattell, 1966) guided the development of the Woodcock Johnson Psychoeducational Battery—Revised (Woodcock & Johnson, 1989) and subsequently CHC theory guided development of the WJ-III (Woodcock et al., 2001). The normative sample for the WJ-III includes over 8,000 individuals, which is very large relative to most test batteries. The normative scores for the WJ-III were weighted based on 11 demographic variables from the 2000 United States census, including census region, community size, sex, race, Hispanic (Yes or No), type of school, type of college/university, education of adults, occupational status of adults, occupation of adults in the labor force, and foreign born. In 2007 the standard scores were reweighted and recalculated in light of population changes according to the 2005 census (WJ-III Normative Update, WJ-III NU, McGrew, Schrank, & Woodcock, 2007), and all scores in the current study are based on the WJ-III NU. Finally, the achievement and cognitive batteries were co-normed, making it an excellent instrument for diagnostic decision

making because discrepancies and patterns of strengths and weaknesses can be based on discrepancy norms created from the normative sample (McGrew et al., 2007).

Group Selection

Several decisions were needed to carry out analyses with the available samples. The clinical sample for this study, as a whole, is large and offers several options for analysis. The only major study which has examined factorial invariance on a cognitive abilities test between individuals with and without SLD is Bowden and colleagues (2008). In their study, the individuals with SLD were analyzed as a single group, and the authors did not specify whether the individuals in the SLD groups were from a single category or if they exhibited difficulties in different categories (e.g., SLD in reading, mathematics, etc.). One important and unique aspect of this research was that data from children and adolescents with SLD diagnoses in reading, mathematics, and writing were analyzed as separate groups. Research has shown that different academic skills are influenced by different sets of cognitive abilities (e.g., McGrew & Wendling, 2010) and it has also been shown that SLD in different academic domains are associated with different cognitive strengths and weaknesses (e.g., Compton et al., in press).

A second issue which must also be addressed here is related to selection. Groups in studies are often from well-defined subpopulations, such as gender, cultural status, or age groups. In these cases, selection should not be an issue because the groups were not created directly on the variables of interest, that is, groups were not created based directly on specific test scores used in the analysis (e.g., all individuals with scores below 80 on Letter-Word Identification are in the SLD Reading group). It has long been known that if individuals are selected to be in groups based directly on the variables which are to be analyzed, this affects the correlational structure of the variables in the analysis (Pearson, 1903; Cohen, 1983). Many times problems

arise because researchers analyze correlations matrices, which are often distorted due to restriction of range. Any selection effects which may be present do not have the same influence on unstandardized parameters as they do on standardized parameters (Keith, 2006, p. 36). This is because the change in standard deviation for truncated variables has a strong effect on standardized parameters, but not on unstandardized parameters. Therefore, covariance matrices will be used, and factorial invariance will be evaluated, not assumed, in this study.

Missing Data

Many methods are available to handle missing data, including listwise or pairwise deletion, mean or regression substitution, expectation maximization substitution, and multiple imputation (Baraldi & Enders, 2010; Enders, 2010; Graham, Hofer, Donaldson, MacKinnon, & Schafer, 1997). Many disadvantages are found with the deletion and substitution methods, including a loss of power and biased parameter estimates, including the strict assumption of Missing Completely at Random (MCAR) that is rarely satisfied in practice (Baraldi & Enders, 2010, Graham et al., 1997). Currently, the recommended practices for missing data are the use of multiple imputation or maximum likelihood estimation (Enders, 2010). One of the major advantages to CFA and SEM methods is that most programs utilize the maximum likelihood algorithm to estimate model parameters. Maximum likelihood does not require cases with missing data to be deleted and it uses all the information provided in the data to estimate parameters which will best reproduce the sample data (Baraldi & Enders, 2010).

One issue with missing data is the manner in which the data is missing, that is, whether the missing values are directly correlated to other scores used in the analysis. There are three different types of missing data (Enders, 2010): Missing Completely at Random (MCAR), Missing at Random (MAR), and Missing Not at Random (MNAR). MCAR indicates that

missing data on one variable (y) is not related to any other variables used in the study. Missing values are randomly distributed among the dataset. MAR indicates that missing data for one variable (y) may be related to values of another variable in the analysis (x), but the probability that a value is missing is not related directly to the value of (y) itself (e.g., those with lower values of y have more missing data than those with higher values on y). For example, missing data on a reading test may be related to scores on a math test taken by the same group of individuals, but the probability that data are missing is not related to performance on the reading test itself (Enders, 2010). Finally, MNAR data occurs when the missing values of variable y are related directly to the values of y itself (e.g., those with lower values on y have more missing data than those with higher values on y). For example, for data to be MNAR some individuals may have missing data on a reading test because they could not finish the test based on poor reading skills. In that case, the probability that a value is missing is directly related to the variable itself.

Unfortunately, it is not possible to formally test the MAR and NMAR hypotheses because the data which would be necessary to test these hypotheses are missing. Despite this, some methods are available to examine missing data. One is to create groups based on missing and nonmissing values for one variable (y), and test the mean differences for all other variables between these groups. However, when a large number of variables are included in the analysis, this method quickly becomes difficult to examine. In addition, this approach does not take the correlation between variables into account, and although many significant differences are found, only one mechanism may account for missing data (Enders, 2010). Little's test for MCAR (Enders, 2010; Little, 1988), which examines this method as a single test, will be used in this study to examine the MCAR hypothesis. If this test is not statistically significant, it indicates that

missing data is MCAR, which is unreasonable in many situations. Maximum likelihood requires the less strict assumption that data are MAR, so a significant value for Little's MCAR test does not rule out the possibility that the data are MAR. Other methods for dealing with incomplete data, such as deletion methods, require the strict assumption of MCAR.

Model Evaluation

Several indices are available for evaluating CFA and SEM models. The most commonly used statistic is the χ^2 test, although this is well known to be very sensitive with large sample sizes (Brown, 2006; Kline, 2011). However, this test is commonly used and reported, and is useful for comparing models, and it will be reported here. Several other goodness-of-fit indices will be used in addition to the χ^2 to examine model fit. These include the comparative fit index (CFI, Bentler, 1990) and the root mean square error of approximation (RMSEA, Steiger & Lind, 1980). The RMSEA used in the multi-group model will be corrected for multiple groups, where the RMSEA value is multiplied by the square root of *K* groups (Steiger, 1998). The Akaike Information Index (AIC, Akaike, 1987), Bayes Information Index (BIC, Schwartz, 1978), and the sample size adjusted Bayes Information Index (aBIC). The likelihood ratio test will be used to compare models.

Schermelleh-Engel, Moosbrugger, and Müller, (2006) provided some guidelines for fit indices. For the CFI, values above .95 indicate acceptable model fit, while values above .97 are indicative of excellent fit. For the RMSEA, values below .05 indicate close fit, while values between .05 and .08 are acceptable. The AIC, BIC, and aBIC are unstandardized measurements and does not have a fixed scale, but they are helpful for comparing non-nested models which may not be directly comparable. For the AIC, BIC, and aBIC, smaller relative values indicate better model fit.

Although there have been rules of thumb for fit indices, several studies suggest that these should not be interpreted without considering other factors, such as the number of variables in the model and model complexity (Kenny & McCoach, 2003) and the size of factor loadings (Heene, Hilbert, Drazler, Ziegler, & Bühner, 2011). Specifically, Kenny and McCoach (2003) show that CFI values tend to show degradation in fit when a greater number of variables are included, whereas the RMSEA tends to show an improvement in fit with more variables. Likewise, Heene and colleagues (2011) show that models with moderate to low factor loadings (greater unique variance) tend to show a decrease in CFI values, although show an improvement in RMSEA values. Previous CFA research with the WJ-III which have used a large number of variables (Floyd et al., 2007; Taub et al., 2008) have had adequate values for the RMSEA and SRMR, although they have had lower values for the CFI (e.g., close to .90). This may be due to the large number of variables used in the analysis and presence of moderate factor loadings. Therefore, setting an absolute value of .95 for the CFI for acceptable model fit will not be used in this study, rather a preponderance of information obtained from all fit indices will be considered in the context of the complex model which will be used here. CFI values approaching .95 will be desired, but this will not be used as an absolute.

The $\Delta\chi^2$, ΔCFI, AIC, BIC, and aBIC will primarily be used to compare differences between models in the measurement invariance portion of the study. Because the χ^2 is sensitive to small differences and changes in model fit, a significance level of .001 will be set for all invariance tests based on the relatively large sample and complex models used in this study. This level is additionally chosen based on the large number of comparisons which will be occurring in this study. Previous studies have also used similar significance levels for the $\Delta\chi^2$ (Keith, et al., 2010; Keith, et al., 2008). According to Cheung and Rensvold (2002), the CFI was one of the

best performing indexes in a simulation study, and they recommended that a change in CFI of more than .01 indicates a significant change in model fit. French and Finch (2006) indicated that this value had a slightly inflated Type I error rate in their simulation study. Meade, Johnson, and Braddy (2008) suggested that a ΔCFI which is less than or equal to .002 suggests very minor changes in model fit, but this value is very strict. Other studies have used the ΔCFI value of .01 as an indication of a substantial change in model fit (Bowden et al., 2008; F. F. Chen et al. 2005), so a value equal to or approaching .01 for the ΔCFI, used in conjunction with the $\Delta\chi^2$ will be adopted for the current study. When comparing the equality of structural paths in the cognitive—achievement in the second part of this study, the likelihood ratio test will be used for all comparisons between models. Statistically significant parameters at the .05 level will be used as a guide for determining which relations should be included in the final model. Again, decisions regarding the inclusion of model constraints will also be guided by the AIC, BIC, and aBIC, where lower values suggest better model fit.

WJ-III CHC model

The WJ-III includes a substantial number of tests designed to measure a wide range of cognitive abilities, and these subtests can be combined to provide composite scores for both narrow and broad cognitive abilities. This allows for a comprehensive CHC model to be created, which will expand the scope of the analysis. Most of the previous research on measurement invariance with the WJ-III has focused only on examining the core subtests of the WJ-III (Edwards & Oakland, 2006; Taub & McGrew, 2004). Several studies have included additional subtests on the WJ-III model (Floyd, et al., 2007; Keith, et al., 2008; Phelps et al., 2005). The subtests chosen to be included in this study will be based on these previous studies and the factor

analyses included in the technical manual for the WJ-III (McGrew & Woodcock, 2001), as well sample specific restrictions on which tests may be used (e.g., sample size, covariance coverage).

All tests from the WJ-III and the broad and narrow abilities they measure are included in Table 3. When possible, narrow abilities will be included as separate latent variables in the analysis. The analysis of measurement invariance with narrow abilities essentially has no coverage in the literature, although some researchers have used models which include narrow abilities when analyzing relations between cognitive and academic abilities (Floyd, et al., 2007; Keith et al., 2008). The proposed CHC model for this study includes 27 subtests from the WJ-III cognitive and achievement. The seven broad abilities mentioned previously will be included in the model (although Gv is measured solely by visualization subtests, making it a measure of the narrow ability factor VZ), and this will also include 11 narrow abilities as well. When there are two subtests which measure a narrow cognitive ability, then a narrow ability latent variable will be included in the model. For example, with Gs there are two subtests which measure perceptual speed, Visual Matching and Cross Out. These two subtests will be included on a latent variable representing Perceptual Speed, and then Perceptual Speed will load on Gs. Two other tests, Decision Speed and Pair Cancellation measure different narrow abilities, and will only load on Gs. The purpose of including narrow abilities is because there has not been a strong focus on narrow abilities in the investigation of the relations between cognitive abilities and achievement with groups of children with SLD. Additionally, more recent research has suggested that specific cognitive deficits may underlie SLD (Compton et al., in press). By including tests which allow for the measurement of narrow abilities, this will help illuminate some of these possible differences. For the second portion of the study, five areas of academic achievement will also be included in the model using tests from the WJ-III ACH. The five achievement areas are Basic

Reading Skills (BRS), Reading Comprehension (RC), Math Calculation Skills (MCS), Applied Math (MR), and Written Expression (WE).

Question 1: Does the WJ-III CHC model have a similar covariance structure among groups of children with and without SLD?

To examine the covariance structure of the WJ-III factor model, invariance tests will be used to test whether certain parameters in the model are equal across groups. The main focus of this investigation is to determine if the covariance structure is similar among groups, and the mean structures will not be tested in this analysis. Invariance of the factor loadings is necessary for quantitative comparison regarding the factor variance and covariances, and thus these tests will be of particular interest for examination of the relations between cognitive and academic skills in the next part of the study.

The order in which measurement invariance is tested will follow typical research on measurement invariance with cognitive ability tests (e.g., Bowden et al., 2008; Keith et al., 2010), where factor loadings will be tested before residual variances, similar to procedures used by Keith et al. (2010); however, tests of mean structures and measurement intercepts were not included because the purpose of this research was not to evaluate the measurement properties of the WJ-III, but to ensure the constructs were similar across groups. A third-order model which includes *g* will be used because this is the model that will be used in the second part of this study. The following steps will be used for this analysis:

Tests of factorial invariance:

1. *Configural Invariance*. The model for each group will be estimated simultaneously using the same factor model. No parameters estimated within groups will be constrained to be

equal across groups, other than those used for identification purposes, will be included in the model for this analysis.

2. *Invariance of factor loadings (weak factorial invariance).* The invariance of factor loadings will be tested in three phases: 1) The equality of factor loadings from the narrow abilities to the subtests will be tested to determine if the subtests which measure the narrow ability factors are equal across groups, 2) The factor loadings of the subtests on the broad abilities (subtests which do not load onto a narrow ability factor) as well as the factor loadings narrow abilities on the broad abilities will be tested for equality, and 3) The factor loadings of broad abilities on *g* will be tested for equality. If a statistically significant change in model fit occurs at any point in these three analyses, individual parameters will be freed to determine which may be leading to model misfit.
3. *Invariance of subtest residuals.* In this model, all subtest residuals will be constrained to be equal across groups. This test of invariance is very stringent and is not necessary to compare relations among latent variables.

Tests of equivalence of substantive parameters:

4. *Invariance of narrow and broad ability residuals.* This model will also consist of two phases. First, the residual variances of the narrow abilities will be constrained to be equal across groups, and second the residual variances for the broad abilities will be constrained to be equal across groups.
5. *g variance.* In the final step of the examination of the covariance structure, the variance for *g* will be constrained to be equal across groups.

Question 2: Does group membership moderate the relations between specific CHC abilities and academic skills among children with and without SLDs?

The second part of this study is designed to determine if relations between cognitive abilities and academic skills are the same between children with and without SLD. Here, the second-order model from Question 1 will be used, and a set of tests will be conducted to examine the relations among broad and narrow CHC abilities and different academic skills. A multi-group model will also be used for this analysis, except structural relations between cognitive abilities and academic skills will also be modeled. Five academic skills will be examined: Basic Reading Skills (measured by Letter-Word Identification and Word Attack), Reading Comprehension (measured by Passage Comprehension and Reading Vocabulary), Math Calculation Skills (measured by Calculation and Math Fluency), Applied Math (measured by Applied Problems and Quantitative Concepts), and Written Expression (measured by Writing Samples and Writing Fluency). A separate analysis will be completed for each academic skill. Each of these skills will be examined through a series of steps designed to elucidate which cognitive abilities are related to the academic skills.

1. *Backward Selection*. First, a backward selection method will be used to initially determine which cognitive abilities are related to each academic skill for individual groups. Academic skills will be regressed on all of the cognitive abilities included in the model simultaneously, and structural paths will be sequentially removed until only statistically significant paths remain. Path removal will begin with the path with the highest negative non-statistically significant path. The new model will then be estimated and the next path with the highest negative estimate removed. This will be repeated until all remaining paths are positive. Next, each path which is not statistically significant at the $p < .05$ level will be removed using the same type of sequence, where paths that are not statistically significant and have the smallest standardized effects will be removed

until all remaining paths are statistically significant. The backward selection method will be used for each group separately to identify which cognitive—achievement relations should be included in the multi-group model. The backwards selection method has been used in previous research on the relations between CHC abilities and academic skills to determine which cognitive abilities relate to academic skills (e.g., Elliot et al., 2010; Floyd et al., 2007). This method was chosen because it is better able to control for specification error compared to forward selection methods (Keith, 2006). The purpose of this step will be to determine which structural paths are statistically significantly different from zero within each group.

2. *Multi-group model*. Once the backward selection has been completed for each group, all paths from cognitive abilities which were statistically significant will be included in a multi-group model and freely estimated across all groups (e.g., all paths which were statistically significant for the Norm group will be estimated for the Norm, SLD Reading, SLD Math, and SLD Writing groups simultaneously). A series of steps will be performed to determine if the magnitude of the structural paths is statistically significantly different across groups. These are interpreted as tests of moderation, where the magnitude of the structural paths depends on group membership. Because researchers approach such tests in different ways, by either adding equality constraints to a set of paths or by beginning with all corresponding paths constrained to be equal across groups and releasing constraints, both ways will be employed in this study for purposes of comparison.
    a. *Individual path constraints*. All structural paths identified as statistically significant for at least one group in the backward selection will be freely estimated across all groups in the multi-group model. Equality constraints will be

added to each corresponding structural path individually across groups to determine if the magnitude of the relation is equal across groups. A statistically significant degradation in model fit suggests that group membership moderates the relation between the cognitive ability and the academic skill. Additionally, if the path was statistically significant for a specific group in the backwards selection, then the path will be freely estimated for that group while it is constrained to be equal for the other groups. For example, if the path from Gc to Reading Comprehension was statistically significant for the SLD Reading group in the backward selection, this path will be included in the multi-group model and estimated freely for all groups. First, the path will be constrained to be equal across all groups. Then, because this path was statistically significant for the SLD Reading group in the backward selection, a second model will be estimated where the path from Gc to Reading Comprehension will be freely estimated for the SLD Reading group, but constrained to be equal for all other groups. If allowing this path to be freely estimated for the SLD Reading group results in a statistically significant improvement in model fit, it suggests that the magnitude of this relation is statistically significantly different from the other groups. This test is determining if the magnitude of the relations between cognitive abilities and academic skills are moderated by group membership. Such moderation is implicitly assumed by modeling the groups separately, but this will provide a formal test.

b. *Sensitivity analysis*. A sensitivity analysis will be completed to verify the findings from the previous procedure. The purpose of the sensitivity analysis is to start

from a different assumption. In this analysis, all structural paths which were statistically significant for any group during the backward selection will be included in the model. Rather than allowing all structural paths to be freely estimated within groups as the initial model, all structural paths will constrained to be equal across groups. Then, the equality constraint on each corresponding structural path will be released one at a time. A statistically significant improvement in model fit indicates differences in the magnitude of the relation among groups. If the path is statistically significant for an individual group in the backward selection, then the equality constraint for that path will be released for the individual group. For example, if the path from Gc to Reading Comprehension was statistically significant in the backward selection for the SLD Reading group, then it will first be freely estimated for all groups, and last it will also be freely estimated for the SLD Reading groups only. If there is a statistically significant improvement in model fit when the equality constraint for the single group is released, then it suggests that group membership moderates the relation between the cognitive ability and the academic skill. The primary purpose of this analysis is to verify the results from the previous analysis using a similar procedure, but instead starting with all paths constrained.

c. *Final model specification*. A final model will be estimated using the findings from the previous two analyses. All paths which did not result in a statistically significant change in model fit when constrained to be equal across groups will be included in the model. Additionally, any individual path constraints which resulted in statistically significant changes in model fit will be included in the

final model as well. For example, if model fit improved by allowing the path from Gc to Reading Comprehension to be freely estimated for the SLD Reading group only, then this path would be free for the SLD Reading group and constrained to be equal across all other groups.

Chapter IV: Results

Missing Data

Missing data analysis was completed to determine the amount of missing data in the samples used for the current analysis. Overall, 22.5% of the data for the 37 variables to be used were missing, and 36.1% of cases had complete data on all variables. Little's test for Missing Completely at Random (MCAR) was statistically significant for the entire sample χ^2 (3710) = 4555.06, $p < .001$, suggesting that the data cannot be assumed to be MCAR. Missing data analysis was also completed at the individual group level to determine the amount of missing data in the different samples and whether the data was MCAR. All variables which were used in the analysis, including both cognitive and academic variables, were included in the missing data analysis. In the Norm sample, 107 (35.7%) of cases had complete data on all variables, and overall 21.1% of the data were missing. Little's MCAR test was significant, χ^2 (2018) = 2284.034, $p < .001$, which indicates the data cannot be assumed to be MCAR. In the SLD Reading sample 55 (30.6%) of cases had complete data on all variables, and overall 30.3% of the data were missing. Little's MCAR test was significant, χ^2 (1214) = 1345.112, $p = .005$, indicating that the data cannot be assumed MCAR. In the SLD Math sample 53 cases had complete data (22.9%), while overall 23.8% of the data were missing. Again, Little's MCAR test was significant for this group, χ^2 (773) = 965.459, $p < .001$. Finally, for the SLD Writing group, Little's MCAR test was not significant, χ^2 (797) = 828.607, $p = .212$, 96 cases had complete data (64.4%) and overall 11.8% of the data were missing.

It is not possible to directly test whether data are MAR (Enders, 2010). However, because the individuals used in this study are gathered from a variety of clinics and from the normative sample for the WJ-III, there is no strong reason to assume that a particular mechanism within the

study is directly related to the missing data. Clinics may differ on which groups of subtests they administer to clients, or different subtests may have been administered to clients based on suspected disability. These reasons for missing data would not be directly related to scores on the tests used in the study. Such an influence should not influence the likelihood of the MAR assumption. Moreover, it should be noted that the subtests with the greatest amount of missing data were supplemental subtests from the WJ-III Diagnostic Supplement. The core cognitive subtests of the WJ-III tend to have larger percentages of data in each of the SLD groups. This can be likened to a reference variable approach, where all or most individuals in the study have complete data on a core set of variables which are most important to the constructs of interest (e.g., WJ-III core cognitive subtests), and other variables in the model are not measured for all individuals, but are measured for a portion of individuals. McArdle (1994) showed that the reference variable approach has similar results when compared to models with complete data. Moreover, although different subtests may have been administered to clients based on suspected disability, such decisions should not affect the MAR assumption because each academic skill is analyzed separately. Thus decisions were not directly related to individual scores on the subtests or to the probability that data are missing. For data to be MNAR, a mechanism such as certain tests not being administered because of a particular score on a subtest or composite used in the study. Regardless, more traditional techniques, including deletion methods, mean substitution, or regression substitution all assume that data are MCAR, which was not met with all groups. The use of maximum likelihood estimation is the most appropriate method for handling incomplete data in this study, as the data are assumed to be MAR.

Univariate Descriptive Statistics

Univariate skewness and kurtosis values of 2 and 7, respectively, may be problematic for maximum likelihood estimation (Curran, West, & Finch, 1997). All variables met the criteria for ML estimation for skewness (values were between -1.39 and 1.52 for all groups, on all variables) and kurtosis (values were between -.754 and 6.00). Means, standard deviations, percentage of data available, skewness, and kurtosis values for the Norm and SLD Reading groups are presented in Table 4, and these results for the SLD Math and SLD Writing groups are in Table 5.

Standard scores, which have a mean of 100 and a standard deviation of 15, were used in this study. Minimum and maximum values for each variable were examined to determine if any values appeared to be out of range. One value appeared to possibly be a data entry error in the SLD Writing group for Visual-Auditory Learning. This value was a standard score of 4.94, which was a *z*-score of -5.49 in SLD Writing group. Because this value was extreme and unlikely, it was removed from the analysis. Univariate outliers were identified for each variable through the use of *z*-scores. These were calculated for each variable and each group separately. Any value which was greater than +/- 3.29 ($p < .001$) was flagged as an outlier, although for studies which use larger sample sizes some outliers would be expected (Tabachnik & Fidell, 2007). The number of outliers for each variable ranged from zero to four in any single group on a single variable, and on average there was less than one outlier per variable in each group. Because there were only a small number of outliers, they were left in the initial analysis in order to preserve as much data as possible. All models were run both with outliers included and with outliers removed to determine if there were any substantive changes in the findings based on the removal of outliers. There were no differences in substantive findings when outliers were included or excluded from the models. Therefore, the results which included outliers in the analysis are reported here.

Question 1: Does the WJ-III CHC model have a similar covariance structure between groups of children with and without SLD?

Individual Group Models. The initial CHC model is presented in Figure 1. It includes 27 subtests, nine first-order narrow ability factors, six second-order broad ability factors, and a third-order *g* factor. Each of the narrow ability factors is indicated by two subtests. There are only six second-order broad ability factors because the tests used for measure Gv are both measures of VZ. Thus, this factor is best defined as a narrow ability. There were two cross-loadings to subtests which were included in the initial model based on previous research (McGrew & Woodcock, 2001), including a cross-loading for Memory for Sentences on LS and MS, and a cross-loading for Cross Out on both PS and VZ. Cross-loadings were considered acceptable in this study because it is unlikely that cognitive tests are pure measures of a single construct, and simple structure may be an ideal that is not always achieved in practice (Meredith & Horn, 2001). Because the purpose of this research was not to evaluate the structure of the test itself, but to investigate the relations among the latent constructs, such cross-loadings were considered acceptable because forcing simple structure when it is not appropriate may inflate correlations among factors. Additionally, the narrow ability RN for the WJ-III loaded both on Gc and Gs (Kaufman, Reynolds, Liu, Kaufman, & McGrew, in press). This is different from other research which has loaded RN on Glr. The loading of RN on Gc for the WJ-III is consistent with the test content, which includes Rapid Picture Naming and Retrieval Fluency. Both of these tests are influenced by background knowledge, and therefore the loading on Gc is warranted. A loading on Glr was included, but this loading was not statistically significant ($p > .05$ for all groups), while the loading from Gc to RN was statistically significant for all groups ($p < .01$).

Before testing for factorial invariance, a WJ-III factor model was fitted to each group's data individually to determine if the proposed model was appropriate for the analysis. There were several group specific model modifications which needed to be made in order to obtain adequate model fit for all groups. For the Norm group, the residual variance for Verbal Comprehension was negative, indicating a possible problem with the factor VD. A negative residual variance on a measured variable indicates that it is perfectly correlated with the factor, suggesting that the variable is perfectly reliable. Because this is highly unlikely, it is more likely that there is a problem with the factor VD. The problem was likely due to the shared content between Verbal Comprehension and Picture Vocabulary (Verbal Comprehension includes a Picture Vocabulary component). The VD factor was removed, and both factor loadings for Verbal Comprehension and Picture Vocabulary were loaded directly on Gc. In this model there was no longer a negative residual for Verbal Comprehension. This same issue was found when all groups were tested against the initial CHC model, so the VD factor was removed. Both Verbal Comprehension and Picture Vocabulary loaded directly on Gc in all subsequent models.

The residual variance for PC in the Norm group was fixed to zero because it was negative. This negative variance is referred to as a Heywood case, and has occurred in previous research with the WJ-III (e.g., Floyd et al., 2008; Taub & McGrew, 1996). This problem is not uncommon in models which include only two indicators on a factor (Loehlin, 1998). Heywood cases were noted, and if Heywood cases were present for other groups, then constraining the residual variance to zero for all groups was considered. In this model, fixing the residual variance of a narrow ability to zero effectively removed that factor from the model. Statistically significant residual covariances in this model were between Gc and Ga, Gc and Gf, Picture Vocabulary and Verbal Comprehension, Picture Vocabulary and General Information, Visual

Matching and Numbers Reversed, Number Series and Visual Matching, and Number Matrices and Pair Cancellation.

In the SLD Reading group, the residual variance for Memory for Words was negative, indicating a possible problem with the MS factor. This was possibly due to the cross loading for Memory for Sentences on LS and MS. The paths for this cross loading were set to be equal to each other, which did not result in a statistically significant change in model fit, and there was no longer a negative residual variance for Memory for Words. Three cross-loadings were necessary for this group in order to obtain adequate fit. These included Visual Matching on RQ (due to the use of numbers on Visual Matching), Auditory Attention on Gs (processing speed may be a factor in this test because the examinee must respond within a short amount of time before the next item is presented), and Story Recall on Gf (Story Recall has been related to Gf in previous research, Keith et al., 2008). The residual variances for the LS, PC, and Glr factors were set to zero because the standardized loadings were greater than one. Statistically significant residual covariances for the SLD Reading group were between Gc and Ga, Verbal Comprehension and Picture Vocabulary, Picture Vocabulary and General Information, Number Matrices and Pair Cancellation, Oral Comprehension and General Information, and Spatial Relations and Memory for Sentences.

The model with the cross loadings from the SLD Reading group was used to estimate the model for the SLD Math group. In the SLD Math group no additional cross-loadings were necessary for adequate model fit. The cross-loadings for Visual Matching on RQ and Story Recall on Gf were not statistically significant in the SLD math group. These loadings, however, were retained until they could be tested in the multiple group model. These loadings may not be statistically significant due to lack of power in the single group model, or they may be important

in the SLD Reading Group, but not the SLD Math group. The residual variances for PS and LS were set to zero because the residual variances were negative. Statistically significant residual covariances were between Gc and Gf, Verbal Comprehension and Picture Vocabulary, Picture Vocabulary and General Information, Visual Matching and Numbers Reversed, Number Matrices and Pair Cancellation, Academic Knowledge and Story Recall, Decision Speed and Retrieval Fluency, and Oral Comprehension and Memory for Sentences.

For the SLD Writing group no residual variances were negative. All paths and cross loadings which were identified in the SLD Reading group model were statistically significant for the SLD Writing group. Statistically significant residual covariances in the SLD Writing group were between Gc and Ga, Gc and Gf, Visual Matching and Pair Cancellation, Academic Knowledge and Story Recall, and Retrieval Fluency and Memory for Sentences.

A model was created which included all residual covariances which were statistically significant from the individual group models. Because PC and LS had either negative residual variances for more than one group or were not statistically significantly different from zero in the other groups, these residual variances were set to zero for all groups. By doing this, the factors of PC and LS are perfectly correlated with their corresponding broad factors, and the factors essentially collapsed, which resulted in six narrow ability factors. The final model to be used in the study is presented in Figure 2. Each group was individually tested against a model which included all the residual covariances identified above, whether or not they were statistically significant in that group. Model fit statistics for each individual group are presented in Table 6. The Norm group and SLD Math group had the best model fit overall, while the SLD Reading and SLD Writing groups had adequate fit. Next, the multi-group model was created to test for measurement invariance.

It should be noted that the purpose of this study was not to investigate every correlation between specific factors (residual variances) or cross-loadings of certain tests on broad or narrow factors. Specific factors are likely to arise when a large number of variables are included in the analysis, and it is not surprising that they may crop up when studying different groups (Meredith & Horn, 2001). Small adjustments made within each group should not detract from the study of the common factors, which have been found in many studies (e.g., Carroll, 1993). The findings do, however, suggest that these specific factors may require more investigation in future research in which the factor structure of the cognitive test is the focus. Most of the residual covariances included in this study were between subtests with shared content (e.g., Picture Vocabulary and Verbal Comprehension) or shared method/stimuli (e.g., subtests which rely on numbers, including subtests such as Number Matrices, Numbers Reversed, and Visual Matching), and were generally small in magnitude. It would be important that they appear in many other studies, and not simply due to differential sampling.

Results of Factorial Invariance Tests

Configural model. The Configural model, in which all parameters were freely estimated between groups, fit fairly well according to fit indices, χ^2 (1171) = 1659.29, $p < .001$, CFI = .939, adjusted RMSEA = .044, AIC = 6457.29, BIC = 17871.96, aBIC = 10253.38 (see Table 7, Model 1 [7.1[1]]). Because the Configural model fit relatively well, further tests of factorial invariance were conducted.

Invariance of narrow ability factor loadings. The corresponding factor loadings from the narrow abilities to the subtests (KO, LS, RN, PS, MS, RQ, MA, and PC) were constrained to be equal across all groups. Based on the criteria set previously, there was not a statistically significant

[1] Note that 7.1 indicates that this is the first model in Table 7. This notation will be maintained throughout the Results and Discussion sections.

change in model fit when compared to the Configural model, $\Delta\chi^2$ (33) = 54.17, p = .012 (Table 7.2). This finding suggests that the factor loadings from subtests to the narrow abilities are equal across groups.

Invariance of broad ability factor loadings. All corresponding factor loadings from broad abilities to both subtests and narrow ability factors were constrained to be equal across groups. Adding these constraints did not lead to a statistically significant degradation in model fit, $\Delta\chi^2$ (46) = 53.05, p = .221 (Table 7.3), which suggests that all factor loadings for the subtests are invariant across groups. With these equality constraints, all factor loadings were statistically significant. This finding suggests that the cross-loadings included in the previous models are important across groups, and non-statistically significant factor loadings for individual groups may have been due to a lack of power within each group.

Invariance of g factor loadings. Corresponding paths from *g* to the broad abilities were constrained to be equal across all groups. These additional constraints led to a nearly statistically significant change in model fit according to the likelihood ratio test based on the criteria set above, $\Delta\chi^2$ (18) = 42.01, p = .001 (Table 7.4). However, the change in CFI was -.003, which was smaller than the criteria of -.01. Additionally, the BIC and aBIC were lower for this model compared to the previous model, and the BIC and aBIC indicated that this was the best fitting model out of all the invariance tests previously conducted. Based on the preponderance of information for these fit indices the factor loadings were judged to be essentially equal. This finding suggests that the influence of the latent factors on the subtests is essentially equal among the groups. Different ways of identifying the model (e.g., using different subtests to scale the factors), will result in the same rescalings of parameters within each sample. Or simply, the latent constructs are the same across groups and quantitative comparisons of factor variances and

covariances across groups are valid. Therefore, the second portion of this study will be possible once it is determined that the achievement factors are also invariant.

Equality of subtest residuals. Equality constraints were placed on all corresponding subtest residuals to determine if they were equal across groups. Adding these constraints resulted in a statistically significant change in model fit, $\Delta\chi^2$ (80) = 299.45, $p < .001$ (Table 7.5). Each subtest residual was released to identify localized strain on the model. There were nine subtest residuals which resulted in a statistically significant improvement in model fit when they were freely estimated among groups. These included Story Recall (freely estimated for the Norm group only), Rapid Picture Naming (freely estimated for the Norm group only), Pair Cancellation (freely estimated for all groups), Spatial Relations (freely estimated for all groups), Number Matrices (freely estimated for all groups), Concept Formation (freely estimated for the Norm group only), Analysis Synthesis (freely estimated for all groups), Incomplete Words (freely estimated for the SLD Math group only), and Auditory Attention (freely estimated for all groups). Allowing these residual variances to be free no longer resulted in a statistically significant change in model fit, $\Delta\chi^2$ (62) = 89.30, $p = .011$. The invariance of subtest residuals is a stringent test and not required for comparisons of latent variable covariances. This finding suggests that there are some differences in subtest residual variances among groups. Because the differences or non-differences may indicate differences in specific factors or error variance, some researchers advocate that this test is not meaningful (e.g., Little, 1997). Moreover, invariance of the residual variances is not critically important for the overall purpose of the current research because differences in residual variances do not affect relations between latent factors.

Results of Differences in Latent Construct Variances

Invariance of narrow ability residuals. The next step was to determine if the corresponding residuals for the narrow ability factors were equal across groups. This test is of substantive interest. Adding equality constraints to all the narrow ability residuals did not lead to a statistically significant degradation in model fit, $\Delta\chi^2$ (18) = 31.32, p = .026 (Table 7.6), suggesting these are equal among groups. The groups relied on a similar range of these factors across groups.

Invariance of broad ability residuals. Equality constraints were added to the broad ability residual variances, and this did lead to a statistically significant degradation in fit, $\Delta\chi^2$ (21) = 69.70, p < .001 (Table 7.7). Each residual was released to identify localized strain on the model. The residuals for Gc and Ga appeared to be affecting model fit the most. Both the SLD Reading and SLD Writing had larger residual variances compared to the Norm and SLD Math groups for Gc and Ga, indicating more heterogeneity in those factors for those groups. These differences appeared to be the cause of localized strain on the model. When both Ga and Gc residual variances were freely estimated across groups, there was no longer a statistically significant change in model fit, $\Delta\chi^2$ (15) = 26.39, p = .034.

g variance. The final step of examining the covariance structure of the model was to constrain the *g* variance equal across groups. This constraint did not lead to a statistically significant change in model fit, $\Delta\chi^2$ (3) = 1.73, p = .629 (Table 7.8), suggesting that all groups have similar variability in latent *g*.

Summary

The model tested here met the requirement for weak factorial invariance, where all factor loadings are invariant. The latent constructs are similar across groups. Based on these results, it

is appropriate to move on to the second part of the study, where the relations between cognitive abilities and academic skills were examined.

Question 2: Does group membership moderate the relations between specific CHC abilities and academic skills among children with and without SLDs?

The second part of the study included individual academic skills in the model to determine which cognitive abilities were related to the academic skills across groups. For each academic skill, the systematic set of steps described above was conducted to identify cognitive abilities that had statistically significant effects on each academic skill. These models were focused on determining whether specific factors, beyond *g*, were important (i.e., statistically significantly different from zero) in explaining variance in the latent academic skills within each group. Modeling the effects separately within groups, however, implies the effects are moderated by group. Therefore, the second step of constraining paths across groups was completed to explicitly test for moderation. This set of steps was repeated for the five academic skills being tested in the current study.

First, it was necessary to determine if the factor loadings from the narrow latent academic skills to the corresponding achievement subtests were invariant across groups. A measurement model was created which included only the five academic skills and their corresponding subtests. Setting all corresponding factor loadings to be equal across groups did not lead to a statistically significant change in model fit according to previously established criteria, $\Delta\chi^2$ (15) = 35.81, p = .002, ΔCFI = -.006. Also, the BIC and aBIC were smaller for the model with factor loading constraints, suggesting the corresponding factor loadings are essentially equal across groups.

Results for Basic Reading Skills

The indicators for the Basic Reading Skills (BRS) factor were the Letter-Word Identification and Word Attack subtests. Initial backward selection analysis with individual groups indicated that MA, RQ, MS, and K0 were statistically significant for the Norm group, K0, Gsm, and PS were statistically significant for the SLD Reading group, Gsm and Gc were statistically significant for the SLD Math group, and RQ, MS, and K0 were statistically significant for the SLD Writing group. These findings indicated that groups differed on which specific factors had statistically significant effects on BRS when using the backward selection procedure.

Next, a multi-group model which regressed BRS on Gc, Gsm, MA, RQ, K0, PS, and MS simultaneously and freely within each group was estimated (see Table 8.1). Because K0 had a high factor loading on Gc (standardized loading = .97), the model used here was estimated with K0 alone. When K0 and Gc were both included in the model and constrained to be equal across groups, Gc was negative and not statistically significant, indicating that it may be acting as a suppressor variable (Kline, 2011). Gc was identified as statistically significant for the SLD Math group in the backward selection. When K0 was included instead of Gc, there was a negligible difference in model fit. K0 was retained for initial model estimation because it was more common in the backwards selection analysis. All models were reestimated using Gc instead of K0, and there were no changes in substantive findings. Model fit was generally better with K0, and based on this K0 was retained for the final model.

All model results are presented in Table 8. Each corresponding structural path from the cognitive abilities to BRS was constrained to be equal across groups, one at a time, to determine if there were differences in specific paths. If a path was statistically significant for a single group in the backwards selection, the path was released for that group to determine if it was statistically

significantly different from the other groups in the multi-group model. For example, PS was only statistically significant for the SLD Reading group in the backward selection. First, an equality constraint was placed on the path from PS to BRS for all groups. Next, the equality constraint for the path from PS to BRS was released for the SLD Reading group only to determine if the magnitude of the path was statistically significantly different from the other groups. If there was a statistically significant change in model fit, the magnitude of the path differed in the SLD Reading group compared to the other groups.

Adding equality constraints individually on the paths from K0, PS, Gsm, MA, and RQ did not result in a statistically significant degradation in model fit. Only the model where the path from MS to BRS was constrained to be equal across groups resulted in a statistically significant degradation in model fit (Table 8, Models 5-5c). MS was statistically significant for the Norm and SLD Writing groups in the backward selection analysis. When the two were freed in combination, the result was a statistically significant improvement in model fit (Table 8.5c). It appears that the Norm and Writing groups may rely on MS for basic reading skills more than the other two groups.

For the sensitivity analysis, all structural paths were constrained to be equal across groups and then the equality constraint on each structural path was released individually to determine if there was a statistically significant improvement in model fit. Similar to the previous analysis, there were no statistically significant changes in model fit when the equality constraints were released for K0, Gsm, RQ, or MA. However, there was a statistically significant improvement in model fit when the path from PS to BRS was freely estimated for the SLD Reading group only (Models 10 and 10a). Releasing the equality constraint on MS for all groups resulted in a statistically significant improvement in model fit (Models 12-12c), but allowing only the Norm

or SLD Writing groups to be released individually did not lead to statistically significant improvements in model fit, although when released for the SLD Writing group the change approached significance (p = .086).

The final model was estimated by constraining K0, RQ, Gsm, and MA to be equal across all groups (Model 15). Because releasing the constraint on PS was statistically significant for the SLD Reading group in the sensitivity analysis, and it was statistically significant when the groups were modeled separately, the path from PS to BRS was freely estimated for the SLD Reading group and was constrained to be equal for all other groups. Additionally, because releasing the path from MS to BRS for the Norm and SLD Writing groups was statistically significant in the previous analyses, MS was freely estimated for these groups and constrained to be equal for the SLD Reading and SLD Math groups. This model was estimated, but the path from MS to BRS for the Norm group was not statistically significant. The path from MS to BRS was constrained to be equal for the Norm, SLD Reading, and SLD Math groups, and was freely estimated for the SLD Writing group. There was not a statistically significant change in model fit, so this constraint was retained (Model 15a). The final model is presented in Figure 3, and consisted of only two group specific differences. PS was freely estimated for the SLD Reading group, and MS was freely estimated for the SLD Writing group. This finding suggests that the effects of these two abilities on BRS are moderated by group membership. Overall the findings related to the effects that were not moderated by group indicate that several specific abilities are important for understanding individual differences in Basic Reading Skills, and these influences are equally important across groups.

Table 9 includes total, direct, and indirect effects from cognitive abilities to BRS. Direct effects are the direct influence of one variable on another, that is, there is a direct path from one

latent variable to another. Indirect effects are the influence of one variable mediated through another variable. Total effects include both direct and indirect effects on the endogenous variable. Both unstandardized and standardized effects from cognitive abilities to BRS are included. Unstandardized effects are interpreted similar to unstandardized regression coefficients (Keith, 2006). For instance, the relationship between K0 and BRS suggests that a one point increase in K0 is associated with a .312 point increase in BRS, regardless of group (i.e., all factor loadings were constrained equal across groups). The interpretation of unstandardized effects depends on how the latent variables are scaled. In this study, all narrow abilities are scaled using observed variables, and the effects of the narrow abilities can be interpreted approximately on the scale of standard scores (M = 100, SD = 15). Not all broad abilities were scaled on observed variables, rather, some (specifically Gsm) were scaled on the narrow abilities. For standardized effects, several rules of thumb are used to interpret the magnitude of the effects. Keith (2006) suggests that in research related to education, standardized effects between .05 and .09 are considered small, effects between .10 and .24 are considered moderate, and effects greater than .25 are considered large.

Similar to previous research, the total effects of a *g* on basic reading were large, but were only indirect, and mediated by the broad and narrow cognitive abilities (Elliot et al., 2010; Floyd et al., 2007). Therefore, although *g* has a large total effect on basic reading, the mechanism by which *g* influences basic reading operates through more specific abilities. Statistically significant paths from cognitive abilities to BRS which were equal among groups were Gsm, K0, and RQ. This suggests that the differential effects of short-term memory (which includes both memory span and working memory), background knowledge, and quantitative reasoning had the same influence on BRS across all groups. These effects were all direct. Short-term memory appeared

to have the largest effect, followed by K0, with RQ having the smallest. The effects of Gsm and K0 would be considered large, while the effects of RQ would be considered moderate. The indirect effect of Gc, which was mediated by K0, was large (β's ranged between .320 and .348). There was also an indirect effect of Gf which was mediated by RQ, and this effect was moderate (β's ranged from .139 and .150).

Effects from other specific cognitive abilities were moderated by SLD group membership. The SLD Writing had a statistically significant path from MS in addition to Gsm, suggesting that MS also has a unique effect on BRS. That is, in the SLD Writing group the specific skill of memory span has an influence on BRS above and beyond the influence of broad short-term memory skills. It also indicates a larger total effect of Gsm, although the effect is indirect. Last, the most interesting finding was a statistically significant path from PS to BRS in the SLD Reading group, suggesting that perceptual speed is a significant predictor of basic reading skills for individuals with SLD in reading, but not for individuals without reading difficulties. Although the findings were slightly inconsistent, the magnitude of the effect (b = .237; β = .191) suggests that indeed the effect is important. According to Keith's rules of thumb for influences on school learning, this effect is considered to be moderate in magnitude (Keith, 2006). There was also a moderate indirect effect from Gs in the SLD Reading group, β = .155, and this indirect effect from Gs was very small for all other groups, β's ranged from .004-.006.

Results for Reading Comprehension

The relative effects of cognitive abilities on reading comprehension were investigated in each group. For the Norm group, paths from Gc and PS were statistically significant, paths from K0 and RQ were statistically significant for the both the SLD Reading and SLD Math groups, and for the SLD Writing group, paths from K0 and RN were statistically significant. Again, these

findings indicate some potentially moderated relations between specific cognitive skills and reading comprehension due to group membership. Similar to the BRS analysis, Gc was left out of the initial analysis because of the high factor loading from K0 to Gc. All models were reestimated including Gc instead of K0, but no substantive findings were changed. Model fit was slightly better with K0, so the path from K0 to RC was retained for the final model.

Results for all RC analyses are presented in Table 10. RC was regressed simultaneously on PS, RQ, K0, and RN in the multi-group model, with all paths freely estimated across groups (Table 10.1). Adding equality constraints individually to each structural path did not result in a statistically significant change in model fit for any path (Table 10.2-10.4). Structural paths that were statistically significant for individual groups from the backward selection were also tested. The only structural path which resulted in a statistically significant improvement in model fit when it was released for a single group was the path from RQ to RC for the SLD Reading group (Table 10.5a). This finding suggests that the magnitude of the path from RQ to RC is different for the SLD Reading group compared to the other groups. The same results were obtained in the sensitivity analysis, where the only statistically significant improvement in model fit was when the structural path from RQ to RC was freely estimated for the SLD Reading group (Table 10.10a). Those who have been identified as having a SLD in reading, are likely to rely more on reasoning skills for reading comprehension.

For the final model, all structural paths were constrained to be equal across groups except for the path from RQ to RC, which was freely estimated for the SLD Reading group and constrained to be equal for all other groups (Table 10.11). The final model is presented in Figure 4, and the unstandardized and standardized effects for the final RC model are presented in Table 11. Similar to the BRS model, *g* had large total effects on RC for all groups, but these effects

were indirect only. The effects of K0, PS, and RN on RC were equal across all groups, but the effects of PS and RN were not statistically significant for any groups in the multi-group model. The parameter estimates for these would be considered to be small and potentially important, but it cannot be said that they are reliably different from zero. Moreover, although the statistical significance of the total indirect effect from Gs factor was not estimated, the size of this total effect suggests that Gs had a small to moderate total effect on RC (See Table 11), K0 had the largest direct effect on RC (*b* = .569), and this effect was equal across all groups. Finally, there was also a large indirect effect from Gc to RC for all groups (*b* = .580).

The most interesting finding was that the effect of RQ on RC was moderated by group membership. This path was equal, and statistically significant, for the Norm, SLD Math, and SLD Writing groups, but was freely estimated for the SLD Reading group. The effect of RQ on RC was statistically significant and larger in magnitude in the SLD Reading group compared to the other groups. This standardized effect would be considered large for the SLD Reading group and moderate for all other groups. The effect from RQ to RC also resulted in a larger indirect effect from Gf on RC for the SLD Reading groups. This finding indicates that individuals in the SLD Reading group who had better general reasoning skills also had better reading comprehension skills. Individuals with SLD in reading may rely on their reasoning skills more to comprehend what they read when compared to individuals without SLD in reading.

Results for Math Calculation Skills

The backward selection procedure was completed with individual groups for Math Calculation Skills (MCS) to determine which cognitive ability factors had direct influence on math calculation skills within each group. The results of the backward selection procedure indicated that direct effects from RQ, PS, and RN were statistically significant for the Norm

group, RQ and PS were statistically significant for the SLD Reading group, Gc and Gsm were statistically significant for the SLD Math group, and RQ and PS were statistically significant for the SLD Writing group. These differences suggest potential heterogeneity in which cognitive abilities are related to MCS among the groups. Interestingly, the relations between cognitive abilities and MCS were quite different for the SLD Math group compared to the other groups. For the other groups, quantitative reasoning skills and processing speed skills were related to MCS, but it the SLD Math group background knowledge and short-term memory skills were related to MCS.

Results for all MCS analyses are presented in Table 12. First, MCS was regressed simultaneously on RQ, PS, RN, Gc, and Gsm in the multi-group model, with all paths freely estimated across groups (Table 12.1). Adding an equality constraint on the path from Gc to MCS resulted in a nearly statistically significant degradation in model fit (Table 12.2). Allowing the SLD Math group to be freely estimated while other groups were constrained also led to a nearly statistically significant improvement in model fit (Table 12.2a). Adding an equality constraint on the path from RQ to MCS also resulted in a statistically significant degradation in model fit (Model 4), and allowing the SLD Math group to be freely estimated on this parameter led to a statistically significant improvement in model fit (Table 12.4c).

The sensitivity analysis had similar results, but several other paths also had statistically significant changes in model fit. Results were similar for Gc, where releasing the equality constraints for all groups led to a nearly statistically significant change in model fit (Table 12.8), and allowing only the SLD Math group to be freely estimated on Gc also resulted in a nearly statistically significant change in model fit (Table 12.8a). There was a statistically significant improvement in model fit by allowing RN to be freely estimated for all groups (Table 12.9),

which did not occur in the previous analysis. Allowing RN to be freely estimated for the SLD Reading and SLD Math groups individually resulted in statistically significant improvements in model fit (Table 12.9b and 12.9c). Results were similar for PS, where there was an overall improvement in model fit when PS was freely estimated (Table 12.10), which did not occur in the previous analysis. Again, allowing the path from PS to MCS to be freely estimated for the SLD Reading and SLD Math groups individually resulted in statistically significant improvements in model fit (Table 12.10b and 12.10c). When the equality constraint from RQ to MCS was released, there was a statistically significant change in model fit (Table 12.11), and again allowing the SLD Reading and SLD Math groups to be freely estimated resulted in statistically significant improvements in model fit (Table 12.11b and 12.11c) Finally, allowing Gsm to be freely estimated across groups also led to a statistically significant improvement in model fit (Table 12.12). Allowing the only SLD Math group to be freely estimated on this path resulted in a statistically significant improvement in model fit (Table 12.12a).

There were several important differences between the two analyses performed above. Adding constraints individually to the model did not result in many differences between groups, but constraining all groups to be equal on all paths and individually releasing equality constraints did result in several statistically significant differences between groups. This likely occurred because of an interaction between groups in the pattern of equality constraints added to the structural paths. Because of these differences, several different models were tested to determine the best-fitting model.

All differences from the sensitivity analysis were included in an initial final model. In this model the path from Gc to MCS was freely estimated for the SLD Math group and constrained to be equal for all other groups, and the path from Gsm to MCS was freely estimated

for the SLD Math group and constrained to be equal for all other groups. The paths from PS to MCS, RQ to MCS, and RN to MCS were freely estimated for the SLD Reading and SLD Math groups and constrained to be equal for the Norm and SLD Writing groups. When this model was estimated, the path from Gc to MCS was negative and statistically significant for the Norm, SLD Reading, and SLD Writing groups. Here, Gc appears to be acting as a suppressor variable, because it is unlikely that *lower* scores for Gc would be related to *higher* scores for math calculation skills. Based on this, the path from Gc to MCS was set to zero for these groups, but continued to be freely estimated for the SLD Math group. This did result in a statistically significant degradation in model fit, but because the relation would be negative without the constraint, this path was set to zero for these groups were set to zero (Table 12.13a).

Next, equality constraints were added to other structural paths to determine if there were statistically significant degradations in model fit. The sensitivity analysis indicated that these paths may be statistically significantly different from one another, but the initial analysis did not indicate any differences between groups. It is possible in the sensitivity analysis that the equality constraint from RQ to MCS across all groups was forcing differences into other paths because there was such a large difference in the magnitude of the path from RQ to MCS in the SLD Math and SLD Reading groups. Adding equality constraints across all groups to the paths from RN to MCS, PS to MCS, and Gsm to MCS did not result in statistically significant changes in model fit (Table 12.13b-12.13d). This suggests that these paths are equal across groups, which was expected because there were no differences in the initial analysis where structural paths were individually constrained to be equal across groups.

Adding an equality constraint across all groups on the path from RQ to MCS did result in a statistically significant degradation in model fit (Table 12.13e), suggesting that the SLD

Reading and SLD Math groups are different from each other on this path. The equality constraint on the path from RQ to MCS may have been forcing differences in other structural paths in the sensitivity analysis. The final model was estimated allowing the SLD Math group to be freely estimated on the paths from Gc to MCS and RQ to MCS, and the SLD Reading group was freely estimated on the path from RQ to MCS (Table 12.14).

The final model included equal structural paths across groups for RN to MCS, PS to MCS, and Gsm to MCS. The path from Gc to MCS was freely estimated for the SLD Math group only and set to zero for all other groups, and the path from RQ to MCS was freely estimated for both the SLD Reading and SLD Math groups and set to be equal for the Norm and SLD Writing group. The final model for MCS is presented in Figure 5, and the total, direct, and indirect effects from this model are presented in Table 13. Again, *g* had large effects on MCS, but these effects were indirect and were mediated by more specific abilities. The paths from Gsm to MCS and PS to MCS were statistically significant and equal across all groups. When interpreting standardized effects, both of these paths had moderate relations to MCS. Additionally, the indirect effect of Gs on MCS was moderate for all groups, and this includes effects from both PS and RN.

The path from RQ to MCS had the largest differences between groups. This magnitude of this path was larger for the SLD Reading group, and smaller for the SLD Math group when compared to the Norm and SLD Writing groups. In fact, the path from RQ to MCS was not statistically significant for the SLD Math group, but approached statistical significance (p = .056) and the standardized path would be considered a large effect (β = .344). Consequently, Gf also had large indirect effects on MCS across groups. Quantitative reasoning skills were important for all groups except those with a SLD in math.

Also, the path from Gc to MCS was not statistically significant for the SLD Math group, but this effect would be moderate according to the standardized parameter (β = .183). Finally, the path from RN to MCS was not statistically significant for any of the groups, although the effect of RN was small according to the rules of thumb (β's ranged from .061-.098 across groups).

Results for Applied Math

The effects of specific and broad CHC abilities on Applied Math beyond *g* were investigated separately first within each group. For Applied Math (AM), paths from K0 and RQ were statistically significant for the Norm group, only the path from RQ was statistically significant for the SLD Reading group, paths from K0 and RQ were statistically significant for the SLD Math group, and paths from K0, RQ, and VZ were statistically significant for the SLD Writing group. VZ approached statistical significance for the SLD Math group (p = .07). Because this was also statistically significant for the SLD Writing group the path from VZ to AM will also be individually tested for the SLD Math group in subsequent models.

Results for all analyses with AM are presented in Table 14. MCS was regressed simultaneously on K0, RQ, and VZ in the multi-group model and freely estimated across all groups (Table 14.1). Adding equality constraints to the path from K0 to AM did not result in a statistically significant change in model fit (Table 14.2-14.2a). Adding equality constraints on the path from RQ to AM did result in a statistically significant degradation in model fit (Table 14.3). When this path was freely estimated for the SLD Reading and SLD Math groups individually, there was a statistically significant improvement in model fit (Table 14.3c-3d). Similarly, adding an equality constraint on the path from VZ to AM resulted in statistically significant degradation in model fit (Table 14.4). Allowing the SLD Writing group to be freely estimated on the path from VZ to AM resulted in a statistically significant improvement in model fit, but allowing the

SLD Math group to be freely estimated on this path did not result in a statistically significant improvement in model fit, although it approached statistical significance (p = .087, Table 14.4a-14.4c). Results from the sensitivity analysis were similar. When the equality constraint on the path from RQ to AM was released for the SLD Reading and SLD Math groups individually, there was a statistically significant improvement in model fit (Table 14.7b and 14.7c). When the equality constraint on the path from VZ to AM was individually released for the SLD Math and SLD Writing groups, this led to a statistically significant improvement in model fit (Table 14.8a and 14.8b).

The final model included an equality constraint for all groups on the path from K0 to AM. Because there were statistically significant differences in model fit when the path from RQ to MR was freely estimated for the SLD Math and SLD Reading groups, this path was freely estimated for both of these groups, and this path was constrained to be equal for the Norm and SLD Writing groups. Also, because there was a statistically significant improvement in model fit when the path from VZ to AM was freely estimated for the SLD Math and SLD Writing groups in the sensitivity analysis, this path was freely estimated for these two groups and constrained to be equal for the Norm and SLD Reading groups. In this model (Table 14.9), the path from VZ to AM approached significance for the SLD Math group (p = .059), but the unstandardized parameters for the path from VZ to AM were very similar for the SLD Math and SLD Writing groups (.196 and .204, respectively). It is possible that this path was not statistically significant for the SLD Math group due to power, so a model was estimated where the path from VZ to AM was constrained to be equal for the SLD Math and SLD Writing group (Table 14.9a). There was not a statistically significant degradation in model fit, indicating the effects of VZ on AM were equal for the SLD Math and SLD Writing groups. This effect was also statistically significant. A

second test was completed where an equality constraint was placed on the path from RQ to AM for the SLD Reading and SLD Math groups. This equality constraint did result in a statistically significant degradation in model fit, indicating that the effect of RQ on AM is different for the SLD Reading and SLD Math groups, in addition to being different from the Norm and SLD Writing groups (Table 14.9b).

The final model for AM is presented in Figure 6, and the total, direct, and indirect effects for the AM model are presented in Table 15. The final model again demonstrated more specific CHC abilities are important for math beyond *g*, and indicated that group membership moderated several of these effects. The direct effect from K0 to MCS was statistically significant and equal across all groups. RQ needed to be freely estimated for the SLD Reading and SLD Math groups. The effects of RQ were stronger for the SLD Reading group when compared to the Norm and SLD Writing groups, and the effects of RQ were weaker for the SLD Math group when compared to the Norm and SLD Writing groups. These findings were similar to the findings from math calculation. Additionally, the path from VZ to AM was statistically significant for the SLD Math and SLD Writing groups (b = .203) only, suggesting that VZ was only important for these two groups.

Results for Written Expression

The effects of CHC abilities on Written Expression were investigated separately within each group. In the model for Written Expression (WE), the statistically significant paths from the backward selection procedure for each group indicated that Gc, RN, and VZ were statistically significant for the Norm group, RN and RQ were significant for the Reading group, Gsm and Gc were significant for the SLD Math group, and RN, RQ, and MS were statistically significant for the SLD Writing group.

A model which regressed WE simultaneously on Gc, VZ, Gsm, RN, RQ, and MS was estimated (Table 16.1). Each structural path was individually constrained to be equal across all groups. When Gc was constrained to be equal across groups, this led to a statistically significant degradation in model fit (Table 16.2). Allowing the SLD Math group to be freely estimated on the path from Gc to WE resulted in a statistically significant improvement in model fit (Table 16.2b), suggesting that the path from Gc to WE is different for the SLD Math group. Adding equality constraints to the other structural paths did not result in a statistically significant change in model fit (Table 16.3-16.7), although Gsm approached significance. The sensitivity analysis did not indicate any statistically significant difference when releasing the equality constraints on individual paths (Table 16.7-16.14).

A final model was created where all structural paths were included in the model and constrained to be equal across groups. Because allowing the path from Gc to WE to be freely estimated for the SLD Math group resulted in a statistically significant change in model fit in the previous analysis, the only path which was freely estimated was the path from Gc to WE for the SLD Math group. Also, because allowing the path from Gsm to WE resulted in a nearly statistically significant change in model fit for the SLD Math group, this path was also freely estimated for the SLD Math group (Table 16.15). The final WE model is presented in Figure 7, and total, direct, and indirect effects are presented in Table 17. In this model, only the paths from Gc, RQ, and RN were statistically significant. The path from VZ to WE approached statistical significance (p = .08), and neither Gsm nor MS were statistically significant. This suggests that all groups relied on background knowledge, quantitative reasoning, and rapid naming skills on the written expression tests. The only difference between groups was the path from Gc to WE for the SLD Math groups. The path from Gsm to WE was not statistically significant for the SLD

Math group, so a model was tested where this path was constrained to be equal across all groups. There was not a statistically significant change in model fit when this constraint was added, suggesting groups were equal on this path (Table 16.15a).

Because both Gsm and MS were included in the model, two alternative models were tested. It is possible that including both Gsm and MS in the model may be masking the magnitude of the effects from each. Potentially, only including a path to WE from one of these two memory factors will change the magnitude of the effects from memory. Two models were estimated to determine if including a path from either Gsm or MS to WE in the model resulted in statistically significant differences in model fit. In the first model (Table 16.16), the path from MS to WE was set to zero for all groups, and the path from Gsm to WE was set to be equal across groups. In the second model (Table 16.17) the path from Gsm to WE was set to zero and the path from MS to WE was constrained to be equal for all groups. Neither model resulted in statistically significant change in model fit from the final model (Table 16.15). The relation between Gsm and WE was statistically significant when the path from MS to WE was constrained to zero for all groups, and the relation between MS and WE was statistically significant when the path from Gsm to WE was constrained to be zero for all groups. This suggests that short-term memory processes are significantly related to WE, but regressing WE on both Gsm and MS in the structural model may mask the effects of each on WE. The model where the path from the path from MS to WE was set to be equal across groups and the path from Gsm to WE was removed from the model fit better according to the AIC, BIC, and aBIC. Gsm did have large indirect effects on WE, however.

The alternative WE model is presented in Figure 8, and estimates from the alternative WE model are included in Table 18. In addition to the path from MS to WE being statistically

significant in this model, it is notable that the path from VZ to WE was also statistically significant. This effect only approached statistical significance in the previous final model.

Chapter V: Discussion

The purpose of this study was to determine if the relations between cognitive abilities and academic skills were similar among groups of children with and without SLDs. A large body of research has examined differences in cognitive skills for children with SLD (e.g., Johnson et al., 2010), but there has been little work examining which cognitive abilities are related to different academic skills in children with SLD. Rather, most of this research has been conducted with normally developing children (Floyd et al., 2007; McGrew & Wendling, 2010; Taub et al., 2008). Although this study was primarily an exploratory analysis, the results suggest several important findings regarding the similarities and differences in cognitive—achievement relations among groups of children with and without SLDs. The similarities and differences offer several important practical and theoretical implications regarding the nature of learning disabilities.

This discussion is composed of five major sections, including (a) the original hypotheses provided in the methods section will be reviewed, (b) a discussion regarding the similarities and differences between groups will be examined for each of the academic skills included in the analysis, (c) implications for theory and practice will be reviewed, (d) the strengths and limitations of the current study will be presented, and (e) general conclusions and future directions will be explored.

Support of Hypotheses

Hypothesis 1: The factor loadings among the Norm and SLD groups will be invariant for the CHC model.

The first hypothesis was supported in the current research. The factor loadings of the WJ-III were statistically equal across groups. When first identifying the model, there were some group-specific differences which were required to obtain adequate model fit (e.g., cross

loadings), but when these were included in the multi-group model there were no differences between groups in the tests of factorial invariance. This suggests that the CHC latent constructs were the same across groups of children with and without SLDs, both at the general, broad, and narrow ability level.

This finding supports previous research which has examined the factorial invariance of other cognitive ability tests for individuals with SLDs (Bowden et al., 2008). This study also extends previous research because it uses a clinical sample of children and adolescents, whereas previous research used a clinical sample of college students. The separation of different types of SLD in reading, mathematics, and writing is also an extension of previous research, in which all individuals with SLD were included in a single group. In general, it can be assumed that the CHC constructs measured by the WJ-III are similar for groups of children and adolescents with and without SLDs. This step was important because it was necessary for the comparison of structural relations between cognitive abilities and academic skills in the second portion of this study.

Hypothesis 2: Narrow cognitive abilities will be most strongly related to academic skills.

This hypothesis was generally supported as well. For the most part, narrow cognitive abilities had direct effects on academic skills. Direct effects from broad cognitive abilities were less common, and there were no direct effects from *g*. It is possible that if the academic skills were more broadly defined (e.g., basic reading skills and reading comprehension were included on a single reading factor) then the broad cognitive abilities may have been better predictors of the academic skills. The academic skills used in this study were narrowly defined (e.g., Basic Reading Skills, Math Calculation Skills), and this supports the notion that narrow cognitive skills best predict performance on narrow academic skills (McGrew & Wendling, 2010). It may also be

the case if the academic constructs were even more narrowly defined (e.g., direct influences from narrow cognitive abilities to academic subtests) there may have been even more narrow ability effects. There were no instances where *g* was identified as a statistically significant direct predictor of the academic skills examined in this study, consistent with previous research on cognitive—achievement relations (Elliot et al., 2010, Floyd et al., 2007; Taub et al., 2008). When considering total effects; however, *g* had the largest effects on academic skills. This finding is a function of the model specification, and positive parameter estimates. In the CHC model used in this study, *g* was placed at the apex of the hierarchical model. This is similar to including *g* in the first step of a hierarchical regression, with the broad abilities entered second, and narrow abilities entered last in the equation. Thus, all *g* effects were indirect and were mediated through the broad and narrow abilities.

When broad abilities did have direct effects on the academic skills, the effects tended to occur for broad abilities which have been identified as important predictors of the academic skill in previous research. Additionally, when relations between broad abilities and academic skills did occur, they tended to occur for broad abilities which were not identified by multiple narrow abilities. For example, Gsm was directly related to BRS in the current study. Previous research has identified that working memory, one of the components of Gsm, is related to BRS (Floyd et al., 2008). In the current study, the inclusion of a specific working memory factor was not possible because there were not two subtests included in the analysis which measured working memory. Rather, the working memory subtest used in the model (Numbers Reversed) loaded directly on Gsm. Without a specific working memory factor, it was not possible to determine if the relation between Gsm and BRS is best explained by the specific working memory factor or broad Gsm.

It is interesting to note that some narrow abilities, namely K0 and RQ were related to nearly every academic skill included in the analysis. Table 19 presents a summary of cognitive—achievement relations identified in the current study. This table clearly shows that K0 and RQ were pervasively directly related to academic skills, and RQ was directly related to all academic skills included in these analyses. When considering the skills measured by K0 and RQ, it is possible that these two narrow cognitive skills may be more related to academic achievement than other cognitive abilities. K0 may be related to school experience, verbal comprehension, and cultural experience because it is a measure of background knowledge, including tests of general knowledge (General Information) and knowledge which would be obtained through school experience (Academic Knowledge). It should be noted, though, that K0 is broader than academic knowledge. It may also be considered the ability to infer correlates between pieces of information a person has been exposed to that are more verbal in nature. K0 measures a broad range of knowledge, more than would be typically taught directly in school. Experience in the environment is essential to acquire knowledge, but the ability to infer relations between different pieces of knowledge is not acquired within the environment itself. RQ may partially be related to academic experience as well because of the reliance on numbers. But, the ability to reason with numbers is one form of inferring relations between stimuli. Learning specific mathematical operations may be related to school experience, but performance on RQ tasks does not require specific math calculations, rather it requires one to be able to determine relations between stimuli, which are numbers in the case of RQ. The reasoning skills required to complete the items on RQ are broad, not more specific skills related to completing specific math calculation problems. However, it is possible that these two abilities were more consistently related to the academic skills because they are closely aligned with the investment of general and broad

abilities into verbal and quantitative domains, and these thus crystallize to form these broad abilities that more readily affect achievement.

Only Gsm and Gc were the other broad abilities which was related to an academic skills. However, like Gsm, Gc only included K0 and RN as specific narrow abilities (the residual for LS was constrained to zero for all groups, which effectively removes that specific factor) and both of the tests which would comprise a verbal development factor loaded directly on Gc. Gc was related to WE, but it is possible that a narrow ability not included in the model would have accounted for the relation between Gc and WE, but this would have to be investigated in future research.

In sum, the relations between cognitive abilities and academic skills tended to occur at the narrow ability level. Additionally, several of the differences among groups in cognitive—achievement relations occurred at the narrow ability level. These will be examined in more detail in the next section.

Hypothesis 3: There will be differences in cognitive—achievement relations among groups. The greatest differences in cognitive—achievement relations will be related to SLD groups for whom the academic skill is related to their disability (e.g., SLD Reading will have significant relations between different cognitive abilities and Basic Reading Skills).

The results of this study are mixed for this hypothesis. There were many similarities in the cognitive—achievement relations among the groups of children and adolescents with and without SLDs, but there were some important differences as well. Overall, the general pattern of cognitive—achievement relations was similar among groups, which is not consistent with some previous research. For example, Elliot et al. (2010) found that a nearly completely different set of cognitive skills predicted decoding for children with reading difficulties when compared to

children without reading difficulties. It is important to note that the analysis used by Elliott et al. was similar to the backward selection procedure used in this study. However, Elliot et al. did not compare the equality of unstandardized paths from cognitive abilities to decoding in a multi-group model, thus moderation of those effects was assumed. Unstandardized parameters were compared for equality in the current study. It is possible that testing these differences formally in the study by Elliot et al. would not have resulted in such disparate differences between groups. By formally testing the equality of cognitive—achievement relations among groups, it was possible to identify whether there were statistically significant differences *between* groups, and not only if they were statistically significant *within* groups. If the cognitive—achievement relations in this study were identified only through the backward selection process, there would have been many more differences in cognitive—achievement relations among the groups. On the other hand, it is also likely that some of the statistical tests in this study lacked power to detect small and even moderate differences in the magnitudes of cognitive—achievement relations across groups.

Nevertheless, testing the equality of the relations across groups provided a more formal test of moderation. These formal tests, however, did indeed suggest that group membership moderated the cognitive—achievement relations. That is, the direct effects of some cognitive abilities on academic skills depended on group membership. The similarities and differences are summarized separately for each of the academic areas investigated.

Basic Reading Skills

The cognitive abilities which predicted performance on BRS were similar among groups, but there were some important specific differences as well. The direct influences of Gsm, K0, and RQ on BRS were statistically significantly different from zero, and the influences of each

were equal in magnitude across all groups. Group membership did not moderate the effects of short-term memory, background knowledge, and quantitative reasoning on BRS. Considering the magnitude of these effects, both Gsm and K0 had similar standardized effects across groups, suggesting that these abilities have comparable influences on BRS, and these influences were large according to rules of thumb by Keith (2006); all standardized structural paths across groups were greater than .25. RQ had smaller standardized effects across groups, suggesting that it was less influential, although still statistically significant. All of these relations are consistent with previous research on reading and the WJ-III (Floyd et al., 2007), as well as the review of cognitive—achievement relations by McGrew and Wendling (2010).

It is not surprising that both background knowledge and short-term memory skills influence reading skills. K0 may be related to basic reading because it represents a broad base of knowledge an individual has obtained from experience in their environment, and it is highly related to cultural experience (Schneider & McGrew, 2012). Reading is acquired within one's culture (especially in the institution of education), so it makes sense that individuals who are able to better acquire background knowledge from their environment would also be able to acquire basic reading as well. Also, as Floyd et al. (2007) point out, reading and writing were both included as part of Gc in Carroll's (1993) original taxonomy of cognitive abilities because the acquisition of these skills are highly related to cultural experience. It is also important to point out the possible reciprocal nature of reading skills and knowledge, where those with better reading skills will likely obtain a broader base of background knowledge from what they read (Floyd et al., 2007).

The relationship between Gsm and reading skills is consistent with previous research as well (Floyd et al., 2008; Swanson & Alexander, 1997) which has interpreted working memory

through the model proposed by Baddeley (Baddeley & Hitch, 1994). In this model, working memory is governed by an overall executive processing system which controls two types of memory storage, the phonological loop (memory for language) and the visual-spatial sketchpad (memory for visual information). The effects of short-term memory on reading are important because the information used to decode words would be stored in the phonological loop. Individuals with better memory skills related to the phonological loop may be able decode words easier because there is less overall cognitive load in the decoding process. The tests of memory on the WJ-III are all auditory and language based, Gsm does not include visual memory tasks. Nevertheless, individuals who have more efficient short-term memory processes in general are likely able to decode words more efficiently because of a lower cognitive load.

The statistically significant influence of RQ indicates that reasoning skills may also be influential in reading, including basic reading skills, which has also been identified in previous research (McGrew & Wendling, 2010). However, relatively little work has been done in the examination of RQ and BRS relations. Interestingly, Swanson and Alexander (1997) examined the relations between several different cognitive processes and reading skills. They found that *g* was a better predictor of decoding skills for individuals with SLD in reading, while phonological processing was better predictor of decoding for normally developing readers. This indicates that more general reasoning abilities (of which RQ would be a part) may be important for basic decoding tasks, and this relation might be strongest for individuals with SLD in reading. The normally developing readers in Swanson and Alexander's study were between 8 and 12 years old. It is possible that *g* has more of an effect on basic reading skills for younger children who are just beginning to develop their reading skills. If a child has reading skills which have become automatic, the influence of *g* would be minimal. But, for a child who continues to have difficulty

with basic reading skills, the cognitive load placed on the decoding process may show more influence of *g*. However, no differences in the magnitude of the relation between RQ and BRS were found in this study, but there were differences in the relation between RQ and reading comprehension, which will be discussed below.

In the BRS analysis, there were two cognitive—achievement relations which were moderated by group membership. First, the SLD Writing group had a direct effect from MS on BRS, which suggests that individuals with SLD in writing may rely directly on narrow memory span skills in addition to broad Gsm. This may be related to differences in how children with SLD in Writing process orthographic symbols, which includes both the decoding of words for reading and the encoding of words for writing. The MS tests on the WJ-III may be related to the phonological loop based on content (Memory for Words, Memory for Sentences). Another explanation for this is that the SLD Writing group did have more males than females, and orthographic coding skills are known to be lower in males with writing difficulties (Berninger, Nielsen, Abbott, Wijsman, & Raskind, 2008). It is possible that this is partially related to a gender difference, or this may be a disability by gender interaction.

The only difference for the SLD Reading group on BRS was a statistically significant relation from PS to BRS. PS was not related to BRS for any of the other groups, but this effect was moderate for the SLD Reading group. The studies by Floyd et al. (2007) and McGrew and Wendling (2010) found that processing speed was more strongly related to BRS for children under the age of eight. Additionally, Elliot et al. (2010) found that processing speed was related to basic reading skills for children with reading difficulties, but not for children without reading difficulties. The findings from the current study are consistent with this previous research. The current study does not differentiate between age groups, but this may provide some insight into

the developmental process of reading for individuals with SLD. If individuals with SLD in reading have delayed decoding skills, then processing speed may continue to predict decoding skills across development because basic decoding skills do not become automatic. Processing speed may continue to be related for individuals with decoding problems because it would influence the speed at which individuals with SLD in reading are able to decode words. For individuals without SLD in reading, processing speed becomes less important as they develop basic decoding skills because decoding becomes an automatic process. Therefore, processing speed no longer differentiates between normally developing readers. These findings suggest that processing speed may continue to predict reading skills for children and adolescents with SLD in reading, long after processing speed no longer differentiates decoding skills for children without SLD because their decoding skills have become automatic.

An alternative explanation for this relation may be related to visual discrimination skills. The PS factor in this study was indicated two tests that require visual discrimination skills. Visual discrimination is an important component of sight word recognition, and deficits in these areas may be related to surface or orthographic dyslexia, which may subsequently affect rapid naming skill (Bowers & Wolf, 1996). Future research may want to consider testing whether this factor influences the specific skill of sight word recognition (e.g., Letter-Word Identification) as opposed to a broad basic reading skills. Similar to different subtypes of SLD in math (Geary, 2003), individuals with SLD in reading may have a deficit in phonological processing, rapid naming speed, or both (Bowers & Wolf, 1996). For example, those with surface dyslexia may have more difficulties with sight word recognition, visual discrimination, and visual memory, whereas those with phonetic dyslexia may show more difficulty with phonological processes and would likely struggle more pseudoword decoding, such as the Word Attack subtest from the WJ-

III. As Bowers and Wolf (1996) argue, those with deficits in both areas will have more impaired reading skills than individuals with deficits only one of these areas.

Finally, MA was related to BRS in the backward selection, but only for the Norm group. However, when included in the multi-group analysis the magnitude of the relation between MA and BRS was not statistically significantly different among all groups, and the path from MA to BRS was not statistically significant from zero. This is likely due to the lack of age differentiation in this study. Previous research (Floyd et al., 2007) found that MA was related to BRS for normally developing children, but only for those between the ages of 5 and 6. It is possible that this relation would have been stronger if different age groups could have been included in the analysis.

Reading Comprehension

The results for RC were consistent across the Norm, SLD Math, and SLD Writing groups, indicating that the influence of cognitive abilities on reading comprehension is similar for individuals without SLD in reading. However, there was an important difference in this model for the SLD Reading group, where the relation between RQ and RC was statistically significantly larger for the SLD Reading group compared to the other groups.

In the final multi-group model, only the paths from K0 to RC and RQ to RC were statistically significant. Paths from PS and RN to RC were included in the model, but were not statistically significant for the final model, although this may have been due to a lack of power. The influence of K0 on RC was largest for all groups, indicating that the influence of background knowledge on reading comprehension is most important for individuals both with and without SLDs. The relation between K0 and RC is consistent with previous research, where background knowledge is strongly related to comprehension (Evans et al., 2001; Floyd, Bergeron, & Alfonso,

2006; McGrew & Wendling, 2010), and the direction of the path from background knowledge to reading comprehension is supported in findings that background knowledge drives positive changes in reading comprehension (Reynolds & Turek, 2012).

The relation between RQ and RC has not been comprehensively explored in research. RQ is a narrow ability of under the umbrella of Gf (Schneider & McGrew, 2012), and McGrew and Wendling (2010) suggest that the relation between Gf and RC is tentative until more evidence is collected. The effects of RQ on RC were smaller than the effects of K0 on RC, but this relation was statistically significant, and large in magnitude, for all groups. RQ had a larger effect on RC for the SLD Reading group when compared to other groups, and this difference was statistically significant. This finding makes sense theoretically, since it is possible that individuals with better reasoning abilities would rely more on their reasoning skills to comprehend meaning in text they have difficulty decoding (Nation & Snowling, 1997). Individuals from other groups would not need to rely on reasoning processes as much to comprehend text because they do not have difficulty with decoding. Rather, they are able to use background knowledge to understand what they read. For individuals with SLD in reading, relying more on reasoning may help compensate for imperfect decoding skills. It is important to note that background knowledge is still important for the SLD Reading group. Even if some individuals with SLD in reading have strong reasoning skills, background knowledge is still necessary to adequately comprehend what is read. Reasoning skills may help put the pieces of what they read together into a coherent message, but background knowledge is still essential for an overall understanding.

Finally, paths from PS to RC and RN to RC were statistically significant for the Norm and SLD Writing groups, respectively, in the backward selection. When included in the multi-group model, these paths were not statistically significant for these groups or any other groups in

the final model. Rapid naming has been identified as a predictor of basic reading skills across a wide range of studies (Bowers & Ishaik, 2003), and rapid naming skills in kindergarten is a significant predictor of reading skills in first and second grades (Schatschneider, Fletcher, Francis, Carlson, Foorman, 2004). The lack of an effect here could be due to the rapid naming tasks used on the WJ-III. The RN factor here is composed of Retrieval Fluency and Rapid Picture Naming. Other studies have used different types of stimuli for RN tasks, such as a mixture of pictures and letters (e.g., Schatschneider et al., 2004) or letters and numbers (e.g., Bell, McCallum, & Cox, 2003). Rapid naming of orthographic symbols, as opposed to pictures, may be more related to reading because of shared content of the tasks. Also, McGrew and Wendling (2010) found that PS was related moderately to RC at younger ages, but was less related to RC for older age groups. It is possible that the lack of age differentiation affected the relation between PS and RC. For example, Evans et al. (2001) show that processing speed is only slightly related to reading comprehension at younger ages, but this relation becomes negligible over the course of development. Based on the findings here, it appears that knowledge and reasoning are the most important skills for comprehending written text. Reasoning, however, is even more important for those with a SLD in reading.

Math Calculation Skills

There were several important differences among the groups for the model with math calculation skills. First, there appeared to be a complex interaction between the cognitive—achievement relations when equality constraints were included on the structural paths in the model. This appeared to be related to RQ, mostly because there was a very large difference in the unstandardized paths from RQ to MCS for the SLD Reading and SLD Math groups. The path from RQ to MCS was larger for the SLD Reading group, and it was smaller for the SLD Math

group. In fact, the path from RQ to MCS for the SLD Math group was not statistically significantly different from zero (although it approached statistical significance, and the standardized path coefficient was above .30, which would be considered a large effect). This difference suggests that quantitative reasoning skills were not as strongly related to MCS for the SLD Math group when compared to the other groups. This is not a surprising finding, because difficulty with math reasoning would be expected for individuals with SLD in math. Children with a SLD in math would likely try to draw upon other cognitive abilities (or other types of resources) apart from quantitative reasoning skills when working on math tasks. Indeed, although this study focused on the interrelations among the skills, this group did have lower means in measures of both RQ and MCS (see Table 5).

The relation between RQ and MCS was stronger for the SLD Reading group when compared to other groups. This may indicate that individuals with SLD in reading also rely more strongly on quantitative reasoning skills to complete math problems. Whereas others may be able to retrieve math facts from memory, it is possible that individuals with SLD in reading may rely more on quantitative reasoning to solve each problem individually. Additionally, a larger proportion of variance in MCS was accounted for by the cognitive abilities included in the model for the SLD Reading group and SLD Writing group when compared to the Norm and SLD Math groups. This is important because it suggests that these two groups rely more on cognitive abilities than other factors to solve math calculation problems, and most of the resources they use to complete the problems are cognitive. Less variance in MCS is accounted for in this model for the Norm and SLD Math groups, and other resources not accounted for in this model are influencing their math calculation skills. This may suggest that individuals in the SLD Reading

group rely more on cognitive resources overall across academic skill areas, and are less influenced by extraneous resources.

The effects of PS and Gsm on MCS were statistically significant for all groups. Additionally, these relations were equal across groups, suggesting that these processes have equal influence on MCS for individuals with and without SLD. Both PS and Gsm (specifically, working memory) have been implicated as important processes for MCS across age groups (Fuchs et al., 2006; McGrew & Wendling, 2010). The relation between these two processes and MCS make sense, where individuals who are able to process numbers quickly and retrieve answers to simple math problems quickly will have an advantage when completing math calculations. Additionally, for calculations which require several steps, Gsm would be important to both recall which steps have been taken and help organize and recall the appropriate steps in the math problems.

It is important to note that McGrew and Wendling (2010) indicate that the relation between PS and MCS may be influenced partially by the tests used on the WJ-III. One of the tests for PS is Visual Matching, which requires the matching of similar numbers. One of the tests for Gsm is Numbers Reversed, which also uses numbers as test stimuli. It is possible that these relations could partially be explained by a separate method factor or Numerical Facility factor (Schneider & McGrew, 2012; McGrew & Wendling, 2010). However, this influence would not be present in the common factor when using latent variables. Rather, these influences would be between residual variances. It would be useful for future research to verify these relationships using tests which include other types of stimuli for processing speed and working memory.

Rapid Naming was statistically significantly related to MCS for the Norm group in the backward selection, but it was not statistically significant for any groups in the final model.

There was not a statistically significant improvement in model fit when the Norm group was freely estimated on RN, indicating that even though this relation was statistically significant in the backward selection for the Norm group, the unstandardized parameter from RN to MCS was essentially the same among all groups. The relation between RN and MCS makes sense because math calculation skills require the recall of math facts. For the Norm group, it is likely that these individuals have memorized basic math facts and they are able to retrieve these quickly from memory, thus the relation between retrieval fluency and MCS. However, in the multi-group analysis, the results indicated that there were no statistically significant differences between the unstandardized coefficients for the relation of RN to MCS, and this relation was not statistically significantly different from zero.

Finally, in the backward selection Gc was statistically significantly related to MCS for the SLD Math group, but not for other groups. This relation was positive and moderately large (β = .183) for the SLD Math group in the multi-group model, but not statistically significant. The relation from Gc to MCS was negative for the other groups, suggesting that it may have been acting as a suppressor variable rather than a predictor of math calculation. Gc was related to math skills for older individuals in the study by Taub et al. (2008). However, their study included both math calculation and applied problem solving, which would be a measure of broad math achievement rather than the more specific skill of math calculation skills. The synthesis of the literature by McGrew and Wendling (2010) found that Gc was consistently related to MCS after the age of nine. Other research (Niileksela & Reynolds, submitted for publication) has indicated that individuals with SLD in math also show an asset in Gc, suggesting they may use background knowledge as a compensation strategy. Such a finding would be consistent with the potential relation found in this study. That is, students with an SLD in Math may try to draw more upon

their background knowledge when performing math calculation. However, this finding is speculative until more research can be completed. The relation between Gc and MCS for the SLD Math group was not statistically significant, although the magnitude of the path was moderate. Research with a larger group may allow for a better understanding of this possible compensation strategy.

It should be noted, however, that there is some conflicting research with this finding. According to Geary (2003), as children grow older they rely less on their quantitative skills to complete math problems and instead rely more on background knowledge. This change occurs because children begin to rely less on problem-solving strategies and more on background knowledge. Specifically, as children get older they are able to retrieve math facts from memory. Children with SLDs, however, continue to use problem-solving strategies (e.g., implicit counting, Geary, Widaman, Little, & Cormier, 1987). The findings from the current study are inconsistent with this notion. Individuals in the SLD Math group do not rely on strongly on RQ for MCS, but they do not rely strongly on Gc either. Individuals who have difficulty with quantitative skills but may have learned some strategies or algorithms for completing math problems, this would implicate Gc as a possible mechanism for compensating for deficits in RQ. The answer to this difference may be due to different subtypes of mathematics disability. Geary (2003) identifies that there is a procedural subtype (difficulty with mathematics procedures or algorithms), a semantic memory subtype (difficulty retrieving math facts), and a visuospatial subtype (difficulty with visual-spatial representation of numerical relationships). It is possible that a mixture of these subtypes is present in the sample used for this study, thus masking differential effects of these subtypes.

Applied Math

The cognitive abilities which were related to AM for all groups were K0 and RQ. K0 was equal across groups and had a relatively small effect. This finding is not surprising in light of previous research, which indicates that background knowledge is important for AM due to the requirements for language and cultural understanding to complete applied math problems (McGrew & Wendling, 2010). Not surprisingly, RQ was strongly related to MR. This relation was equal for the Norm and SLD Writing groups. Similar to the MCS model, RQ had to be freely estimated for the SLD Reading and SLD Math groups. For SLD Reading, the relation between RQ and AM was larger when compared to the other groups, and for the SLD Math group the relation between RQ and AM was smaller when compared to the other groups. Both of these differences were statistically significant. Again, this finding indicates that the relation between RQ and AM was moderated by group, where individuals in the SLD Reading group relied more on quantitative reasoning when completing applied math problems compared to other groups, but individuals in the SLD Math group relied less on quantitative reasoning.

An interesting finding from the AM analysis was that the relation between VZ and AM was statistically significant for the SLD Math and SLD Writing groups, but not for the Norm or SLD Reading group. Visualization skills have been implicated as important for the development of math skills, but this has not been consistent across research. McGrew and Wendling (2010) found no significant relations between Gv abilities and academic skills in their review of the research. However, they also point out that the visual-spatial skills which could be related to mathematics skills may not be adequately represented in current batteries of cognitive abilities. They also point out that the variables often measured in achievement batteries do not sample items from higher-level mathematics areas (e.g., geometry, trigonometry, calculus), which may require more visualization skills.

It is known that visualization skills are important to success in science, technology, engineering, and mechanical (STEM) domains (Wai, Lubinski, & Benbow, 2009), all areas for which mathematics reasoning are important. Visual-spatial deficits in children are related to difficulties in the development of math skills (Geary, 1993), and visual-spatial deficits have also been found in adolescents with SLD in math (Swanson, 2011). Additionally, the visuospatial sketchpad component of working memory (Baddeley & Hitch, 1994) predicts the development of math skills for elementary students (Geary, 2011). When considered together, this research does suggest that visualization skills are important for mathematics skills, but these skills may not be adequately assessed on current cognitive ability tests. However, more research in this domain would be necessary to better understand the relationship between visualization skills and the different types of math skills. When considering the results of this study, it makes sense that individuals with SLD in math may perform better if they have better visualization skills. For example, if some individuals with SLD in math are able to better visualize the components of an applied math problem, this may provide an advantage over individuals with difficulty with visualization, even when calculation skills are similar. Individuals with SLD in math may rely more on visualization skills in order to compensate for deficits in quantitative reasoning.

The relation between VZ and AM was also statistically significant for the SLD Writing group. It is not clear why the SLD Writing group would have a similar path from VZ to MR, but this could be due to orthographic coding skills related to writing (Abbott & Berninger, 1993; Berninger & O'Malley May, 2011). Berninger & Amtmann (2003) review research that suggests orthographic coding is especially important for written language because it involves the processing of orthographic codes, which is different from reading difficulties because reading requires the processing of both orthographic and phonological codes. Some previous research

suggests that there may be deficits in Gv abilities for individuals with SLD in writing, specifically visual memory (Niileksela & Reynolds, submitted for publication), although the SLD Writing group appears to have scores similar to other groups in the tests related to the Gv factor. It is possible that visualization skill differences in the SLD Writing group also provides an advantage in math reasoning, where better visualization skills predict better performance on applied math problems, similar to the possible connection of visualization and math reasoning the SLD Math group.

The identification of VZ as a significant predictor of mathematics skills is in contrast to other research, which has not identified Gv abilities as measured by common intelligence batteries as an important predictor of math reasoning skills in samples of normally developing children (e.g., Taub et al., 2008) or individuals with SLD in math (e.g., Proctor, in press). However, the factor used in this study was specifically visualization, which includes Spatial Relations and Block Rotation. Other studies with the WJ-III have not used this specific factor, rather they have used a more broadly defined Gv factor (e.g., Proctor, in press; Proctor et al., 2006; Taub et al., 2008).

It is important to note that several relations were not statistically significant in this analysis, namely for short-term memory or working memory. Proctor (in press) found that working memory was statistically significantly related to math reasoning skills in a sample of college students with SLD in math, and McGrew and Wendling (2010) also show that working memory is related to mathematics reasoning. The relations between K0, RQ, and VZ account for most of the variance in RQ (over 95% for all SLD groups). It is possible that these processes are accounting for variance usually attributed to short-term memory processes in other studies. More

comprehensive research including all of these areas would be helpful in determining which of these processes are most important for math reasoning skills.

Written Expression

The results for WE were very similar across all groups, and this was the only area in which the group with a SLD in the academic skill area did not show any moderated relations between cognitive abilities and the academic skill. In the backward selection analysis, six cognitive abilities were identified as statistically significant predictors of WE. However, when all of these paths were included in a single model, only three were statistically significant. The paths which were statistically significant for all groups were Gc, RN, and RQ. The relationship between Gc and writing is not surprising, because adequate writing skills would require background knowledge. However, it was interesting that this relation was not K0, as it has been in the previous analyses in this study. This suggests that other skills subsumed under Gc, such as verbal development or listening skills, play an important role in written language beyond the effects of the specific knowledge factor. The relationship between RN and written expression also makes sense theoretically, because individuals who are able to retrieve words or ideas from memory faster may be able to better express themselves in writing. However, some previous research has not found a relationship between rapid naming skills and writing. For example, Berninger, Abbott, Thomson, & Raskind (2001) found that rapid naming deficits were only associated with difficulties in reading. Again, this difference may be related to task differences, where Berninger et al. (2001) used rapid naming of letters and numbers, whereas the WJ-III rapid naming tasks are both free retrieval and rapid picture naming. It is possible that these retrieval processes do play a role in written expression. Finally, RQ was also related to WE for all groups. The relationship between fluid reasoning processes and written expression has also

been found in previous literature, but only for adolescents in one study (Floyd et al., 2008), and only for younger children in another study (McGrew & Knopik, 1993). Both of these studies used the broad Gf cluster, not the more specific ability RQ in their analyses. It is possible that RQ may have more specific effects on written expression. The use of fluid reasoning processes in written expression makes sense because individuals must organize ideas in order to express them, but clearly more research into fluid reasoning would need to be done to verify these relationships.

In the initial model, Gsm, MS, and VZ were not statistically significant for any groups. However, the standardized relations between memory (Gsm and MS) and WE were moderate, and Gsm has been identified as an important predictor of writing in previous research (e.g., Floyd et al., 2008). Two alternative models were tested which were designed to evaluate the effects of Gsm and MS on WE. First, the relation between MS and WE was removed from the model while the relation between Gsm and WE was constrained to be equal across groups. Second, the path from Gsm to WE was removed from the model and the path from MS to WE was constrained to be equal across groups. Neither of these models were statistically significantly different from the model which included Gsm and MS simultaneously in the model. Overall, the model where Gsm was removed and the path from MS to WE was freely estimated across groups fit slightly better, so this model will be discussed.

In the model where Gsm was removed, all five remaining paths were statistically significant for all groups, indicating that including both Gsm and MS in the model was affecting the size of the relations for other abilities. Gc was most important based on standardized effects, followed by RQ and RN, with MS and VZ providing the smallest effects. This model indicates that both individuals with and without SLDs rely on a wide range of skills for written expression.

The addition of VZ indicates that visual skills are important for written expression, and because writing is a complex process it is not surprising that several different processes are significant predictors of written expression.

Summary of Results

Taken together, these results provide an interesting look at the relations between academic skills and cognitive abilities. An overall summary of the results for all cognitive—achievement relations are presented in Table 19. In this table, the names of the groups are included in cells for which there were paths from the cognitive abilities to the academic skills. If the path was statistically significant for a group, the name of the group is in bold in the table. If the path was not statistically significant in the model, the name of the group is italicized. Finally, groups which are statistically significantly different from other groups are underlined, and arrows next to the group name are included to indicate if the difference in the magnitude of the relation is higher or lower.

It is clear that RQ had the most consistent effect on academic skills, both across groups and across academic skills. RQ was statistically significantly related to all academic skills and across all groups except for the SLD Math group and MCS (although it approached statistical significance). The next most consistent relation between cognitive and achievement skills was K0, indicating that general background knowledge plays an important part in predicting performance on reading and math skills in general, for children both with and without SLD. The relation between K0 and academic skills were equal for all academic areas and all groups, there were no statistically significant differences between groups in the influence of K0. This may be due to the multiple influences that are captured in a broad “knowledge” factor, which includes, but is not limited to opportunity to learn and the investment of general cognitive abilities via

motivation and academic interest. That is, knowledge may represent the accumulation of all of these influences, whereas a general factor, such as *g*, represents prior levels of cognitive ability, or as what Cattell (1987) would refer to as historical Gf.

It is important to note that several abilities which were statistically significant for individual groups in the backward selection were not statistically significant in the multi-group model. This is likely due to either sampling error (measurement or statistical) or a lack of power. But, the results from the multi-group model were weighted more heavily in this study. This was based on the assumption that if there was not a statistically significant difference between groups in the magnitude of a cognitive—achievement relation, then it suggests that the path is statistically equal across groups. The constrained paths were favored over unconstrained paths in this analysis, as long as there was not a statistically significant degradation in model fit when adding equality constraints to the structural relations. The lack of statistically significant effects, even for individual groups, could be due to a lack of power. Even though the samples sizes for all groups were adequate, the model was complex and many variables were included in the model, reducing overall power. Moreover, fewer indicators per factor is likely to have a substantial influence on power, and some of the effects that were found would be considered small or moderate in magnitude, even though they were not statistically significant. A decision was made to include the narrow abilities in this model. If the broad abilities were the focus, there may have been more statistically significant findings because they would have had more indicators per broad ability.

Lack of Ga effects

One way in which the current study differs from previous research on cognitive—achievement relations is that auditory processing was not related to any of the academic skills.

The deficit in phonological processing is one of the hallmarks of differences between children with and without SLD in reading (Hoskyn & Swanson, 2000; Johnson et al., 2010; Stanovich & Siegel, 1994). Looking at the sample statistics, it might be surprising that score for Sound Blending was not different between the SLD Reading and Norm groups. However, it is important to note that even if children with SLD in reading have a deficit in phonological skills (and they did not in this study), this does not necessarily indicate that there should be a relation between Ga and academic skills. For example, Swanson and Alexander (1997) found that *g* was the best predictor of pseudoword decoding for children with SLD in reading, while phonological awareness was the best predictor of pseudoword decoding for children with normally developing reading skills. This suggests that there are differences in which cognitive processes are important across groups, and some differences were found in this study. The lack of effects from Ga is consistent with previous research using the WJ-III for both basic reading (Floyd et al., 2007), math skills (Proctor, in press; Proctor et al., 2006; Taub et al., 2008), and writing skills (Floyd et al., 2008). When including several abilities in the model, any variance which would typically be accounted for by Ga if it were individually included in the model may be accounted for by other processes, such as background knowledge or memory. Floyd et al. found that Ga was a statistically significant predictor of basic reading skills, but only for adults. The findings from this study also show that the effects of Ga on basic reading skills are not statistically significant for individuals with SLDs either. This lack of finding may also reflect the trend in education. Phonics and phonemic awareness training have become a standard part of nearly all reading programs and curricula. Previous research may have reflected educational practice at the time, and an overall increase in phonemic awareness across the population of children and adolescents will affect whether tests like Sound Blending show up as deficits. It is possible that some

children who now have normative deficits compared to the current population will appear average when compared to the sample of children used as the normative sample for the WJ-III, which occurred between 1996 and 1999 (McGrew & Woodcock, 2001).

This presents an important issue regarding specification error, where important independent variables are not included in the analysis, even though they are essential to understanding the phenomenon of interest (Keith, 2006; Kline, 2011). Based on the current research, Ga did not play a statistically significant role in predicting academic skills for children and adolescents with and without SLD. This is not to say that phonological processing is not important for the development of some skills in reading and writing, but it may not be a primary process which predicts performance in reading when accounting for a variety of cognitive skills.

Implications for Research and Practice

Many of the findings from the current study support theoretical notions of SLD. First, the differences in cognitive—achievement relations suggest that individuals with SLD in particular academic areas may have some compensatory strategies for learning. For instance, the SLD Reading group appeared to rely more heavily on reasoning processes to comprehend text. If these individuals have difficulty decoding words but also have better reasoning skills, they may be able to come up with an adequate understanding what they are reading through the use of reasoning processes. This would explain why individuals with SLD in reading who have better reasoning processes also have better reading comprehension, and this relation was stronger for the SLD Reading group when compared to other groups. Similarly, individuals with SLD in math may rely more on their visualization skills and less on quantitative reasoning skills for math reasoning.

The findings from this study also indicate that the narrow abilities measured by the WJ-III are factorially invariant across groups of children with and without SLDs. Because they were invariant, it suggests that the tests used here may be good indicators of these skills, and the comparison of differences between children or adolescents with and without SLDs would be valid from a measurement perspective. Knowing that the measurement model in studies is valid is important, if accurate comparisons are going to be made. This is especially important for researchers in a variety of fields, including education, psychology, and neuroscience. Without adequate measurement models for adequate comparisons between groups, and invariance in those constructs, confidence in the accuracy of findings is diminished, and the possibility of error increases. Moreover, the use of latent variables, such as those used in this research, are more likely to result in findings that generalize across different samples of people and measurement instruments (e.g., Keith et al. 2008).

Although this study is primarily exploratory, there are several important implications for practice related to these results. When considering the identification of SLDs, the results from this study further support the importance of assessing specific cognitive skills as part of the evaluation process. The results show that *g* did not have any direct relations to the academic skills examined, and almost all relations between cognitive abilities and academic skills were for narrow cognitive abilities. This is not to say that *g* is unimportant (*g* did have the largest total effects across all cognitive abilities, although these effects were all mediated through more specific abilities), but these results indicate that more specific cognitive skills should also be considered during the assessment process, consistent with more recent models of SLD identification (e.g., Flanagan et al., 2010).

The group specific relations may provide some indication of what type of treatment might be useful for individuals with SLD in particular areas. However, practitioners should always keep the individual in mind when designing interventions. The population of individuals with SLD is heterogeneous (Fletcher et al., 2007), and intervention choice should be tailored to individual strengths and needs. The results of this study may suggest a better understanding of specific processes underlying SLD's, and a better understanding may lead to the development of particular interventions which may be useful for specific cognitive deficits. For example, visualization skills predicted math reasoning performance for the SLD Math group. If an individual has good visualization skills, providing interventions which can accommodate or supplement visualization as a method to improve math performance may be useful. Teaching specific strategies or providing guidance on how to increase performance in these possible compensatory areas may be useful for practitioners to know when making recommendations for intervention.

Gsm skills were related to several academic skills in this study. Recently, there has been an increase in research on the training and improvement of working memory and how it is related to other improvements in other areas of functioning. Working memory training has been related to improvements in working memory in children, (e.g., Holmes, Gathercole, & Dunning, 2009) and some recent research has identified a transfer of improvement to reading skills (Loosli et al., 2011), improvements in mathematics skills (Holmes et al., 2009), and impulsive decision making (Bickel, Yi, Landes, Hill, & Baxter, 2011). However, the research on transfer of working memory training is still very new and should be considered speculative until these effects have been replicated. The current study did not include a specific working memory factor, but working memory is a component of Gsm, which was related to BRS, MCS, and WE in this study. If

working memory training programs are useful in improving working memory along with other skills, interventions which focus on the development of working memory may be useful for children with SLD in improving academic skills. Theoretically, this should also be beneficial for children with SLDs because group membership did not moderate the relations between Gsm and academic skills in any part of this study. This assumption, however, would need to be formally examined in an empirical context.

For BRS, perceptual speed was related to decoding skills for the SLD Reading group. This suggests that individuals with faster information processing skills also have better decoding skills. Reading fluency requires both automatic and accurate processing of orthographic and phonological codes (Wolf & Katzir-Cohen, 2001). One intervention program designed by Wolf and colleagues focuses on the development of faster retrieval to help make the reading process more automatic (Retrieval, Automaticity, Vocabulary Elaboration, and Orthography [RAVE-O], Wolf, Miller, & Donnelly, 2000). Based on the double-deficit hypothesis (Wolf & Bowers, 1999), the RAVE-O program is designed to improve automatic retrieval of letter patterns in an effort to increase overall reading fluency. Although this program does not target processing speed directly, it does target speeded processes which are known to be related to reading (e.g., rapid naming). Other studies have shown improved processing speed through training (Mackey, Hill, Stone, & Bunge, 2011), and processing speed training programs have been successful in improving processing speed both in the short-term and long-term for older adults (e.g., Vance et al., 2007). Specific programs targeting the improvement of broad processing speed and how this is related to improvements in reading do not appear to have been examined extensively in children.

The major difference in reading comprehension between groups in this study was that individuals in the SLD Reading group relied more on reasoning skills to comprehend text. This

was specifically related to quantitative reasoning, and suggests that if individuals with difficulty in reading are able to improve their quantitative reasoning skills, this may help them improve their ability to comprehend text despite possible concurrent difficulties with decoding. Some research does suggest that improvements in reasoning processes are possible with training. For instance, Mackey et al. (2011) found that children increased in Gf by an average of 10 IQ points after an intervention where the children played both computer and non-computer based reasoning games. Whether these improvements would transfer to reading comprehension skills or other measures of fluid reasoning is unknown. Fletcher et al. (2007) point out that higher-level cognitive processes are important for reading comprehension, especially skills such as making inferences and metacognition. When comprehending written text, the reader must interpret what is read in a broader context of their own knowledge base. Basically, individuals must use their prior knowledge about the reading passage and make inferences about the meaning of the passage using reasoning abilities. The development of specific inference skills may be a helpful in activating reasoning and background knowledge to help better understand the meaning in text may be helpful for individuals who have trouble with decoding. For children with SLD in reading, these reasoning skills are, by definition, typically intact. Therefore, they do not need to be taught “reasoning” skills, but interventions would focus on how to use these reasoning skills in an efficient manner when reading so that they can work around the deficit related to word reading efficiency.

The results from this study indicated that quantitative reasoning did not predict math calculation skills for the SLD Math group, but Gsm and PS were statistically significant predictors. It is unknown if the deficits in the SLD Math group were due to retrieval difficulties or reasoning difficulties when considering the subtypes of math SLD outlined in Geary (2003).

However, Holmes et al. (2009) indicated that working memory training resulted in improved math performance six months after the training. Based on the findings from the current study, interventions which focus on the development of memory or processing speed skills may help improve math calculation performance for individuals with SLD in math.

For math reasoning, the main difference for the SLD Math group was that they relied less on quantitative reasoning skills and more on visualization skills. That is, better performance on tasks of math reasoning was related to better visualization skills. It is possible that direct training on visualization strategies may help these students improve their ability in completing applied problems. The ability to visualize how a problem could be solved may provide an advantage because it this may rely more on the visual aspect of working memory. One strategy which may help with solving applied math problems is the use of a graphic organizer, which may provide some assistance with visual processing. As an example, the use of graphic organizers has been effective in teaching adolescents with SLD to solve linear equations (Ives, 2007). By using a graphic organizer, this may take some load off of the visuospatial sketchpad, reducing overall cognitive load. However, Fletcher et al. (2007) note that little work has been completed examining specific interventions for improving math reasoning skills, and most of these interventions have focused on the implementation of specific strategies for solving math problems as opposed to changing cognitive skills (e.g., Fuchs, Fuchs, Hamlett, and Appleton, 2002). Another strategy may be to use a curriculum which focuses on a greater understanding practical math and practical problems. Continually providing intervention for memorizing math facts may not be helpful, but focusing on the practical applications of math may be more useful for individuals who have other means of solving calculation problems (e.g., teaching calculator use, number lines or number grids).

Finally, there were no differences between the SLD Writing group and other groups in the cognitive—achievement relations for Written Expression. It is interesting to note that there were a wide variety of skills related to written expression (Gc, RN, RQ, MS, and VZ), indicating that it is a complex task which relies on several different cognitive processes, regardless of SLD status. It is difficult to determine if there are any specific interventions which can be identified based on the results of this study. However, researchers such as Berninger and Colleagues (Berninger & O'Malley May, 2011) have provided extensive work on evidence-based diagnosis and intervention for SLD in written language. Additionally, Mather and Roberts (1995) provide a number of helpful interventions for written expression. They indicate that metacognition is especially important for written expression, and strategies which focus on the use of metacognitive strategies may be effective. For instance, the use of a graphic organizer may help individuals organize their thoughts and provide a concrete structure to the writing process. They also suggest that improving other skills, such as word retrieval strategies, can be helpful in improving written expression because it helps develop vocabulary and can help introduce a variety of words which can be used during the writing process.

Strengths of the Current Study

There are several important strengths to this study. First, this study used data from the WJ-III, which is a commonly used test of cognitive abilities designed from a CHC perspective. The use of the WJ-III also allowed for several narrow abilities to be included in the models, which has not been included often in previous research. This is unique in comparison to previous studies on cognitive—achievement relations.

The sample of children and adolescents with SLD were able to be differentiated by academic difficulty, rather than including all individuals in a single group. The results of this

study show that differentiating the groups by academic difficulty was important, because there were some specific differences in cognitive—achievement relations.

The methods used in the study also have several strengths. First, the use of CFA and SEM has not been employed often to examine the structure of cognitive ability tests or the relations between cognitive abilities and achievement in children with SLD. Only Bowden and colleagues (2008) have examined the measurement properties of a major cognitive ability test (WAIS-III) with individuals with SLDs. Besides this, little work has examined differences in cognitive abilities between children with and without SLDs using latent variable methods. The use of CFA is especially advantageous because latent variables represent error-free constructs, so relations among latent variables are based on reliable common variance. Moreover, this study obviated potential effects related to the selection of groups based on test scores (e.g., restriction of range) by utilizing covariance matrices and multi-group models, which allowed for a common metric to be used across groups. Last, this is one of the few studies which include both broad and narrow CHC abilities in the analysis. Floyd et al. (2007) included narrow abilities in their analysis of CHC abilities and reading, but other researchers have not included this level of analysis.

Methodologically, there were several ways in which the relations between cognitive abilities and academic skills could have been examined. The use of backward selection at the individual group level and the use of a multi-group model for examining similarities and differences between groups were selected for several reasons. The backward selection procedure in individual groups has been used previously by researchers (Elliot et al., 2010; Floyd et al., 2007; Keith et al., 2008) and helps reduce specification error, which is more likely with forward selection methods (Floyd et al., 2007). The purpose of using backward selection was to determine which relations were statistically significant from zero for individual groups. By

including all statistically significant relations in a multi-group model, equality constraints were added to each structural path to determine if the relation was equal across groups. This test is designed to determine if the cognitive—achievement relations identified in the backward selection were statistically significantly different between groups. This method has not been used in previous studies of cognitive-achievement relations for children with SLDs.

Limitations of the Current Study

There are several limitations to this study which need to be considered when interpreting the results. First, although the sample sizes for each of the SLD groups were adequate, a larger sample for each group would have been desirable. Larger sample sizes would have more power to detect smaller effects, and some relations between cognitive abilities and academic skills may have been overlooked due to a lack of power in this study. In addition to power, several variables which would have been desirable in the study could not be included due to missing data. One of the most important of these was Auditory Working Memory, which would have allowed for a specific working memory factor to be included. It is unknown whether the relations between Gsm and different academic skills were due to broad Gsm or because the single working memory subtest included in the analysis loaded directly on Gsm. Previous research indicates that working memory itself is important for a variety of academic skills (McGrew & Wendling, 2010), but it was not possible to verify in the current study due to missing data on this variable.

Next, the samples used in this study were not collected specifically for this investigation. The data for this study was obtained from a clinical database managed by the Woodcock-Muñoz Foundation. Although the database of children and adolescents with SLD was large and had a substantial number of individuals, this means that data was not collected based on any standard procedures other than the standardization of test administration. Different clinics may rely more

on specific groups of tests, and this resulted in a substantial portion of missing data for some individual tests in the SLD groups. Missing data is not an issue for maximum likelihood estimation if the data were collected in this fashion because this method of data collection does not directly violate the MAR assumption. However, it should be noted that planning for missing data is better practice because the specific mechanism regarding missing data is known (McArdle, 1994). In fact, it is probably a strength that not all participants completed all possible tests (37 tests total) included in the study because this would reduce any effects of fatigue on the part of the examinee. However, planning for missing data would be best practice in knowing exactly why data are missing from the sample.

One major limitation of the current analysis is the lack of age differentiation within groups. Each of the groups included children and adolescents between the ages of 6 and 19. It is well known that the relations between cognitive abilities and academic skills change over development (Floyd et al., 2007, 2008; McGrew & Wendling, 2010; Taub et al., 2008). The sample sizes for the SLD groups prevented the creation of separate groups based on age. Because groups could not be differentiated by age, some statistically significant relations may not have been identified, and not all relations should be interpreted invariantly across age groups. The findings here should be combined with other studies which have been able to differentiate groups based on age. Future research will need to examine the differential effects due to age more closely with larger samples of children and adolescents with SLD.

The model used in this study was very complex. One purpose of this study was to attempt to include narrow cognitive abilities in the factor model to predict performance on academic tasks, and this model was more inclusive of variables which could be used based on the data available. Most studies have only focused on broad abilities (e.g., Proctor, in press; Taub et al.,

2008), and few have specifically examine a wide range of broad abilities (Floyd et al., 2007). The inclusion of narrow ability factors added to the complexity of the model. As the number of variables are added to a model, the possible combinations and interactions increase geometrically, which can complicate the interpretation of what which cognitive skills are most importantly related to academic skills. By adding equality constraints to one cognitive—achievement relation at a time, this may change the magnitude of other relations in the model. However, this is one reason why the equality of structural paths were examined by adding equality constraints to a model where all structural paths were freely estimated across groups, and then removing equality constraints from a model where all structural paths were constrained to be equal across groups. The use of a multi-group model to examine similarities and differences in the relations between cognitive abilities and academic skills has not been used often (though see Keith, 1999), and methods on determining the best order of testing and inclusion of variables are not well-defined. The two methods of examining equality constraints did have some differences. For instance, in the MCS model the sensitivity analysis indicated that there should be statistically significant differences between groups on nearly all structural relations. However, this did not occur in the models where all structural paths were freely estimated and each was constrained individually. It was apparent that some differences found during the sensitivity analysis occurred because equality constraints on some structural paths (e.g., the path from RQ to MCS) may have been forcing differences between groups into other structural paths in the model. Using two different methods helps better understand these relations, hopefully making it more likely that differences between groups were adequately identified.

Another limitation of the current study is that only a single theoretical model was tested, the CHC model of cognitive abilities. While this provides the best interpretation of the WJ-III (McGrew & Woodcock, 2001), there are other possible models which could have been tested, such as a bifactor, or nested factors model where the effects of *g* are included directly on the subtests (e.g., Gustafsson & Balke, 1993), or a different model which groups subtests in a slightly different manner, such as the Verbal-Perceptual-Rotation model (Johnson & Bouchard, 2005a, 2005b). The bifactor model, however, would only focus on direct effects of all of the variables, including the *g* factor, so it would have a slightly different interpretation. However, these models were not included in the analysis because one of the purposes of this study was to examine the relations between more narrow cognitive abilities and academic skills specifically from a CHC perspective. CHC theory provides a theoretically and empirically consistent taxonomy of cognitive abilities (Schneider & McGrew, 2012), although other interpretations of the structure of the cognitive tests used in this study may be viable.

Conclusions and Future Directions

The results of this study show that there are some important differences in the cognitive abilities used by children with and without SLDs when completing academic tasks. The results from this study are important for several reasons. First, the results verify that the CHC constructs measured by the WJ-III are essentially equal across groups of children with and without SLD. Next, the results of this study show that individuals with SLD in different academic rely on different cognitive abilities when engaging in academic tasks. It is possible that these differences are compensation strategies for weaknesses in other cognitive processes and difficulties in academic performance. There were several interesting findings from this study, but there are also many unanswered questions which can help guide future research.

First, it would be very beneficial to examine the cognitive—achievement relations from a developmental perspective. With longitudinal studies, it is possible to better understand how development or deficits in one area affect the development or deficits in other areas over time. For example, considering the smaller relationships between RQ and math achievement for those in the SLD Math group, do reasoning processes start out more highly related to math skills and then become less important, or is this smaller relation between RQ and math skills pervasive across development? Similarly, does RQ affect growth in reading comprehension for children with SLD in reading more than background knowledge, and when does this change begin to take place? Longitudinal studies could help provide a number of answers regarding the nature of how the relations between cognitive abilities and academic skills change over time, and this could provide important guidance for intervention and prevention. The results of the current study may help inform which variables might need to be included in future research. Research has shown that SLDs are not developmental lags in cognitive processing or academic skill development (O'Shaughnessy & Swanson, 1998; Stanovich & Siegel, 1994). Rather, they are better viewed as deficits in cognitive processing which affects academic skill development (Flanagan et al., 2010). Long-term strategies, guidance, and interventions that are developmentally informed will thus have much greater influence on a person's life, however, such interventions and guidance are rarely considered.

This is one of the first studies to closely examine cognitive—achievement relations for groups of children and adolescents with different forms of SLD. It would be beneficial for future research to attempt to differentiate these groups even more. For instance, if there are different subtypes of SLD in mathematics (Geary, 2003), then there may be specific cognitive differences and cognitive—achievement relations for these different subtypes. Individuals with difficulty in

math calculation skills may rely on a very different set of skills to complete applied math problems than individuals with difficulty in mathematics problem solving. Determining if these subtypes differ in their cognitive skills and intervention needs would be beneficial to providing services which best fit the individual. A better understanding of SLD in general will require even more fine-grained analysis.

Finally, future research would also need to verify these findings with different tests and a different sample. The current study used an archived sample, but a well-planned study which employs a variety of measures from a variety of cognitive and achievement batteries would provide important external validity to the findings from this study.

References

Abbott, R. D., & Berninger, V. W. (1993). Structural equation modeling of relationships among developmental skills and writing skills in primary- and intermediate-grade writers. *Journal of Educational Psychology, 85,* 478-508.

Akaike, H. A. (1987). Factor analysis and AIC. *Psychometrika, 52,* 317-332.

Alfonso, V. C., Flanagan, D. P., & Radwan, S. (2005). The impact of the Cattell-Horn-Carroll theory on test development and interpretation of cognitive and academic abilities. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (2nd ed., pp. 185-202). New York: Guilford Press.

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text revision). Washington, DC: Author.

Baddeley, A. D., & Hitch, G. H. (1994). Developments in the concept of working memory. *Neuropsychology, 8,* 485-493.

Baraldi, A. N., & Enders, C. K. (2010). An introduction to modern missing data analyses. *Journal of School Psychology, 48,* 5-37. doi: 10.1016/j.jsp.2009.10.001

Bell, S. M., McCallum, R. S., & Cox, E. A. (2003). Toward a research-based assessment of dyslexia: Using cognitive measures to identify reading disabilities. *Journal of Learning Disabilities, 36,* 505-516.

Benson, N. (2007). Cattell—Horn—Carroll cognitive abilities and reading achievement. *Journal of Psychoeducational Assessment, 26,* 27-41. doi: 10.1177/0734282907301424

Benson, N., Hulac, D. M., & Kranzler, J. H. (2010). Independent examination of the Wechsler Adult Intelligence Scale-Fourth Edition (WAIS-IV): What does the WAIS-IV measure? *Psychological Assessment, 22,* 121-130. doi: 10.1037/a0017767

Bentler, P. M. (1990). Comparative fit indexes in structural equation modeling. *Psychological Bulletin, 107,* 238-246.

Berninger, V. W., Abbott, R. D., Thomson, J. B., & Raskind, W. H. (2001) Language phenotype for reading and writing disability: A family approach. *Scientific Studies of Reading, 5,* 59-106.

Berninger, V. W., & Amtmann, D. (2003). Preventing written expression disabilities through early and continuing assessment and intervention for handwriting and/or spelling problems: Research into practice. In H. L. Swanson, K. R. Harris, & S. Graham (Eds.), *Handbook of learning disabilities* (pp. 140-157). New York: Guilford Press.

Berninger, V. W., Nielsen, K. H., Abbott, R. D., Wijsman, E., & Raskind, W. (2008). Gender differences in severity of writing and reading disabilities. *Journal of School Psychology, 46,* 151-172. doi:10.1016/j.jsp.2007.02.007

Berninger, V. W., & O'Malley May, M. (2011). Evidence-based diagnosis and treatment for specific learning disabilities involving impairments in written and/or oral language. *Journal of Learning Disabilities, 44,* 167-183.

Bickel, W. K., Yi, R., Landes, R. D., & Baxter, C. (2011). Remember the future: Working memory training decreases delay discounting among stimulant addicts. *Biological Psychiatry, 69,* 260-265.

Bowden, S. C., Cook, M. J., Bardenhagen, F. J., Shores, E. A., & Carstairs, J. R. (2004). Measurement invariance of core cognitive abilities in heterogeneous neurological and community samples. *Intelligence, 32,* 363-389. doi: 10.1016/j.intell.2004.05.002

Bowden, S. C., Gregg, N., Bandalos, D., Davis, M., Coleman, C., Holdnack, J. A., & Weiss, L. G. (2008). Latent mean and covariance differences with measurement equivalence in

college students with developmental difficulties versus the Wechsler Adult Intelligence Scale-III/Wechsler Memory Scale-III normative sample. *Educational and Psychological Measurement, 68,* 621-642. doi: 10.1177/0013164407310126

Bowden, S. C., Ritter, A. J., Carstairs, J. R., Shores, E. A., Pead, J., Greeley, J. D., . . . Clifford, C. C. (2001). Factorial invariance for combined Wechsler Adult Intelligence Scale-Revised and Wechsler memory Scale-Revised scores in a sample of clients with alcohol dependency. *The Clinical Neuropsychologist, 15,* 69-80.

Bowers, P. G., & Ishaik, G. (2003). RAN's contribution to understanding reading disabilities. In H. L. Swanson, K. R. Harris, & S. Graham (Eds.), *Handbook of learning disabilities* (pp. 140-157). New York: Guilford Press.

Bowers, P. G., & Wolf, M. (1993). Theoretical links between naming speed, precise timing mechanisms and orthographic skills in dyslexia. *Reading and Writing: An Interdisciplinary Journal, 5,* 69-85.

Bradley, R., Danielson, L., & Hallahan, D. P. (Eds.). (2002). *Identification of learning disabilities: Research to practice*. Mahwah, NJ: Lawrence Erlbaum Associates.

Brown, T. A. (2006). *Confirmatory factor analysis for applied research*. New York: Guilford Press.

Brown-Chidsey, R., & Steege, M. W. (2010). *Response to intervention: Principles and strategies for effective practice* (2nd ed.). New York: Guilford Press.

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor analytic studies*. New York: Cambridge University Press.

Carroll, J. B. (2005). The three-stratum theory of cognitive abilities. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (2nd ed.) (pp. 69-76). New York: Guilford Press.

Cattell, R. B. (1943). The measurement of adult intelligence. *Psychological Bulletin, 40,* 153-193.

Cattell, R. B. (1963). Theory of fluid and crystallized intelligence: A critical experiment. *Journal of Educational Psychology, 54,* 1-22.

Cattell, R. B. (1987). *Intelligence: Its structure, growth, and action*. New York: Elsevier.

Chen, F. F., Sousa, K. H., & West, S. G. (2005). Testing measurement invariance of second-order factor models. *Structural Equation Modeling, 12,* 471-492.

Chen, H., Keith, T. Z., Weiss, L. G., Zhu, J., & Li, Y. (2010). Testing for multigroup invariance of second-order WISC-IV structure across China, Hong Kong, Macau, and Taiwan. *Personality and Individual Differences, 49,* 677-682. doi: 10.1016/j.paid.2010.06.004

Chen, H., & Zhu, J. (2008). Factor invariance between genders of the Wechsler Intelligence Scale for Children – Fourth edition. *Personality and Individual Differences, 45,* 260-266. doi: 10.1016/j.paid.2008.04.008

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling, 9,* 233-255.

Ciancilo, A. T., & Sternberg, R. J. (2004). *Intelligence: A brief history*. Malden, MA: Blackwell Publishing.

Ciszek, G. J. (2003). Review of the Woodcock-Johnson III. In B. S. Plake & J. C. Impara (Eds.), *The fifteenth mental measurements yearbook* (pp. 1020-1024). Lincoln, NE: Buros Institute of Mental Measurements.

Cohen, J. (1983). The cost of dichotomization. *Applied Psychological Measurement, 7,* 249-253.

Compton, D. L., Fuchs, L. S., Fuchs, D., Lambert, W., & Hamlett, C. (in press). The cognitive and academic profiles of reading and mathematics learning disabilities. *Journal of Learning Disabilities*. doi: 10.1177/0022219410393012

Curran, P. J., West, S. G., & Finch, J. F. (1996). The robustness of test statistics to nonnormality and specification error in confirmatory factor analysis. *Psychological Methods, 1,* 16-29.

Edwards, O. W., & Oakland, T. D. (2006). Factorial invariance of Woodcock-Johnson III scores for African Americans and Caucasian Americans. *Journal of Psychoeducational Assessment, 24,* 358-366. doi: 10.1177/0734282906289595

Elliot, C. D. (2007). *Differential ability scales* (2nd ed.): *Introductory and technical manual*. San Antonio, TX: Psychological Corporation.

Elliot, C. D., Hale, J. B., Fiorello, C. A., Dorvil, C., & Moldovan, J. (2010). Differential Ability Scales-II prediction of reading performance: Global scores are not enough. *Psychology in the Schools, 47,* 698-720. doi: 10.1002/pits.20499

Enders, C. K. (2010). *Applied missing data analysis.* New York: Guilford Press.

Evans, J. J., Floyd, R. G., McGrew, K. S., & Leforgee, M. H. (2001). The relations between measures of Cattell-Horn-Carroll (CHC) cognitive abilities and reading achievement during childhood and adolescence. *School Psychology Review, 31,* 246-262.

Fiorello, C. A., Hale, J. B., & Snyder, L. E. (2006). Cognitive hypothesis testing and response to intervention for children with reading problems. *Psychology in the Schools, 43,* 835-853. doi: 10.1002/pits.20192

Flanagan, D. P., Fiorello, C. A., & Ortiz, S. O. (2010). Enhancing practice through application of Cattell-Horn-Carroll theory and research: A “third method” approach to specific learning disability identification. *Psychology in the Schools*, 739-760. doi: 10.1002/pits.20501

Flanagan, D. P., & Mascolo, J. T. (2005). Psychoeducational assessment and learning disability diagnosis. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (pp. 521-544). New York: Guilford Press.

Flanagan, D. P., & McGrew, K. S. (1998). Interpreting intelligence tests from contemporary Gf-Gc theory: Joint confirmatory factor analysis of the WJ-R and KAIT in a non-white sample. *Journal of School Psychology, 36,* 151-182.

Flanagan, D. P., Ortiz, S. O., Alfonso, V. C., & Dynda, A. M. (2006). Integration of response to intervention and norm-referenced tests in learning disability identification: Learning from the Tower of Babel. *Psychology in the Schools, 43,* 807-825. doi: 10.1002/pits.20190

Fletcher, J. M., Lyon, G. R., Fuchs, L. S., & Barnes, M. A. (2007). *Learning disabilities: From identification to intervention*. New York: Guilford Press.

Fletcher, J. M., Shaywitz, S. E., Shankweiler, D. P., Katz, L., Liberman, I. Y., Stuebing, K. K.,…Shaywitz, B. A. (1994). Cognitive profiles of reading disability: Comparisons of discrepancy and low achievement definitions. *Journal of Educational Psychology, 86,* 6-23.

Floyd, R. G., Bergeron, R., & Alfonso, V. C. (2006). Cattell-Horn-Carroll cognitive abilities profiles of poor comprehenders. *Reading and Writing, 19,* 427-456.

Floyd, R. G., Keith, T. Z., Taub, G. E., & McGrew, K. S. (2007). Cattell-Horn-Carroll cognitive abilities and their effects on reading decoding skills: g has indirect effects, more specific

abilities have direct effects. *School Psychology Quarterly, 22,* 200-233. doi: 10.1037/1045-3830.22.2.200

Floyd, R. G., McGrew, K. S., Barry, A., Rafael, F., & Rogers, J. (2009). General and specific effects on Cattell-Horn-Carroll broad ability composites: Analysis of the Woodcock-Johnson-III Normative Update Cattell-Horn-Carroll factor clusters across development. *School Psychology Review, 38,* 249-265.

Floyd, R. G., McGrew, K. S., & Evans, J. J. (2008). The relative contributions of the Cattell-Horn-Carroll cognitive abilities in explaining writing achievement during childhood and adolescence. *Psychology in the Schools, 45,* 132-144. doi: 10.1002/pits.20284

French, B. F., & Finch, W. H. (2006). Confirmatory factor analytic procedures for the determination of measurement invariance. *Structural Equation Modeling, 13,* 378-402.

Fuchs, L. S., Fuchs, D., Compton, D. L., Powell, S. R., Seethaler, P. M., Capizzi, A. M.,...Fletcher, J. M. (2006). The cognitive correlates of third-grade skill in arithmetic, algorithmic computation, and arithmetic word problems. *Journal of Educational Psychology, 98,* 29-43. doi: 10.1037/0022-0663.98.1.29

Fuchs, L. S., Fuchs, D., Hamlett, C. L., & Appleton, A. C. (2002). Explicitly teaching for transfer: Effects on the mathematical problem-solving performance of students with mathematics disabilities. *Learning Disabilities Research and Practice, 17,* 90-106.

Geary, D. C., (1993). Mathematical disabilities: Cognitive, neuropsychological, and genetic components. *Psychological Bulletin, 114,* 345-362.

Geary, D. C. (2003). Learning disabilities in arithmetic: Problem-solving differences and cognitive deficits. In H. L. Swanson, K. R. Harris, & S. Graham (Eds.), *Handbook of learning disabilities* (pp. 199-212). New York: Guilford Press.

Geary, D. C. (2011, September 26). Cognitive predictors of achievement growth in mathematics: A 5-year longitudinal study. *Developmental Psychology.* Advance online publication. doi: 10.1037/a0025510

Geary, D. C., Widaman, K. F., Little, T. D., & Cormier, P. (1987). Cognitive addition: Comparison of learning disabled and academically normal elementary school children. *Cognitive Development*, 2, 149-169.

Glutting, J., Adams, W., & Shelslow, D. (2002). *Wide Range Intelligence Test*. Wilmington, DE: Wide Range.

Graham, J. W., Hofer, S. M., Donaldson, S. I., MacKinnon, D. P., & Schafer, J. L. (1997). Analysis with missing data in prevention research. In K. J. Bryant, M. Windle & S. G. West (Eds.), *The science of prevention: Methodological advances from alcohol and substance abuse research* (pp. 325-366). Washington, DC: American Psychological Association.

Gregg, N. (2009). *Adolescents and adults with learning disabilities and ADHD: Assessment and accommodation*. New York: Guilford Press.

Gustafsson, J. E. (1984). A unifying model for the structure of intellectual abilities. *Intelligence, 8*, 179-203.

Gustafsson, J. E., & Balke, G. (1993). General and specific abilities as predictors of school achievement. *Multivariate Behavioral Research, 28*, 407-434.

Hale, J. B., Kaufman, A., Naglieri, J. A., & Kavale, K. A. (2006). Implementation of IDEA: Integrating response to intervention and cognitive assessment methods. *Psychology in the Schools, 43*, 753-770. doi: 10.1002/pits.20186

Hallahan, D. P., Keller, C. E., Martinez, E. A., Byrd, E. S., Gelman, J. A., & Fan, X. (2007). How variable are interstate prevalence rates of learning disabilities and other special education categories? A longitudinal comparison. *Exceptional Children, 73,* 136-146.

Hallahan, D. P., & Mercer, C. D. (2002). Learning disabilities: Historical perspectives. In R. Bradley, L. Danielson, & D. P. Hallahan (Eds.). *Identification of learning disabilities: Research to practice* (pp. 1-67). Mahwah, NJ: Lawrence Earlbaum.

Hallahan, D. P., & Mock, D. R. (2003). A brief history of the field of learning disabilities. In H. L. Swanson, K. R. Harris, & S. Graham (Eds.). *Handbook of learning disabilities* (pp. 16-29). New York: Guilford Press.

Henne, M., Hilbert, S., Drazler, C., Zeigler, M., & Bühner, M. (2011). Masking misfit in confirmatory factor analysis by increasing unique variances: A cautionary not on the usefulness of cutoff values of fit indices. *Psychological Methods, 16,* 319-336. doi:10.1037/a0024917.

Holmes, J., Gathercole, S. E., & Dunning, D. L. (2009). Adaptive training leads to sustained enhancement of poor working memory in children. *Developmental Science,* DOI: 10.1111/j.1467-7687.2009.00848.x

Horn, J. L., & Blankson, N. (2005). Foundations for better understanding of cognitive abilities. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (2nd ed., pp. 41-68). New York: Guilford Press.

Horn, J. L., & Cattell, R. B. (1966). Refinement and test of the theory of fluid and crystallized general intelligence. *Journal of Educational Psychology, 57,* 253-270.

Horn, J. L., & McArdle, J. J. (1992). A practical and theoretical guide to measurement invariance in aging research. *Experimental Aging Research, 18,* 117-144.

Hoskyn, M., & Swanson, H. L. (2000). Cognitive processing of low achievers and children with reading disabilities: A selective meta-analytic review of the published literature. *School Psychology Review, 29,* 102-119.

Hu., L., & Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological Methods, 3,* 424-453.

Hu. L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal, 6,* 1-55. doi: 10.1080/10705519909540118

Ives, B. (2007). Graphic organizers applied to secondary algebra instruction for students with learning disorders. *Learning Disabilities Research and Practice, 22,* 110-118.

Jackson, D. L., Gillaspy, J. A., & Purc-Stephenson, R. (2009). Reporting practices in confirmatory factor analysis: An overview and some recommendations. *Psychological Methods, 14,* 6-23. doi:10.1037/a0014694

Jensen, A. R. (1998). *The* g *factor: The science of mental ability*. Westport, CT: Praeger.

Johnson, E. S., Humphery, M., Mellard, D. F., Woods, K., & Swanson, H. L. (2010). Cognitive processing deficits and students with specific learning disabilities: A selective meta-analysis of the literature. *Learning Disability Quarterly, 33,* 3-18.

Johnson, W., & Bouchard, T. J. (2005a). The structure of human intelligence: It is verbal, perceptual, and image rotation (VPR), not fluid and crystallized. *Intelligence, 33,* 393-416.

Johnson, W., & Bouchard, T. J. (2005b). Constructive replication of the visual-perceptual-image rotation model in Thurstone's (1941) battery of 60 tests of mental ability. *Intelligence, 33,* 417-430.

Jöreskog, K. G. (1971). Simultaneous factor analysis in several populations. *Psychometrika, 36,* 409-426.

Kamphaus, R. W., Winsor, A. P., Rowe, E. W., & Kim, S. (2005). A history of intelligence test interpretation. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (2nd ed., pp. 23-38). New York: Guilford Press.

Kan, K. J., Keivit, R., Dolan, C., & van der Maas, H. (in press). On the interpretation of the CHC factor Gc. *Intelligence*. doi:10.1016/j.intell.2011.05.003

Kaufman, A. S., & Kaufman, N. L. (2004). *Kaufman Assessment Battery for Children* (2nd ed.): *Technical manual*. Pines, MN: American Guidance Service.

Kaufman, S. B., Reynolds, M. R., Liu, X., Kaufman, A. S., & McGrew, K. S. Are cognitive *g* and academic achievement *g* one and the same *g*? An exploration on the Woodcock-Johnson and Kaufman tests. *Intelligence* (2012), doi:10.1016/j.intell.2012.01.009

Kavale, K. A. (2002). Discrepancy models in the identification of learning disability. In R. Bradley, L. Danielson, & D. P. Hallahan (Eds.), *Identification of learning disabilities: Research to practice* (pp. 369-426). Mahwah, NJ: Lawrence Earlbaum.

Kavale, K. A., & Forness, S. R. (1984). A meta-analysis of the validity of Wechsler scale profiles and recategorizations: Patterns or parodies? *Learning Disability Quarterly, 7,* 136-156.

Keith, T. Z. (1999). Effects of general and specific abilities on student achievement: Similarities and differences across ethnic groups. *School Psychology Quarterly, 14,* 239-262.

Keith, T. Z. (2006). *Multiple regression and beyond*. Boston: Pearson.

Keith, T. Z., Fine, J. G., Taub, G. E., Reynolds, M. R., & Kranzler, J. H. (2006). Higher order, multisample, confirmatory factor analysis of the Wechsler Intelligence Scale for Children - Fourth Edition: What does it measure? *School Psychology Review, 35,* 108-127.

Keith, T. Z., Kranzler, J. H., & Flanagan, D. P. (2001). What does the Cognitive Assessment System measure? A joint confirmatory factor analysis of the CAS and Woodcock-Johnson Tests of Cognitive Ability (3rd Edition). *School Psychology Review, 30,* 89-119.

Keith, T. Z., Low, J. A., Reynolds, M. R., Patel, P. G., & Ridley, K. P. (2010). Higher-order factor structure of the Differential Ability Scales-II: Consistency across ages 4 to 17. *Psychology in the Schools, 47,* 676-697. doi: 10.1002/pits.20498

Keith, T. Z., & Reynolds, M. R. (2010). Cattell-Horn-Carroll abilities and cognitive tests: What we've learned from 20 years of research. *Psychology in the Schools, 47,* 635-650. doi: 10.1002/pits 20496

Keith, T. Z., Reynolds, M. R., Patel, P. G., & Ridley, K. P. (2008). Sex differences in latent cognitive abilities ages 6 to 59: Evidence from the Woodcock–Johnson III Tests of Cognitive Abilities. *Intelligence, 36,* 502-525. doi: 10.1016/j.intell.2007.11.001

Kenny, D. A., & McCoach, D. B. (2003). Effect of the number of variables on measures of fit in structural equation modeling. *Structural Equation Modeling: A Multidisciplinary Journal, 10,* 333-351. doi: 10.1207/S15328007SEM1003_1

Kline, R. B. (2011). *Principles and practice of structural equation modeling* (3rd ed.). New York: Guilford Press.

Klingberg, T. (2010). Training and plasticity in working memory. *Trends in Cognitive Science, 14,* 317-324.

Krane, W. R., & Slaney, K. L. (2005). A general introduction to the common factor model. In A. Mayadeu-Olivares & J. J. McArdle (Eds.), *Contemporary psychometrics: A festschrift for Roderick P. McDonald* (pp. 125-151). Mahwah, NJ: Lawrence Erlbaum Associates.

Kranzler, J. H., Keith, T. Z., & Flanagan, D. P. (2000). Independent examination of the factor structure of the cognitive assessment system (CAS): Further evidence challenging the construct validity of the CAS. *Journal of Psychoeducational Assessment, 18,* 143-159.

Little, R. J. A. (1988). A test of missing completely at random for multivariate data with missing values. *Journal of the American Statistical Association, 83,* 1198–1202.

Little, T. D. (1997). Mean and covariance structures (MACS) analyses of cross-cultural data: Practical and theoretical issues. *Multivariate Behavioral Research, 32,* 53-76.

Locke, S., McGrew, K. S., & Ford. L. (2011). A multiple group confirmatory factor analysis of the structural invariance of the Cattell-Horn-Carroll theory of cognitive abilities across matched Canadian and U.S. samples. *WMF Press Bulletin, No. 1*. Retrieved from the Woodcock-Munoz Foundation website: http://woodcock-munoz-foundation.org/press/pressbulletins.html

Loehlin, J. C. (1998). *Latent variable models: An introduction to factor, path, and structural analysis.* Mahwah, NJ: Lawrence Earlbaum.

Loosli, S. V., Buschkuel, M., Perrig, W. J., & Jaeggi, S. M. (2012). Working memory training improves reading processes in typically developing children. *Child Neuropsychology, 18,* 62-78. doi: 10.1080/09297049.2011.575772

Lyon, G. R., Fletcher, J. M., & Barnes, M. C. (2003). Learning disabilities. In E. J. Mash & R. A. Barkley (Eds.), *Child psychopathology* (pp. 520–586). , 2nd ed. New York: Guilford Press.

Mackey, A. P., Hill, S. S., Stone, S. I., & Bunge, S. A. (2011). Differential effects of reasoning and speed training in children. *Developmental Science, 14,* 582-590. doi: 10.1111/j.1467-7687.2010.01005.x

Mather, N., & Kaufman, N. (2006a). Special issue, part one: Integration of cognitive assessment and response to intervention [Special Issue]. *Psychology in the Schools, 43*, 745-825.

Mather, N., & Kaufman, N. (2006b). Special issue, part one: Integration of cognitive assessment and response to intervention [Special issue]. *Psychology in the Schools, 43*, 827-908.

Mather, N, & Roberts, R. (1995). *Informal assessment and instruction in written language: A practitioner's guide for students with learning disabilities.* New York: Wiley.

Mayes, S. D., Calhoun, S. L., & Crowell, E. W. (2000). Learning disabilities and ADHD: Overlapping spectrum disorders. Journal of Learning Disabilities, 33, 417-424. doi: 10.1177/002221940003300502

McArdle, J. J. (1994). Structural factor analysis experiments with incomplete data. *Multivariate Behavioral Research, 29,* 409-454.

McDermott, P. A., Goldberg, M. M., Watkins, M. W., Stanley, J. L., & Glutting, J. J. (2006). A nationwide epidemiologic modeling study of LD: Risk, protection, and unintended impact. *Journal of Learning Disabilities, 29,* 230-251.

McGrew, K. S. (2005). The Cattell-Horn-Carroll theory of cognitive abilities: Past, present, and future. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (2nd ed., pp. 136-181). New York: Guilford Press.

McGrew, K. S. (2009). CHC theory and the human cognitive abilities project: Standing on the shoulders of the giants of psychometric intelligence research. *Intelligence, 37,* 1-10.

McGrew, K. S., & Flanagan, D. P. (1998). *The intelligence test desk reference (ITDR): Gf-Gc cross battery assessment*. Boston: Allyn & Bacon.

McGrew, K. S., & Knopik, S. N. (1993). The relationship between the WJ-R Gf-Gc cognitive clusters and writing achievement across the life-span. *School Psychology Review, 22,* 687-695.

McGrew, K. S., Schrank, F. A., & Woodcock, D. W. (2007). *Technical manual. Woodcock-Johnson III Normative Update.* Rolling Meadows, IL: Riverside Publishing.

McGrew, K. S., & Wendling, B. J. (2010). Cattell-Horn-Carroll cognitive-achievement relations: What we have learned from the past 20 years of research. *Psychology in the Schools, 47,* 651-675. doi: 10.1002/pits.20497

McGrew, K. S., & Woodcock, R. W. (2001). *Technical manual. Woodcock-Johnson III.* Itasca, IL: Riverside Publishing.

Meade, A. W., Johnson, E. C., & Braddy, P. W. (2008). Power and sensitivity of alternative fit indices in tests of measurement invariance. *Journal of Applied Psychology, 93,* 568-592. doi: 10.1037/0021-9010.93.3.568

Meredith, W. (1993). Measurement invariance, factor analysis, and factorial invariance. *Psychometrika, 58,* 525-543.

Meredith, W., & Horn, J. L. (2001). The role of factorial invariance in modeling growth and change. In L. M. Collins & Sayer, A. G. (Eds.), *New methods for the analysis of change* (pp. 203-240). Washington, DC: American Psychological Association.

Meredith, W., & Teresi, J. A. (2006). An essay on measurement and factorial invariance. *Medical Care, 44,* S69-S77.

Millsap, R. E. (2011). *Statistical approaches to measurement invariance*. New York: Routledge.

Millsap, R. E., & Meredith, W. (2007). Factorial invariance: Historical perspectives and new problems. In R. Cudeck & R. C. MacCallum (Eds.), *Factor analysis at 100: Historical*

developments and future directions (pp. 131-152). Mahwah, NJ: Lawrence Erlbaum Associates.

Mulaik, S. A. (2010). *Foundations of factor analysis* (2nd ed.). Boca Raton, FL: Chapman & Hall/CRC.

Muthén, B. O. (1989). Latent variable modeling in heterogeneous populations. *Psychometrika, 54,* 557-585.

Nation, K., & Snowling, M. J. (1997). Assessing reading difficulties: the validity and utility of current measures of reading skill. *British Journal of Educational Psychology, 67,* 359-370.

National Joint Committee on Learning Disabilities. Anonymous. (1991). Learning disabilities: Issues on definition. *Asha, 33,* (Suppl. 5), 18-20.

Niileksela, C. R., & Reynolds, M. R. (manuscript submitted for publication). Global, broad, or specific cognitive differences? Using a MIMIC model to examine difference in CHC abilities in children with learning disabilities.

Osterlind, S. J. (2010). *Modern measurement. Theory, principles, and applications of mental appraisal* (2nd ed.). Boston: Pearson.

O'Shaughnessy, T. E., & Swanson, H. L. (1998). Do immediate memory deficits in students with learning disabilities in reading reflect a developmental lag or deficit?: A selective meta-analysis of the literature. *Learning Disability Quarterly, 21,* 123-148.

Payette, K. A., & Clarizio, H. F. (1994). Discrepant team decisions: The effects of race, gender, achievement, and IQ on LD eligibility. *Psychology in the Schools, 31,* 40–48.

Pearson, K. (1903). Mathematical contributions to the theory of evolution. XI. On the influence of natural selection on the variability and correlation of organs. *Philosophical*

Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character, 200, 1-66.

Phelps, L., McGrew, K. S., Knopik, S. N., & Ford, L. (2005). The general (*g*), broad, and narrow CHC stratum characteristics of the WJ-III and WISC-IV tests: A confirmatory cross-battery investigation. *School Psychology Quarterly, 20,* 66-88.

Preacher, K . J., & MacCallum, R. C. (2003). Repairing Tom Swift's electric factor analysis machine. *Understanding Statistics, 2,* 13-43.

Proctor, B. (in press). Relationships between Cattell-Horn-Carroll (CHC) cognitive abilities and math achievement within a sample of college students with learning disabilities. *Journal of Learning Disabilities,* 1-10. doi: 10.1177/0022219410392049

Proctor, B. E., Floyd, R. G., & Shaver, R. B. (2005). Cattell-Horn-Carroll broad cognitive ability profiles of low math achievers. *Psychology in the Schools, 42,* 1-12. doi: 10.1001/pits.20030

Reschly, D. J. (2005). Learning disabilities, identification: Primary intervention, secondary intervention, and then what? *Journal of Learning Disabilities, 38,* 510-515. doi: 10.1177/00222194050380060601

Reynolds, C. R., & Kamphaus, R. W. (2003). *Reynolds Intellectual Assessment Scales*. Lutz, FL: Psychological Assessment Resources.

Reynolds, C. R., & Shaywitz, S. E. (2009). Response to intervention: Ready or not? Or, from wait-to-fail to watch-them-fail. *School Psychology Quarterly, 24,* 130-145.

Reynolds, M. R., & Keith, T. Z. (2007). Spearman's law of diminishing returns in hierarchical models of intelligence for children and adolescents. *Intelligence, 35,* 267-281. doi: 10.1016/j.intell.2006.08.002

Reynolds, M. R., Keith, T. Z., Fine, J. G., Fisher, M. E., & Low, J. A. (2007). Confirmatory factor structure of the Kaufman Assessment Battery for Children-Second Edition: Consistency with Cattell-Horn-Carroll theory. *School Psychology Quarterly, 22,* 511-539. doi: 10.1037/1045-3830.22.4.511

Reynolds, M. R., Keith, T. Z., Ridley, K., & Patel, P. (2008). Sex differences in latent general and broad cognitive abilities for children and youth: Evidence from higher-order MG-MACS and MIMIC models. *Intelligence, 36,* 236-260. doi: 10.1016/j.intell.2007.06.003

Reynolds, M. R., & Turek, J. J. (in press). Developmental dynamics between Gc and reading comprehension: Gc accelerates positive change in reading comprehension.

Roid, G. H. (2003). *Stanford-Binet Intelligence Scales* (5th ed.): *Technical manual*. Itasca, IL: Riverside.

Sanders, S., McIntosh, D. E., Dunham, M., Rothlisberg, B. A., & Finch, H. (2007). Joint confirmatory factor analysis of the Differential Ability Scales and the Woodcock-Johnson Tests of Cognitive Abilities–Third Edition. *Psychology in the Schools, 44,* 119-138. doi: 10.1002/pits.20211

Sandoval, J. (2003). Review of the Woodcock-Johnson III. In B. S. Plake & J. C. Impara (Eds.), *The fifteenth mental measurements yearbook* (pp 1024-1028). Lincoln, NE: Buros Institute of Mental Measurements.

Schatschneider, C., Fletcher, J. M., Francis, D. J., Carlson, C. D., & Foorman, B. R. (2004). Kindergarten prediction of reading skills: A longitudinal comparative analysis. *Journal of Educational Psychology, 96,* 265-282. doi: 10.1037/0022-0663.96.2.265

Schermelleh-Engel, K., Moosbrugger, H., & Müller, H. (2003). Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research Online, 8,* 23-74.

Schneider, W. J., & McGrew, K. S. (2012). The Cattell-Horn-Carroll model of intelligence. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (3rd ed., pp. 99-144). New York: Guilford Press.

Schreiber, J. B., Stage, F. K., King, J., Nora, A., & Barlow, E. A. (2006). Reporting structural equation modeling and confirmatory factor analysis results: A review. *Journal of Educational Research, 99,* 323-337.

Schwarz, G. (1978). Estimating the dimensions of a model. *The Annals of Statistics, 6,* 461-464.

Sörbom, D. (1974). A general method for studying differences in factor means and factor structure between groups. *British Journal of Mathematical and Statistical Psychology, 27,* 229-239.

Spearman, C. (1904). "General intelligence," objectively determined and measured. *The American Journal of Psychology, 15,* 201-292.

Stanovich, K. E., & Siegel, L. S. (1994). Phenotypic performance profile of children with reading disabilities: A regression-based test of the phonological-core variable-difference model. *Journal of Educational Psychology, 86,* 24-53.

Stanovich, K. E., & West, R. F. (1989). Exposure to print and orthographic processing. *Reading Research Quarterly, 24,* 402-433.

Steiger, J. H. (1998). A note on multiple sample extensions of the RMSEA fit index. Structural Equation Modeling, 5, 411-419.

Steiger, J. H., & Lind, J. (1980, June). *Statistically based tests for the number of common factors*. Paper presented at the annual meeting of the Psychometric Society, Iowa City, IA.

Strauss, A. A., & Werner, H. (1943). Comparative psychopathology of the brain-injured child and the traumatic brain-injured adult. *American Journal of Psychiatry, 19,* 835–838.

Swanson, H. L. (1993). Working memory in learning disability subgroups. *Journal of Experimental Child Psychology, 56,* 87-114.

Swanson, H. L. (2011). Cognitive profile of adolescents with math disabilities: Are the profiles different from those with reading disabilities? *Child Neuropsychology,* DOI:10.1080/09297049.2011.589377

Swanson, H. L., & Alexander, J. E. (1997). Cognitive processes as predictors of word recognition and reading comprehension in learning disabled and skilled readers: Revisiting the specificity hypothesis. *Journal of Educational Psychology, 89,* 128-158.

Swanson, H. L., & Ashbaker, M. H. (2000). Working memory, short-term memory, speech rate, word recognition and reading comprehension in learning disabled readers: Does the executive system have a role? *Intelligence, 28,* 1-30.

Swanson, H. L., Kehler, Q., & Jerman, O. (2010). Working memory, strategy knowledge, and strategy instruction in children with reading disabilities. *Journal of Learning Disabilities, 43,* 24-47.

Tabachnik, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5th ed.). Boston, MA: Pearson.

Taub, G. E., Floyd, R. G., Keith, T. Z., & McGrew, K. S. (2008). Effects of general and broad cognitive abilities on mathematics achievement. *School Psychology Quarterly, 23,* 187-198. doi: 10.1037/1045-3830.23.2.187

Taub, G. E., & McGrew, K. S. (2004). A confirmatory factor analysis of Cattell-Horn-Carroll theory and cross-age invariance of the Woodcock-Johnson Tests of Cognitive Abilities III. *School Psychology Quarterly, 19,* 72-87.

Thurstone, L. L. (1938). *Primary mental abilities* (Psychometric Monographs, No. 1). Chicago: University of Chicago Press.

U.S. Department of Education. (2010). *Twenty-ninth annual report to Congress on the implementation of the Individuals with Disabilities Education Act*. Washington, DC: Author.

Vance, D., Dawson, J., Wadley, V., Edwards, J., Roenker, D., Rizzo, M.,...Ball, K. (2007). The accelerate study: The longitudinal effect of speed of processing training on cognitive performance of older adults. *Rehabilitation Psychology, 52,* 89-96, doi: 10.1037/0090-5550.52.1.89

Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods, 3,* 4-70.

Vellutino, F. R., Scanlon, D. M., & Lyon, G. R. (2000). Differentiating between difficult-to-remediate and readily remediated poor readers: More evidence against the IQ-achievement discrepancy definition of reading disability. *Journal of Learning Disabilities, 33,* 223-238. doi: 10.1177/002221940003300302

Vernon, P. E., (1950). *The structure of human abilities*. London: Methuen.

Wai, J., Lubinski, D., & Benbow, C. P. (2009). Spatial ability for STEM domains: Aligning over fifty years of cumulative psychological knowledge solidifies its importance. *Journal of Educational Psychology, 101,* 817-835.

Wakefield, J. C. (1992). The concept of mental disorder: On the boundary between biological facts and social values. *American Psychologist, 47,* 373-388.

Ward, L. C., Bergman, M. A., & Hebert, K. R. (2011, September 19). WAIS-IV subtest covariance structure: Conceptual and statistical considerations. *Psychological Assessment*. Advance online publication. doi:10.1037/a0025614

Wasserman, J. D., & Tulskey, D. S. (2005). A history of intelligence assessment. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (2nd ed., pp. 3-22). New York: Guilford Press.

Watkins, M. W. (2005). Diagnostic validity of Wechsler subtest scatter. *Learning Disabilities: A Contemporary Journal, 3,* 20-29.

Watkins, M. W., Glutting, J. J., & Youngstrom, E. A. (2005). Issues in subtest profile analysis. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (2nd ed., pp. 251-268). New York: Guilford Press.

Watkins, M. W., Kush, J. C., & Glutting, J. J. (1997). Discriminant and predictive validity of the WISC-III ACID profile among children with learning disabilities. *Psychology in the Schools, 34,* 309-319.

Wechsler, D. (1991). *Wechsler Intelligence Scale for Children* (3rd ed.). San Antonio, TX: Psychological Corporation.

Wechsler, D. (2001). *Wechsler Adult Intelligence Scales* (3rd ed.). San Antonio, TX: Psychological Corporation.

Wechsler, D. (2003). *Wechsler Intelligence Scale for Children* (4th ed.). San Antonio, TX: Psychological Corporation.

Weekes, B., Coltheart, M., & Gordon, E. (1997). Deep dyslexia and right hemisphere reading—a regional cerebral blood flow study. *Aphasiology, 11,* 1139-1158.

Widaman, K. F., & Reise, S. P. (1997). Exploring the measurement invariance of psychological instruments: Applications in the substance abuse domain. In K. J. Bryant, M. Windle, & S. G. West (Eds.), *The science of prevention: Methodological advances from alcohol and substance abuse research* (pp. 281-324). Washington, DC: American Psychological Association.

Willcutt, E. G., & Pennington, B. F. (2000). Psychiatric comorbidity in children and adolescents with reading disability. *Journal of Child Psychology and Psychiatry, 41,* 1039-1048.

Wodrich, D. L., Spencer, M. L. S., & Daley, K. B. (2006). Combining RTI and psychoeducational assessment: What we must assume to do otherwise. *Psychology in the Schools, 43,* 797-806. doi: 10.1002/pits.20189

Wolf, M., & Bowers, P. G. (1999). The double-deficit hypothesis for the developmental dyslexias. *Journal of Educational Psychology, 91,* 415-438.

Wolf, M., & Katzir-Cohen, T. (2001). Reading fluency and its intervention. *Scientific Studies of Reading, 5,* 211-238.

Wolf, M., Miller, L., & Donnelly, K. (2000). Retrieval, automaticity, vocabulary elaboration, orthography (RAVE-O): A comprehensive, fluency-based reading intervention program. Journal of Learning Disabilities, 33, 375-386. doi: 10.1177/002221940003300408

Woodcock, R. W. (1990). Theoretical foundations of the WJ-R measures of cognitive ability. *Journal of Psychoeducational Assessment, 8,* 231-258.

Woodcock, R. W., & Johnson, M. B. (1977). *Woodcock-Johnson Psycho-Educational Battery*. Hingham, MA: Teaching Resources.

Woodcock, R. W., & Johnson, M. B. (1989). *Woodcock-Johnson Psycho-Educational Battery—Revised*. Chicago: Riverside.

Woodcock, R. W., McGrew, K. S., & Mather, N. (2001, 2007). *Woodcock-Johnson III*. Rolling Meadows, IL: Riverside Publishing.

Wright-Strawderman, C., & Watson, B. L. (1992). The prevalence of depressive symptoms in children with learning disabilities. *Journal of Learning Disabilities, 25,* 258-264.

Table 1

Demographic information for samples

	Normative Subsample	SLD Reading	SLD Math	SLD Writing
Sex				
Male	146 (48.67)	88 (48.89%)	98 (42.42%)	111 (74.50%)
Female	154 (51.33)	92 (51.11%)	133 (57.58%)	38 (25.50%)
Race				
Caucasian	217 (72.33%)	152 (84.44%)	171 (74.03%)	121 (81.21%)
African-American	50 (16.67%)	15 (8.33%)	30 (12.99%)	14 (9.40%)
Native American	8 (2.67%)	3 (1.67%)	3 (1.30%)	1 (0.67%)
Asian/Native Hawaiian/ Pacific Islander	24 (8.00%)	5 (2.78%)	3 (1.30%)	9 (6.04%)
Other	1 (0.33%)	0 (0.65%)	4 (1.73%)	1 (0.67%)
Ethnicity				
Non-Hispanic	266 (88.67)	149 (82.78%)	214 (92.64%)	135 (90.60%)
Hispanic	34 (11.33)	27 (15.00%)	17 (7.35%)	9 (6.04%)
Mother Education				
Less than 5th grade	7 (2.33%)	2 (1.11%)	2 (0.87%)	3 (2.01%)
Less than HS Diploma	19 (6.33%)	7 (3.89%)	9 (3.90%)	13 (8.72%)
HS Graduate	94 (31.33%)	21 (11.67%)	30 (12.99%)	36 (24.16%)
1 to 3 years of College	93 (31.00%)	15 (8.33%)	14 (6.06%)	22 (14.77%)
Bachelor's Degree or Higher	78 (26.00%)	19 (10.56%)	22 (9.52%)	35 (23.49%)
Missing	9 (3.00%)	116 (64.44%)	154 (66.67%)	40 (26.85%)
Father Education				
Less than 5th grade	4 (1.33%)	2 (1.11%)	1 (0.42%)	4 (2.68%)
Less than HS Diploma	28 (9.33%)	7 (3.89%)	11 (4.76%)	19 (12.75%)
HS Graduate	102 (34.00%)	19 (10.56%)	26 (11.26%)	30 (20.13%)
1 to 3 years of College	67 (22.33%)	16 (8.89%)	14 (6.06%)	14 (9.40%)
Bachelor's Degree or Higher	90 (30.00%)	20 (11.11%)	24 (10.39%)	42 (28.19%)
Missing	9 (3.00%)	116 (64.44%)	155 (67.10%)	40 (26.85%)
Age				
M (*SD*)	11.88 (4.01)	12.00 (4.02)	12.32 (3.92)	11.99 (3.97)

Table 2

Percent in each SLD group with secondary and tertiary diagnoses

Group	Percent with Secondary Diagnosis	Percent with Tertiary Diagnosis
SLD Reading	26.11%	4.03%
SLD Math	12.99%	3.46%
SLD Writing	10.07%	6.11%

Table 3

CHC categorization of cognitive and achievement tests used in study

WJ-III Test Name	CHC Broad Ability	CHC Narrow Ability
Sound Blending	Ga	Phonetic Coding (PC)
Incomplete Words	Ga	Phonetic Coding (PC)
Auditory Attention	Ga	Sound Discrimination (U3)
General Information	Gc	Knowledge (K0)
Academic Knowledge	Gc	Knowledge (K0)
Verbal Comprehension	Gc	Language Development (LD)
Picture Vocabulary	Gc	Language Development (LD)
Story Recall	Gc	Listening Ability (LS)
Oral Comprehension	Gc	Listening Ability (LS)
Analysis-Synthesis	Gf	General Sequential Reasoning (RG)
Concept Formation	Gf	Induction (I)
Number Matrices	Gf	Quantitative Reasoning (RQ)
Number Series	Gf	Sequential Reasoning (RQ)
Visual-Auditory Learning	Glr	Associative Memory (MA)
Memory for Names	Glr	Associative Memory (MA)
Retrieval Fluency	Glr/Gs/Gc	Naming Fluency (NF)
Rapid Picture Naming	Glr/Gs/Gc	Naming Fluency (NF)
Pair Cancellation	Gs	Attention/Concentration
Visual Matching	Gs	Perceptual Speed (PS)
Decision Speed	Gs	Semantic Processing Speed
Cross Out	Gs/Gv	Perceptual Speed (PS)
Memory for Words	Gsm	Memory Span (MS)
Numbers Reversed	Gsm	Working Memory (MW)
Memory for Sentences	Gsm/Gc	Memory Span (MS)
Picture Recognition	Gv	Visual Memory (MV)
Spatial Relations	Gv	Visualization (VZ)
Block Rotation	Gv	Visualization (VZ)
Calculation	Gq	Math Calculation Skills (MCS)
Math Fluency	Gq	Math Calculation Skills (MCS)
Applied Problems	Gq	Applied Math (AM)
Quantitative Concepts	Gq	Applied Math (AM)
Passage Comprehension	Grw	Reading Comprehension (RC)
Reading Vocabulary	Grw	Reading Comprehension (RC)
Letter-Word Identification	Grw	Basic Reading Skills (BRS)
Word Attack	Grw	Basic Reading Skills (BRS)
Writing Fluency	Grw	Written Expression (WE)
Writing Samples	Grw	Written Expression (WE)

Note. Gc = Comprehension/Knowledge, Gf = Fluid Reasoning, Ga = Auditory Processing, Gv = Visual-Spatial Thinking, Glr = Long-term Retrieval, Gsm = Short-term Memory, Gs = Processing Speed, Gq = Quantitative knowledge, Grw = Reading and Writing.

Table 4

Univariate descriptive statistics for the Norm and SLD Reading groups

	Normative Subsample				SLD Reading			
WJ-III Test	Mean (*SD*)	*N* (% present)	Skewness	Kurtosis	Mean (*SD*)	*N* (% present)	Skewness	Kurtosis
Verbal Comprehension	100.25 (14.58)	177 (59.0%)	-.181	.257	98.88 (15.07)	155 (86.1%)	-.139	.310
Visual-Auditory Learning	100.91 (15.43)	225 (75.0%)	.164	.769	94.95 (17.31)	152 (84.4%)	.269	.212
Spatial Relations	100.34 (15.15)	220 (73.3%)	.024	.094	102.67 (13.02)	145 (80.6%)	.137	1.744
Sound Blending	99.21 (14.37)	257 (85.7%)	-.221	-.407	101.32 (14.75)	157 (87.2%)	.305	1.310
Concept Formation	99.89 (15.14)	261 (87.0%)	-.463	.147	104.17 (14.11)	157 (87.2%)	.146	-.179
Visual Matching	99.88 (14.39)	253 (84.3%)	-.003	.189	95.40 (12.92)	153 (85.0%)	-.172	1.006
Numbers Reversed	99.64 (15.71)	236 (78.7%)	-.545	.752	97.10 (16.23)	140 (77.8%)	-.056	.416
Incomplete Words	98.45 (16.02)	231 (77.0%)	-.332	.386	100.01 (16.81)	123 (68.3%)	-.877	.994
General Information	101.36 (15.82)	237 (79.0%)	-.269	.892	98.73 (15.87)	125 (69.4%)	-.875	2.662
Retrieval Fluency	101.10 (14.89)	225 (75.0%)	.208	.215	98.92 (15.39)	135 (75.0%)	-.332	1.576
Picture Recognition	100.90 (15.00)	183 (61.0%)	-.090	.252	102.44 (11.82)	149 (82.8%)	.277	.193
Auditory Attention	100.03 (16.43)	207 (69.0%)	-.043	.530	98.63 (13.55)	121 (67.2%)	-.270	-.123
Analysis-Synthesis	100.32 (15.51)	187 (62.3%)	-.296	.843	104.48 (13.20)	148 (82.2%)	.406	.694
Decision Speed	101.26 (15.84)	217 (72.3%)	.154	.231	101.12 (14.16)	135 (75.0%)	.199	.164
Memory for Words	100.18 (16.91)	257 (85.7%)	-.374	.210	99.82 (15.19)	152 (84.4%)	.459	.611
Rapid Picture Naming	101.53 (17.44)	148 (49.3%)	.111	2.199	96.07 (12.71)	90 (50.0%)	-.160	.290
Pair Cancellation	99.75 (12.12)	231 (77.0%)	-.730	3.347	98.27 (9.59)	82 (45.6%)	-.012	.657
Letter-Word Identification	100.63 (13.88)	282 (94.0%)	-.024	.608	89.31 (14.28)	171 (95.0%)	.013	.057
Story Recall	100.70 (16.02)	219 (73.0%)	.000	.362	103.09 (14.77)	110 (61.1%)	.372	2.183
Calculation	99.13 (15.31)	283 (94.3%)	-.308	.558	102.10 (15.73)	161 (89.4%)	.357	.052
Math Fluency	100.80 (15.38)	264 (88.0%)	.453	.943	95.21 (14.00)	151 (83.9%)	.344	-.155
Writing Fluency	100.11 (14.39)	256 (85.3%)	-.026	.107	96.96 (16.75)	166 (92.2%)	1.153	2.573
Passage Comprehension	100.16 (15.45)	282 (94.0%)	-.085	.016	89.36 (15.67)	171 (95.0%)	-.365	.551
Applied Problems	100.03 (14.52)	283 (94.3%)	-.130	.370	101.79 (14.72)	160 (88.9%)	.546	-.055

(Continued)

Table 4 (cont.)

WJ-III Test	Normative Subsample				SLD Reading			
	Mean (*SD*)	*N* (% present)	Skewness	Kurtosis	Mean (*SD*)	*N* (% present)	Skewness	Kurtosis
Writing Samples	100.93 (15.01)	280 (93.3%)	-.313	1.140	97.94 (13.32)	156 (86.7%)	-.275	.109
Word Attack	100.98 (14.91)	248 (82.7%)	-.322	.237	91.39 (13.93)	134 (74.4%)	.180	.069
Picture Vocabulary	98.90 (14.70)	259 (86.3%)	.097	.054	99.18 (14.62)	101 (56.1%)	-.091	.277
Oral Comprehension	100.57 (15.80)	274 (91.3%)	-.103	.492	102.73 (14.16)	97 (53.9%)	-.630	1.037
Reading Vocabulary	101.01 (14.28)	203 (67.7%)	.121	.155	96.60 (18.21)	89 (49.4%)	.688	.486
Quant. Concepts	99.95 (15.73)	234 (78.0%)	-.480	.611	101.97 (18.58)	82 (45.6%)	.163	-.543
Academic Knowledge	100.08 (15.52)	275 (91.7%)	.035	-.114	98.82 (16.78)	94 (52.2%)	-.374	1.555
Memory for Names	101.29 (15.72)	225 (75.0%)	-.108	1.196	100.96 (15.18)	92 (51.1%)	.323	-.067
Number Series	100.98 (16.57)	272 (90.7%)	-.914	2.441	105.25 (16.90)	66 (36.7%)	-.334	-.112
Number Matrices	97.82 (16.51)	241 (80.3%)	.068	-.688	104.72 (15.64)	66 (36.7%)	-.593	-.754
Cross Out	101.27 (15.47)	206 (68.7%)	.080	.852	100.78 (14.98)	93 (51.7%)	-.348	.300
Memory for Sentences	99.47 (15.22)	259 (86.3%)	.085	.648	100.40 (14.78)	94 (52.2%)	-.110	-.360
Block Rotation	100.75 (15.31)	162 (54.0%)	-.454	4.458	105.56 (18.74)	71 (39.4%)	.536	2.029

Table 5

Univariate descriptive statistics for the SLD Math and SLD Writing groups

	SLD Math				SLD Writing			
WJ-III Test	Mean (*SD*)	*N* (% present)	Skewness	Kurtosis	Mean (*SD*)	*N* (% present)	Skewness	Kurtosis
Verbal Comprehension	98.07 (12.33)	223 (96.5%)	.064	.951	103.11 (15.26)	138 (92.6%)	-.209	1.157
Visual-Auditory Learning	91.49 (14.10)	222 (96.1%)	.396	.550	100.90 (13.97)	140 (94.0%)	-.184	-.353
Spatial Relations	98.17 (12.05)	218 (94.4%)	-.205	.993	102.33 (15.45)	127 (85.2%)	-.357	.068
Sound Blending	100.39 (12.99)	224 (97.0%)	-.407	.212	99.94 (15.90)	142 (95.3%)	.199	.333
Concept Formation	95.96 (13.84)	224 (97.0%)	-.441	.814	104.89 (15.27)	142 (95.3%)	-.410	.525
Visual Matching	94.42 (15.54)	222 (96.1%)	-.315	.579	96.59 (13.93)	141 (94.6%)	.115	.222
Numbers Reversed	96.61 (16.43)	218 (94.4%)	-.550	.651	100.18 (15.70)	127 (85.2%)	-.235	.115
Incomplete Words	105.48 (12.06)	206 (89.2%)	.063	.950	101.43 (14.31)	135 (90.6%)	.052	1.178
General Information	98.69 (12.39)	213 (92.2%)	.147	.163	103.66 (15.72)	123 (82.6%)	-.454	.538
Retrieval Fluency	98.37 (12.67)	214 (92.6%)	-.474	1.280	97.85 (14.63)	125 (83.9%)	-.359	.700
Picture Recognition	99.37 (13.02)	222 (96.1%)	.020	2.908	100.54 (14.54)	140 (94.0%)	.092	.525
Auditory Attention	99.37 (14.71)	206 (89.2%)	-.392	.456	98.12 (15.82)	117 (78.5%)	-1.393	4.247
Analysis-Synthesis	94.42 (13.61)	223 (96.5%)	.098	-.233	103.88 (17.84)	140 (94.0%)	-.741	1.974
Decision Speed	99.19 (16.19)	217 (93.9%)	-.544	1.918	100.11 (15.11)	126 (84.6%)	.468	-.263
Memory for Words	101.10 (14.80)	223 (96.5%)	-.024	.387	101.12 (14.90)	140 (94.0%)	.081	.934
Rapid Picture Naming	97.16 (13.30)	82 (35.5%)	-.073	-.092	97.35 (15.98)	115 (77.2%)	-.389	1.496
Pair Cancellation	97.27 (11.98)	202 (87.4%)	-.405	.689	99.33 (9.68)	118 (79.2%)	.529	.194
Letter-Word Identification	100.17 (11.91)	227 (98.3%)	-.685	3.926	99.43 (16.13)	147 (98.7%)	.221	1.350
Story Recall	100.45 (12.30)	211 (91.3%)	.002	-.132	108.66 (14.42)	126 (84.6%)	-.150	1.071
Calculation	86.47 (14.12)	225 (97.4%)	-1.107	2.004	104.98 (17.71)	144 (96.6%)	-.202	.049
Math Fluency	87.48 (14.33)	223 (96.5%)	.074	.257	98.13 (15.71)	133 (89.3%)	.064	.004
Writing Fluency	100.89 (16.35)	228 (98.7%)	-.195	1.859	85.64 (16.50)	144 (96.6%)	.226	1.247
Passage Comprehension	100.80 (12.82)	228 (98.7%)	-.656	2.783	100.92 (15.32)	147 (98.7%)	-.175	.156
Applied Problems	87.66 (11.89)	225 (97.4%)	-1.070	1.557	107.94 (14.99)	144 (96.6%)	.179	-.049

(Continued)

Table 5 (cont.)

WJ-III Test	SLD Math				SLD Writing			
	Mean (*SD*)	*N* (% present)	Skewness	Kurtosis	Mean (*SD*)	*N* (% present)	Skewness	Kurtosis
Writing Samples	98.30 (11.65)	220 (95.2%)	-.180	2.034	94.48 (14.64)	144 (96.6%)	-1.389	3.235
Word Attack	98.00 (12.23)	220 (95.2%)	-.376	1.910	97.71 (13.73)	146 (98.0%)	-.829	3.033
Picture Vocabulary	100.01 (13.21)	88 (38.1%)	.170	.199	102.27 (15.55)	134 (89.9%)	-.007	.221
Oral Comprehension	101.86 (15.44)	89 (38.5%)	-.007	.571	103.66 (17.41)	132 (88.6%)	.343	1.714
Reading Vocabulary	96.54 (15.71)	74 (32.0%)	-.394	.278	103.84 (16.28)	123 (82.6%)	.249	1.379
Quant. Concepts	87.62 (15.08)	132 (57.1%)	-.435	1.019	102.55 (20.70)	119 (79.9%)	-.579	1.057
Academic Knowledge	97.83 (15.37)	87 (37.7%)	-.112	-.433	104.65 (17.24)	126 (84.6%)	.122	1.367
Memory for Names	101.10 (16.97)	84 (36.4%)	.271	.987	100.92 (15.21)	129 (86.6%)	.286	.359
Number Series	90.52 (19.56)	75 (32.5%)	-.837	1.742	102.09 (20.66)	112 (75.2%)	-.526	.905
Number Matrices	92.17 (15.90)	77 (33.3%)	-.341	-.296	103.34 (15.51)	112 (75.2%)	-.253	-.635
Cross Out	99.63 (18.37)	85 (36.8%)	-.600	.776	99.85 (15.22)	130 (87.2%)	-.056	-.056
Memory for Sentences	101.34 (15.95)	85 (36.8%)	-.281	-.064	103.67 (16.04)	130 (87.2%)	-.177	1.004
Block Rotation	99.44 (14.99)	73 (31.6%)	1.517	5.995	105.97 (15.07)	106 (71.1%)	-.183	1.305

Table 6

Final CHC model results for individual groups

Group	χ^2	*df*	CFI	RMSEA	AIC	BIC	aBIC
Norm	377.87	292	.966	.031	6933.87	22530.90	12120.85
SLD Reading	426.16	294	.910	.050	6978.16	22565.68	12161.97
SLD Math	381.65	292	.955	.037	6937.65	22534.68	12124.63
SLD Writing	472.87	293	.913	.064	7026.87	22619.15	12212.27
Multi-group Model	1659.29	1171	.939	.044*	6457.29	17871.96	10253.37

Note. CFI = Comparative Fit Index, aRMSEA = Root Mean Square Error of Approximation, AIC = Akaike Information Index, BIC = Baysian Information Criterion, aBIC = Sample Size-Adjusted Baysian Information Criterion. *Adjusted for multiple groups, RMSEA*$\sqrt{K}$ groups.

Table 7

Results for measurement invariance tests

Model	χ2 (*df*)	Δχ2 (Δ*df*)	*p*	CFI	ΔCFI	aRMSEA	AIC	BIC	aBIC
1. Configural Model	1659.29 (1171)			.939		.044	6457.29	17871.96	10253.38
2. Narrow Factor Loadings Equal	1713.46 (1204)	54.17 (33)	.012	.937	-.002	.044	6445.46	17703.11	10189.33
3. Broad Factor Loadings Equal	1766.51 (1250)	53.05 (46)	.221	.936	-.001	.044	6406.51	17445.29	10077.59
4. *g* Factor Loadings Equal	1808.53 (1268)	42.01 (18)	.001	.933	-.003	.044	6412.53	17365.66	10055.12
5. Subtest Residuals Equal	2107.98 (1348)	299.45 (80)	.000	.906	-.027	.052	6551.98	17124.46	10067.98
5a. Story Recall Free for Norm only	2065.25 (1347)	256.72 (79)	.000	.911	-.022	.050	6511.25	17088.49	10028.84
5b. Rapid Picture Naming Free for Norm only	2054.57 (1346)	246.04 (78)	.000	.912	-.021	.050	6502.57	17084.57	10021.74
5c. Pair Cancellation Free	1981.64 (1343)	173.12 (75)	.000	.921	-.012	.048	6435.64	17031.92	9959.56
5d. Spatial Relations Free	1957.51 (1340)	148.98 (72)	.000	.923	-.010	.046	6417.51	17028.06	9946.17
5e. Number Matrices Free	1948.36 (1337)	139.83 (69)	.000	.924	-.009	.046	6414.36	17039.18	9947.77
5f. Concept Formation Free for Norm only	1936.17 (1336)	127.64 (68)	.000	.925	-.008	.046	6404.17	17033.75	9939.16
5g. Analysis Synthesis Free	1923.02 (1333)	114.49 (65)	.000	.927	-.006	.046	6397.02	17040.88	9936.76
5h. Incomplete Words Free for SLD Math only	1909.45 (1332)	100.93 (64)	.002	.928	-.005	.046	6385.45	17034.07	9926.77
5i. Auditory Attention Free	1897.83 (1329)	89.30 (61)	.011	.929	-.004	.044	6379.83	17042.72	9925.90
6. Narrow Ability Residual Variances Equal	1929.15 (1347)	31.32 (18)	.026	.928	-.001	.044	6375.15	16952.39	9892.74
7. Broad Ability Residual Variances Equal	1998.85 (1368)	69.70 (21)	.000	.922	-.006	.046	6402.85	16880.17	9887.21
7a. Ga Free	1976.76 (1365)	47.61 (18)	.000	.924	-.004	.046	6386.76	16878.36	9875.87
7b. Gc Free	1955.54 (1362)	26.39 (15)	.034	.926	-.002	.046	6371.54	16877.41	9865.39
8. *g* Variance Equal	1957.27 (1365)	1.73 (3)	.629	.926	.000	.046	6367.27	16858.87	9856.38

Note. CFI = Comparative Fit Index, aRMSEA = Adjusted Root Mean Square Error of Approximation, AIC = Akaike Information Index, BIC = Baysian Information Criterion, aBIC = Sample Size-Adjusted Baysian Information Criterion. Compare each model to the previous model, except models 5a-5i are compared to model 4, and models 7a and 7b are compared to model 6.

Table 8

Results for Basic Reading Skills

Model	$\chi2$ (*df*)	$\Delta\chi2$ (Δdf)	*p*	CFI	aRMSEA	AIC	BIC	aBIC
Part 1: Add constraints to individual structural paths								
1. Initial model (no equality constraints)	2108.13 (1455)			.929	.046	7388.13	19949.50	11565.57
2. K0 constrained	2108.27 (1458)	.14 (3)	.987	.929	.046	7382.27	19929.37	11554.96
3. PS constrained	2111.87 (1458)	3.73 (3)	.292	.929	.046	7385.87	19932.96	11558.55
3a. PS constrained (SLD Reading Free)	2109.49 (1457)	2.38 (1)	.123	.929	.046	7385.49	19937.34	11559.76
4. RQ constrained	2110.05 (1458)	1.92 (3)	.590	.929	.046	7384.05	19931.15	11556.74
4a. RQ constrained (Norm Free)	2110.05 (1457)	.00 (1)	.964	.929	.046	7386.05	19937.90	11560.32
4b. RQ Constrained (SLD Writing Free)	2109.99 (1457)	.07 (1)	.797	.929	.046	7385.99	19937.84	11560.25
4c. RQ Constrained (Norm & SLD Writing Free)	2109.98 (1456)	.07 (2)	.966	.929	.046	7387.98	19944.59	11563.83
5. MS constrained	2116.75 (1458)	8.61 (3)	.035	.928	.046	7390.75	19937.84	11563.43
5a. MS constrained (Norm Free)	2114.83 (1457)	1.92 (1)	.165	.928	.046	7390.83	19942.68	11565.09
5b. MS constrained (SLD Writing Free)	2113.06 (1457)	3.69 (1)	.055	.928	.046	7389.06	19940.91	11563.33
5c. MS Constrained (Norm & SLD Writing Free)	2108.94 (1456)	7.81 (2)	.020	.929	.046	7386.94	19943.55	11562.79
6. MA Constrained	2112.24 (1458)	4.10 (3)	.251	.929	.046	7386.24	19933.33	11558.92
6a. MA constrained (Norm Free)	2109.62 (1457)	2.62 (1)	.106	.929	.046	7385.62	19937.47	11559.89
7. Gsm constrained	2111.25 (1458)	3.12 (3)	.374	.929	.046	7385.25	19932.35	11557.94
7a. Gsm constrained (SLD Math Free)	2109.21 (1457)	2.05 (1)	.153	.929	.046	7385.21	19937.06	11559.47
7b. Gsm constrained (SLD Reading Free)	2111.22 (1457)	.03 (1)	.865	.929	.046	7387.22	19939.08	11561.49
7c. Gsm constrained (SLD Math & SLD Reading Free)	2108.13 (1456)	3.12 (2)	.210	.929	.046	7386.13	19942.75	11561.98
Part 2: Sensitivity Analysis								
8. All structural paths constrained	2130.40 (1473)			.928	.046	7374.40	19850.12	11523.35
9. K0 free	2128.44 (1470)	1.96 (3)	.581	.928	.046	7378.44	19868.44	11532.14
10. PS free	2122.78 (1470)	7.61 (3)	.055	.929	.046	7372.78	19862.78	11526.48
10a. PS free (SLD Reading only)	2125.74 (1472)	4.66 (1)	.031	.929	.046	7371.74	19852.22	11522.27
11. RQ free	2126.69 (1470)	3.71 (3)	.295	.928	.046	7376.69	19866.69	11530.39
11a. RQ free (Norm only)	2130.16 (1472)	.24 (1)	.625	.928	.046	7376.16	19856.64	11526.69
11b. RQ free (SLD Writing)	2130.39 (1472)	.01 (1)	.938	.928	.046	7376.39	19856.87	11526.92
11c. RQ free (Norm & SLD Writing only)	2130.15 (1471)	.25 (2)	.884	.928	.046	7378.15	19863.39	11530.27

(Continued)

Table 8 (cont.)

Model	$\chi2$ (*df*)	$\Delta\chi2$ (Δdf)	*p*	CFI	aRMSEA	AIC	BIC	aBIC
12. MS free	2122.49 (1470)	7.91 (3)	.048	.929	.046	7372.49	19862.49	11526.19
12a. MS free (Norm only)	2129.62 (1472)	.78 (1)	.376	.928	.046	7375.62	19856.10	11526.15
12b. MS free (SLD Writing)	2127.44 (1472)	2.96 (1)	.086	.928	.046	7373.44	19853.92	11523.97
12c. MS free (Norm & SLD Writing)	2125.25 (1471)	5.15 (2)	.076	.929	.046	7373.25	19858.49	11525.37
13. MA free	2128.23 (1470)	2.16 (3)	.539	.928	.046	7378.23	19868.23	11531.93
13a. MA free (Norm only)	2128.78 (1472)	1.62 (1)	.203	.928	.046	7374.78	19855.26	11525.31
14. Gsm free	2127.38 (1470)	3.01 (3)	.389	.928	.046	7377.38	19867.38	11531.08
14a. Gsm free (SLD Math only)	2127.56 (1472)	2.84 (1)	.092	.928	.046	7373.56	19854.04	11524.09
14b. Gsm free (SLD Reading only)	2129.30 (1472)	1.10 (1)	.294	.928	.046	7375.30	19855.78	11525.83
14c. Gsm free (SLD Reading & SLD Math only)	2127.39 (1471)	3.01 (2)	.222	.928	.046	7375.39	19860.62	11527.50
Part 3: Final Model Specification								
15. Final Model	2119.64 (1470)	11.50 (15)	.716	.929	.046	7369.64	19859.63	11523.33
15a. Final Model (MS constrained for Norm)	**2122.69 (1471)**	**3.05 (1)**	**.081**	**.929**	**.046**	**7370.69**	**19855.92**	**11522.80**

Note. K0, = Knowledge, PS = Perceptual Speed, RQ = Quantitative Reasoning, MS = Memory Span, MA = Associative Memory, Gsm = Short-term Memory, CFI = Comparative Fit Index, aRMSEA = Adjusted Root Mean Square Error of Approximation, AIC = Akaike Information Index, BIC = Baysian Information Criterion, aBIC = Sample Size-Adjusted Baysian Information Criterion. In Part 1, compare models that only have numbers to Model 1 (e.g., Model 2 is compared to Model 1). Models with alphabetic characters (e.g., 2a, 2b) are compared to the model without an alphabetic character to the same number (e.g., Models 2a and 2b are compared to Model 2). For the Sensitivity Analysis, all models are compared to the fully constrained model (Model 8). Model 15 is compared to Model 1. The Final Model presented in Figure 3 is in bold.

Table 9

Total, direct, and indirect effects for Basic Reading Skills

	Norm			SLD Reading			SLD Math			SLD Writing		
Cognitive ability	Total	Direct	Indirect	Total	Direct	Indirect	Total	Direct	Indirect	Total	Direct	Indirect
					Unstandardized Effects							
g	.728	--	.728	.827	--	.827	.728	--	.728	.830	--	.830
Glr	.035	--	.035	.035	--	.035	.035	--	.035	.035	--	.035
MA	.035	.035	--	.035	.035	--	.035	.035	--	.035	.035	--
Gc	.310	--	.310	.310	--	.310	.310	--	.310	.310	--	.310
K0	.321	**.321**	**--**	.321	**.321**	**--**	.321	**.321**	**--**	.321	**.321**	**--**
Gf	.149	--	.149	.149	--	.149	.149	--	.149	.149	--	.149
RQ	.149	**.149**	**--**	.149	**.149**	**---**	.149	**.149**	**--**	.149	**.149**	**--**
Gsm	.437	**.364**	**.**073	.437	**.364**	**.**073	.437	**.364**	**.**073	.593	**.364**	.228
MS	.073	.073	--	.073	.073		.073	.073	--	.228	**.228**	--
Gs	.007	--	.007	.237	--	.237	.007	--	.007	.007	--	.007
PS	.007	.007	--	.237	**.237**	--	.007	.007	--	.007	.007	--
					Standardized Effects							
g	.664	--	.664	.708	--	.708	.735	--	.735	.708	--	.708
Glr	.039	--	.039	.027	--	.027	.032	--	.032	.029	--	.029
MA	.041	.041	--	.038	.038	--	.043	.043	--	.028	.028	--
Gc	.332	--	.734	.348	--	.348	.320	--	.320	.320	--	.320
K0	.344	**.344**	--	.347	**.347**	--	.331	**.331**	--	.326	**.326**	--
Gf	.150	--	.150	.139	--	.139	.160	--	.160	.149	--	.149
RQ	.170	**.170**	--	.161	**.161**	--	.186	**.186**	--	.172	**.172**	--
Gsm	.303	**.253**	.050	.349	**.291**	.058	.355	**.296**	.059	.414	**.254**	.159
MS	.076	.076	--	.077	.077	--	.082	.082	--	.226	**.226**	--
Gs	.005	--	.005	.155	--	.155	.006	--	.006	.004	--	.004
PS	.005	.005	--	.191	**.191**	--	.006	.006	--	.005	.005	--
Variance explained in BRS		**56%**			**69%**			**64%**			**68%**	

Note. g = General Cognitive Ability, Glr = Long-term Retrieval, MA = Associative Memory, Gc = Comprehension/Knowledge, K0 = Knowledge, Gf = Fluid Reasoning, RQ = Quantitative Reasoning, Gsm = Short-term Memory, MS = Memory Span, Gs = Processing Speed, PS = Perceptual Speed, BRS = Basic Reading Skills. Direct effects in bold are statistically significant, $p < .05$.

Table 10

Results for Reading Comprehension

Model	χ2 (*df*)	Δχ2 (Δ*df*)	*p*	CFI	aRMSEA	AIC	BIC	aBIC
Part 1: Add constraints to individual structural paths								
1. Initial model (No equality constraints)	2079.26 (1466)			.932	.044	7337.26	19846.29	11497.29
2. K0 constrained	2080.04 (1469)	.78 (3)	.854	.932	.044	7332.04	19826.80	11487.32
3. RN constrained	2082.70 (1469)	3.44 (3)	.329	.932	.044	7334.70	19829.46	11489.98
3a. RN constrained (SLD Writing free)	2079.68 (1468)	3.02 (1)	.082	.932	.044	7333.68	19833.19	11490.54
4. PS constrained	2082.84 (1469)	3.57 (3)	.311	.932	.044	7334.84	19829.59	11490.12
4a. PS constrained (SLD Norm free)	2080.75 (1468)	2.09 (1)	.149	.932	.044	7334.75	19834.26	11491.61
5. RQ constrained	2085.65 (1469)	6.38 (3)	.094	.931	.044	7337.65	19832.40	11492.93
5a. RQ constrained (SLD Reading free)	2079.94 (1468)	5.70 (1)	.017	.932	.044	7333.94	19833.46	11490.81
5b. RQ constrained (SLD Math free)	2085.59 (1468)	.06 (1)	.808	.931	.044	7339.59	19839.10	11496.45
5c. RQ constrained (SLD Reading & SLD Math free)	2079.30 (1467)	6.35 (2)	.042	.932	.044	7335.30	19839.57	11493.74
Part 2: Sensitivity Analysis								
6. All structural paths constrained	2094.57 (1478)			.931	.044	7328.57	19780.50	11469.61
7. K0 free	2093.01 (1475)	1.57 (3)	.667	.931	.044	7333.01	19799.21	11478.79
8. RN free	2091.97 (1475)	2.61 (3)	.457	.931	.044	7331.97	19798.17	11477.75
8a. RN free (SLD Writing Only)	2093.45 (1477)	1.12 (1)	.289	.931	.044	7329.45	19786.14	11472.07
9. PS free	2091.14 (1475)	3.43 (3)	.330	.932	.044	7331.14	19797.35	11476.93
9a. PS free (Norm Only)	2094.55 (1477)	.02 (1)	.893	.931	.044	7330.55	19787.24	11473.17
10. RQ free	2089.30 (1475)	5.27 (3)	.153	.932	.044	7329.30	19795.51	11475.09
10a. RQ free (SLD Reading Only)	2090.02 (1477)	4.55 (1)	.033	.932	.044	7326.02	19782.71	11468.64
10b. RQ free (SLD Math Only)	2094.55 (1477)	.02 (1)	.888	.931	.044	7330.55	19787.24	11473.17
10c. RQ free (SLD Reading & SLD Math Only)	2089.53 (1476)	5.05 (2)	.080	.932	.044	7327.53	19788.98	11471.73
Part 3: Final Model Specification								
11. Final Model	**2090.02 (1477)**	**10.76 (11)**	**.463**	**.932**	**.044**	**7326.02**	**19782.71**	**11468.64**

Note. K0 = Knowledge, RN = Rapid Naming, PS = Perceptual Speed, RQ = Quantitative Reasoning, CFI = Comparative Fit Index, aRMSEA = Adjusted Root Mean Square Error of Approximation, AIC = Akaike Information Index, BIC = Baysian Information Criterion, aBIC = Sample Size-Adjusted Baysian Information Criterion. In Part 1, compare models that only have numbers to Model 1 (e.g., Model 2 is compared to Model 1). Models with alphabetic characters (e.g., 2a, 2b) are compared to the model without an alphabetic character to the same number (e.g., Models 2a and 2b are compared to Model 2). For the Sensitivity Analysis, all models are compared to the fully constrained model (Model 6). Model 11 is compared to Model 1. The Final Model presented in Figure 4 is in bold.

Table 11

Total, direct, and indirect effects for Reading Comprehension

	Norm			SLD Reading			SLD Math			SLD Writing		
Cognitive Ability	Total	Direct	Indirect	Total	Direct	Indirect	Total	Direct	Indirect	Total	Direct	Indirect
					Unstandardized Effects							
g	.685	--	.685	.838	--	.838	.685	--	.685	.685	--	.685
Gc	.580	--	.580	.580	--	.580	.580	--	.580	.580	--	.580
K0	.569	**.569**	--	.569	**.569**	--	.569	**.569**	--	.569	**.569**	--
Gf	.157	--	.157	.309	--	.309	.157	--	.157	.157	--	.157
RQ	.157	**.157**	--	.309	**.309**	--	.157	**.157**	--	.157	**.157**	--
Gs	.113	--	.113	.113	---	.113	.113	--	.113	.113	--	.113
PS	.070	.070	--	.070	.070	--	.070	.070	--	.070	.070	--
RN	.090	.090	--	.090	.090	--	.090	.090	--	.090	.090	--
					Standardized Effects							
g	.734	--	.734	.747	--	.747	.841	--	.841	.746	--	.746
Gc	.727	--	.727	.672	--	.672	.725	--	.725	.759	--	.759
K0	.702	**.702**	--	.632	**.632**	--	.717	**.717**	--	.759	**.759**	--
Gf	.184	--	.184	.297	--	.297	.204	--	.204	.198	--	.198
RQ	.205	**.205**	--	.359	**.359**	--	.235	**.235**	--	.228	**.228**	--
Gs	.095	--	.095	.095	--	.095	.113	--	.113	.094	--	.094
PS	.066	.066	--	.058	.058	--	.075	.075	--	.067	.067	--
RN	.111	.111	--	.069	.069	--	.097	.097	--	.098	.098	--
Variance explained in RC		**89%**			**94%**		.	**98%**			**97%**	

Note. g = General Cognitive Ability, Gc = Comprehension/Knowledge, K0 = Knowledge, Gf = Fluid Reasoning, RQ = Quantitative Reasoning, Gs = Processing Speed, PS = Perceptual Speed, RN = Rapid Naming, RC = Reading Comprehension. Direct effects in bold are statistically significant, $p < .05$.

Table 12

Results for Math Calculation Skills

Model	χ2 (*df*)	Δχ2 (Δ*df*)	*p*	CFI	aRMSEA	AIC	BIC	aBIC
Part 1: Add constraints to individual structural paths								
1. Initial model (No equality constraints)	2128.71 (1458)			.925	.046	7402.71	19949.81	11575.40
2. Gc constrained	2136.23 (1461)	7.52 (3)	.057	.924	.046	7404.23	19937.05	11572.17
2a. Gc constrained (SLD Math free)	2132.54 (1460)	3.69 (1)	.055	.925	.046	7402.54	19940.12	11572.06
3. RN constrained	2132.25 (1461)	3.54 (3)	.316	.925	.046	7400.25	19933.07	11568.19
3a. RN constrained (Norm free)	2131.16 (1460)	1.09 (1)	.296	.925	.046	7401.16	19938.74	11570.68
4. RQ constrained	2136.98 (1461)	8.27 (3)	.041	.924	.046	7404.98	19937.80	11572.92
4a. RQ constrained (Norm free)	2135.51 (1460)	1.47 (1)	.225	.924	.046	7405.51	19943.09	11575.03
4b. RQ constrained (SLD Reading free)	2134.07 (1460)	2.92 (1)	.088	.924	.046	7404.07	19941.65	11573.59
4c. RQ constrained (SLD Math free)	2132.10 (1460)	4.88 (1)	.027	.925	.046	7402.10	19939.68	11571.62
4d. RQ constrained (SLD Writing free)	2136.31 (1460)	.68 (1)	.411	.924	.046	7406.31	19943.89	11575.83
5. Gsm constrained	2130.42 (1461)	1.71 (3)	.635	.925	.046	7398.42	19931.24	11566.36
5a. Gsm constrained (SLD Math free)	2129.97 (1460)	.45 (1)	.504	.925	.046	7399.97	19937.55	11569.49
6. PS constrained	2130.28 (1461)	1.57 (3)	.667	.925	.046	7398.28	19931.10	11566.22
6a. PS constrained (Norm free)	2129.74 (1460)	.54 (1)	.462	.925	.046	7399.74	19937.32	11569.26
6b. PS constrained (SLD Reading free)	2129.80 (1460)	.48 (1)	.488	.925	.046	7399.80	19937.38	11569.32
6c. PS constrained (SLD Math)	2129.57 (1460)	.71 (1)	.400	.925	.046	7399.57	19937.15	11569.09
6d. PS constrained (SLD Writing free)	2129.96 (1460)	.32 (1)	.572	.925	.046	7399.96	19937.54	11569.48
Part 2: Sensitivity Analysis								
7. All structural paths constrained to equality	2156.99 (1473)			.923	.046	7400.99	19876.71	11549.94
8. Gc free	2149.95 (1470)	7.04 (3)	.071	.924	.046	7399.95	19889.95	11553.65
8a. Gc free (SLD Math only)	2153.80 (1472)	3.19 (1)	.074	.924	.046	7399.80	19880.28	11550.33
9. RN free	2145.88 (1470)	11.10 (3)	.011	.924	.046	7395.88	19885.88	11549.58
9a. RN free (Norm only)	2156.75 (1472)	.23 (1)	.629	.923	.046	7402.75	19883.24	11553.29
9b. RN free (SLD Reading)	2149.98 (1472)	7.01 (1)	.008	.924	.046	7395.98	19876.46	11546.51
9c. RN free (SLD Math)	2150.15 (1472)	6.84 (1)	.009	.924	.046	7396.15	19876.63	11546.68
9d. RN free (SLD Writing)	2156.99 (1472)	.00 (1)	.964	.923	.046	7402.99	19883.47	11553.52

(Continued)

Table 12 (cont.)

Model	χ2 (*df*)	Δχ2 (Δ*df*)	*p*	CFI	aRMSEA	AIC	BIC	aBIC
10. PS free	2145.09 (1470)	11.90 (3)	.008	.924	.046	7395.09	19885.09	11548.78
10a. PS free (Norm only)	2156.64 (1472)	.35 (1)	.553	.923	.046	7402.64	19883.12	11553.17
10b. PS free (SLD Reading only)	2149.63 (1472)	7.36 (1)	.007	.924	.046	7395.63	19876.11	11546.16
10c. PS free (SLD Math only)	2149.99 (1472)	7.00 (1)	.008	.924	.046	7395.99	19876.47	11546.52
10d. PS free (SLD Writing only)	2156.38 (1472)	.61 (1)	.437	.923	.046	7402.38	19882.86	11552.91
11. RQ free	2141.01 (1470)	15.98 (3)	.001	.925	.046	7391.01	19881.01	11544.71
11a. RQ free (SLD Reading only)	2145.45 (1472)	11.54 (1)	.001	.925	.046	7391.45	19871.93	11541.98
11b. RQ free (SLD Math only)	2148.97 (1472)	8.01 (1)	.005	.924	.046	7394.97	19875.46	11545.51
11c. RQ free (SLD Writing only)	2156.70 (1472)	.29 (1)	.590	.923	.046	7402.70	19883.18	11553.23
11d. RQ free (Norm only)	2156.26 (1472)	.72 (1)	.395	.923	.046	7402.26	19882.74	11552.80
12. Gsm Released	2146.85 (1470)	10.14 (3)	.017	.924	.046	7396.85	19886.85	11550.55
12a. Gsm free (SLD Math only)	2151.11 (1472)	5.88 (1)	.015	.924	.046	7397.11	19877.59	11547.64
Part 3: Final Model Specification								
13. Final Model	2135.59 (1465)			.925	.046	7395.59	19909.37	11557.19
13a. Set Gc to zero for all but SLD Math	2139.59 (1466)	4.00 (1)	.046	.925	.046	7397.59	19906.62	11557.61
13b. Set RN equal across all groups	2144.16 (1468)	4.57 (2)	.102	.924	.046	7398.16	19897.67	11555.02
13c. Set PS equal across all groups	2141.45 (1468)	1.87 (2)	.393	.925	.046	7395.45	19894.97	11552.32
13d. Set Gsm equal across all groups	2140.38 (1467)	.79 (1)	.373	.925	.046	7396.38	19900.65	11554.82
13e. Set RQ equal across all groups	2148.70 (1468)	9.11 (2)	.011	.924	.046	7402.70	19902.21	11559.56
14. Final model	**2144.98 (1471)**	**5.39 (5)**	**.370**	**.924**	**.046**	**7392.98**	**19878.22**	**11545.09**

Note. Gc = Comprehension/Knowledge, RN = Rapid Naming, RQ = Quantitative Reasoning, PS = Perceptual Speed, Gsm = Short-term Memory, CFI = Comparative Fit Index, aRMSEA = Adjusted Root Mean Square Error of Approximation, AIC = Akaike Information Index, BIC = Baysian Information Criterion, aBIC = Sample Size-Adjusted Baysian Information Criterion. In Part 1, compare models that only have numbers to Model 1 (e.g., Model 2 is compared to Model 1). Models with alphabetic characters (e.g., 2a, 2b) are compared to the model without an alphabetic character to the same number (e.g., Models 2a and 2b are compared to Model 2). For the Sensitivity Analysis, all models are compared to the fully constrained model (Model 7). Model 14 is compared to Model 1. The Final Model presented in Figure 5 is in bold.

Table 13

Total, direct, and indirect effects for Math Calculation Skills

	Norm			SLD Reading			SLD Math			SLD Writing		
Cognitive ability	Total	Direct	Indirect	Total	Direct	Indirect	Total	Direct	Indirect	Total	Direct	Indirect
					Unstandardized Effects							
g	.762	--	.762	.973	--	.973	.605	--	.605	.762	--	.762
Gc	.034	.000	.034	.034	.000	.034	.156	.122	.034	.034	.000	.034
Gf	.488	--	.488	.699	--	.699	.230	--	.230	.488	--	.488
RQ	.488	**.488**	--	.699	**.699**	--	.230	.230	--	.488	**.488**	--
Gs	.231	--	.231	.231	--	.231	.231	--	.231	.231	--	.231
PS	.191	**.191**	--	.191	**.191**	--	.191	**.191**	--	.191	**.191**	--
RN	.090	.090	--	.090	.090	--	.090	.090	--	.090	.090	--
Gsm	.216	**.216**	--	.216	**.216**	--	.216	**.216**	--	.216	**.216**	--
					Standardized Effects							
g	.707	--	.707	.776	--	.776	.696	--	.696	.734	--	.734
Gc	.037	.000	.037	.036	.000	.036	.183	.143	.040	.039	.000	.039
Gf	.500	--	.500	.597	--	.597	.279	--	.279	.529	--	.529
RQ	.555	**.555**	--	.708	**.708**	--	.344	.344	--	.642	**.642**	--
Gs	.186	--	.186	.150	--	.150	.243	--	.243	.187	--	.187
PS	.168	**.168**	--	.147	**.147**	--	.224	**.224**	--	.176	**.176**	--
RN	.098	.098	--	.061	.061	--	.092	.092	--	.084	.084	--
Gsm	.148	**.148**	--	.154	**.154**	--	.194	**.194**	--	.172	**.172**	--
Variance explained in MCS		**65%**			**83%**			**61%**			**78%**	

Note. *g* = General Cognitive Ability, Gc = Comprehension/Knowledge, Gf = Fluid Reasoning, RQ = Quantitative Reasoning, Gs = Processing Speed, PS = Perceptual Speed, RN = Rapid Naming, MCS = Math Calculation Skills. Direct effects in bold are statistically significant, $p < .05$.

Table 14

Results for Applied Math

Model	χ2 (*df*)	Δχ2 (Δ*df*)	*p*	CFI	aRMSEA	AIC	BIC	aBIC
Part 1: Add constraints to individual structural paths								
1. Initial model (No equality constraints)	2071.83 (1464)			.938	.044	7333.83	19852.38	11497.02
2. K0 constrained	2074.11 (1467)	2.27 (3)	.518	.938	.044	7330.11	19834.38	11488.55
2a. K0 constrained (SLD Reading free)	2073.22 (1466)	.88 (1)	.347	.938	.044	7331.22	19840.25	11491.25
3. RQ constrained	2087.01 (1467)	15.18 (3)	.002	.937	.044	7343.01	19847.28	11501.45
3a. RQ constrained (Norm free)	2086.71 (1466)	.30 (1)	.585	.937	.044	7344.71	19853.74	11504.74
3b. RQ constrained (SLD Reading free)	2075.60 (1466)	11.41 (1)	.001	.938	.044	7333.60	19842.63	11493.62
3c. RQ constrained (SLD Math free)	2078.75 (1466)	8.26 (1)	.004	.938	.044	7336.75	19845.78	11496.78
3d. RQ constrained (SLD Writing free)	2086.50 (1466)	.50 (1)	.477	.937	.044	7344.50	19853.53	11504.53
3e. RQ constrained (SLD Math & SLD Reading free)	2072.35 (1465)	14.66 (2)	.001	.938	.044	7332.35	19846.14	11493.96
4. VZ constrained	2085.82 (1467)	13.99 (3)	.003	.937	.044	7341.82	19846.09	11500.27
4a. VZ constrained (SLD Writing free)	2078.49 (1466)	7.33 (1)	.007	.938	.044	7336.49	19845.52	11496.51
4b. VZ constrained (SLD Math Free)	2082.89 (1466)	2.93 (1)	.087	.937	.044	7340.89	19849.92	11500.92
4c. VZ constrained (SLD Math & SLD Writing free)	2071.83 (1465)	13.99 (2)	.001	.938	.044	7331.83	19845.62	11493.44
Part 2: Sensitivity Analysis								
5. All structural paths constrained to equality	2105.90 (1473)			.935	.044	7349.90	19825.63	11498.85
6. K0 free	2096.60 (1470)	9.30 (3)	.026	.936	.044	7346.60	19836.60	11500.30
6a. K0 free (SLD Reading only)	2102.60 (1472)	3.30 (1)	.069	.936	.044	7348.60	19829.08	11499.14
7. RQ free	2087.59 (1470)	18.32 (3)	.000	.937	.044	7337.59	19827.59	11491.29
7a. RQ free (Norm only)	2105.90 (1472)	.00 (1)	.956	.935	.044	7351.90	19832.38	11502.43
7b. RQ free (SLD Reading only)	2098.11 (1472)	7.79 (1)	.005	.936	.044	7344.11	19824.59	11494.64
7c. RQ free (SLD Math only)	2093.02 (1472)	12.89 (1)	.000	.937	.044	7339.02	19819.50	11489.55
7d. RQ free (SLD Writing only)	2102.37 (1472)	3.53 (1)	.060	.936	.044	7348.37	19828.85	11498.90
7e. RQ free (SLD Reading and SLD Math only)	2089.31 (1471)	16.60 (2)	.000	.937	.044	7337.31	19822.55	11489.42
8. VZ free	2095.55 (1470)	10.35 (3)	.016	.936	.044	7345.55	19835.55	11499.25
8a. VZ free (SLD Writing only)	2096.99 (1472)	8.91 (1)	.003	.936	.044	7342.99	19823.48	11493.53
8b. VZ free (SLD Math only)	2100.95 (1472)	4.96 (1)	.026	.936	.044	7346.95	19827.43	11497.48

(Continued)

Table 14 (cont.)

Model	χ2 (*df*)	Δχ2 (Δ*df*)	*p*	CFI	aRMSEA	AIC	BIC	aBIC
8c. VZ free (SLD Writing and SLD Math only)	2095.57 (1471)	10.34 (2)	.006	.936	.044	7343.57	19828.81	11495.68
Final Model Specification								
9. Final Model	2077.69 (1469)	5.86 (5)	.320	.938	.044	7329.69	19824.45	11484.97
9a. Set VZ equal for SLD Writing and SLD Math	**2077.69 (1470)**	**.00 (1)**	**.964**	**.938**	**.044**	**7327.69**	**19817.69**	**11481.39**
9b. Set RQ equal for SLD Reading and SLD Math	2102.63 (1471)	24.93 (1)	.000	.936	.044	7350.63	19835.87	11502.74

Note. K0 = Knowledge, RQ = Quantitative Reasoning, VZ = Visualization. CFI = Comparative Fit Index, aRMSEA = Adjusted Root Mean Square Error of Approximation, AIC = Akaike Information Index, BIC = Baysian Information Criterion, aBIC = Sample Size-Adjusted Baysian Information Criterion. In Part 1, compare models that only have numbers to Model 1 (e.g., Model 2 is compared to Model 1). Models with alphabetic characters (e.g., 2a, 2b) are compared to the model without an alphabetic character to the same number (e.g., Models 2a and 2b are compared to Model 2). For the Sensitivity Analysis, all models are compared to the fully constrained model (Model 5). Model 9 is compared to Model 1. The Final Model presented in Figure 6 is in bold.

Table 15

Total, direct, and indirect effects for Applied Math

	Norm			SLD Reading			SLD Math			SLD Writing		
Cognitive Ability	Total	Direct	Indirect	Total	Direct	Indirect	Total	Direct	Indirect	Total	Direct	Indirect
					Unstandardized Effects							
g	.774	--	.774	.909	--	.909	.731	--	.731	.892	--	.892
Gc	.142	--	.142	.142	--	.142	.142	--	.142	.142	--	.142
K0	.148	**.148**	--	.148	**.148**	--	.148	**.148**	--	.148	**.148**	--
Gf	.657	--	.657	.792	--	.792	.496	--	.496	.657	--	.657
RQ	.657	**.657**	--	.792	**.792**	--	.496	**.496**	--	.657	**.657**	--
VZ	-.003	-.003	--	-.003	-.003	--	.203	**.203**	--	.203	**.203**	--
					Standardized Effects							
g	.815	--	.815	.804	--	.804	.880	--	.880	.789	--	.789
Gc	.176	--	.176	.166	--	.166	.175	--	.175	.154	--	.154
K0	.182	**.182**	--	.165	**.165**	--	.182	**.182**	--	.157	**.157**	--
Gf	.773	--	.773	.754	--	.754	.620	--	.620	.680	--	.680
RQ	.820	**.820**	--	.858	**.858**	--	.717	**.717**	--	.769	**.769**	--
VZ	-.002	-.002	--	-.003	-.003	--	.192	**.192**	--	.147	**.147**	--
Variance explained in AM		**92%**			**95%**			**97%**			**94%**	

Note. *g* = General Cognitive Ability, Gc = Comprehension/Knowledge, K0 = Knowledge, Gf = Fluid Reasoning, RQ = Quantitative Reasoning, VZ = Visualization, AM = Applied Math. Bold values are statistically significant at the $p < .05$ level.

Table 16

Results for Written Expression

Model	χ2 (*df*)	Δχ2 (Δ*df*)	*p*	CFI	aRMSEA	AIC	BIC	aBIC
Part 1: Add constraints to individual structural paths								
1. Initial model (No equality constraints)	2139.49 (1459)			.923	.046	7411.49	19953.83	11582.60
2. Gc constrained	2149.05 (1462)	9.56 (3)	.023	.922	.046	7415.05	19943.12	11581.41
2a. Gc constrained (SLD Norm free)	2148.82 (1461)	.24 (1)	.628	.922	.046	7416.82	19949.64	11584.76
2b. Gc constrained (SLD Math free)	2140.12 (1461)	8.93 (1)	.003	.923	.046	7408.12	19940.94	11576.06
2c. Gc constrained (Norm & SLD Math free)	2139.49 (1460)	9.56 (2)	.008	.923	.046	7409.49	19947.07	11579.01
3. RN constrained	2143.40 (1462)	3.90 (3)	.272	.923	.046	7409.40	19937.46	11575.75
3a. RN constrained (Norm free)	2142.71 (1461)	.69 (1)	.406	.923	.046	7410.71	19943.53	11578.64
3b. RN constrained (SLD Reading free)	2142.44 (1461)	.95 (1)	.329	.923	.046	7410.44	19943.26	11578.38
3c. RN constrained (SLD Writing free)	2142.14 (1461)	1.26 (1)	.262	.923	.046	7410.14	19942.96	11578.08
4. RQ constrained	2145.43 (1462)	5.94 (3)	.115	.922	.046	7411.43	19939.49	11577.79
4a. RQ constrained (SLD Reading free)	2143.89 (1461)	1.54 (1)	.215	.922	.046	7411.89	19944.71	11579.83
4b. RQ constrained (SLD Writing free)	2144.35 (1461)	1.08 (1)	.298	.922	.046	7412.35	19945.17	11580.28
4c. RQ constrained (SLD Reading & SLD Writing free)	2141.68 (1460)	3.75 (2)	.153	.923	.046	7411.68	19949.25	11581.20
5. Gsm constrained	2142.66 (1462)	3.17 (3)	.367	.923	.046	7408.66	19936.72	11575.02
5a. Gsm constrained (SLD Math free)	2139.58 (1461)	3.08 (1)	.079	.923	.046	7407.58	19940.40	11575.52
6. VZ constrained	2145.15 (1462)	5.66 (3)	.129	.922	.046	7411.15	19939.21	11577.51
6a. VZ constrained (Norm free)	2143.19 (1461)	1.97 (1)	.161	.923	.046	7411.19	19944.01	11579.13
7. MS constrained	2141.16 (1462)	1.67 (3)	.644	.923	.046	7407.16	19935.22	11573.52
7a. MS constrained (SLD Writing free)	2140.40 (1461)	.76 (1)	.382	.923	.046	7408.40	19941.22	11576.34
Part 2: Sensitivity Analysis								
8. All structural paths constrained to equality	2160.50 (1477)			.922	.046	7396.50	19853.19	11539.12
9. Gc free	2158.59 (1474)	1.91 (3)	.592	.922	.046	7400.59	19871.56	11547.96
9a. Gc free for Norm only	2160.24 (1476)	.26 (1)	.608	.922	.046	7398.24	19859.69	11542.44
9b. Gc free for SLD Math	2158.60 (1476)	1.90 (1)	.168	.922	.046	7396.60	19858.05	11540.80
9c. Gc free for Norm & SLD Math only	2158.60 (1475)	1.90 (2)	.386	.922	.046	7398.60	19864.80	11544.38
10. RN free	2160.24 (1474)	.26 (3)	.967	.922	.046	7402.24	19873.20	11549.61

(Continued)

Table 16 (cont.)

Model	$\chi2$ (*df*)	$\Delta\chi2$ (Δdf)	*p*	CFI	aRMSEA	AIC	BIC	aBIC
10a. RN free for Norm only	2160.40 (1476)	.10 (1)	.755	.922	.046	7398.40	19859.85	11542.61
10b. RN free for SLD Reading only	2160.50 (1476)	.00 (1)	.975	.922	.046	7398.50	19859.95	11542.70
10c. RN free for SLD Writing only	2160.25 (1476)	.25 (1)	.616	.922	.046	7398.25	19859.70	11542.45
11. Gsm free	2160.00 (1474)	.50 (3)	.920	.922	.046	7402.00	19872.97	11549.37
11a. Gsm free for SLD Math only	2160.11 (1476)	.39 (1)	.531	.922	.046	7398.11	19859.56	11542.31
12. VZ free	2159.93 (1474)	.57 (3)	.902	.922	.046	7401.93	19872.89	11549.29
12a. VZ free for Norm only	2160.41 (1476)	.09 (1)	.767	.922	.046	7398.41	19859.86	11542.61
13. MS free	2160.16 (1474)	.34 (3)	.953	.922	.046	7402.16	19873.13	11549.53
13a . MS free for SLD Writing only	2160.49 (1476)	.01 (1)	.916	.922	.046	7398.49	19859.94	11542.69
14. RQ free	2159.84 (1474)	.66 (3)	.883	.922	.046	7401.84	19872.81	11549.21
14a. RQ free for SLD Reading only	2160.50 (1476)	.00 (1)	.975	.922	.046	7398.50	19859.95	11542.70
14b. RQ free for SLD Writing only	2160.16 (1476)	.34 (1)	.562	.922	.046	7398.16	19859.61	11542.37
14c. RQ free for SLD Reading & SLD Writing only	2160.13 (1475)	.36 (2)	.833	.922	.046	7400.13	19866.34	11545.92
Part 3: Final Model Specification								
15. Final Model (Gc & Gsm free for math)	2157.47 (1475)			.922	.046	7397.47	19863.67	11543.25
15a. Final model (set Gsm equal for all)	**2158.60 (1476)**	**1.13 (1)**	**.288**	**.922**	**.046**	**7396.60**	**19858.05**	**11540.80**
Alternative Models								
16. Gsm equal for all groups, MS removed	2160.86 (1477)	3.39 (2)	.183	.922	.046	7396.86	19853.55	11539.48
17. MS equal for all groups, Gsm removed	2159.66 (1477)	2.19 (2)	.334	.922	.046	7395.66	19852.35	11538.28

Note. Gc = Comprehension/Knowledge, RN = Rapid Naming, RQ = Quantitative Reasoning, Gsm = Short-term Memory, MS = Memory Span, VZ = Visualization, CFI = Comparative Fit Index, aRMSEA = Adjusted Root Mean Square Error of Approximation, AIC = Akaike Information Index, BIC = Baysian Information Criterion, aBIC = Sample Size-Adjusted Baysian Information Criterion. In Part 1, compare models that only have numbers to Model 1 (e.g., Model 2 is compared to Model 1). Models with alphabetic characters (e.g., 2a, 2b) are compared to the model without an alphabetic character to the same number (e.g., Models 2a and 2b are compared to Model 2). For the Sensitivity Analysis, all models are compared to the fully constrained model (Model 8). Model 15 is compared to Model 1. The Final Model for the initial analysis presented in Figure 7 is in bold. The alternative model presented in Figure 8 is Model 17.

Table 17

Total, direct, and indirect effects for Written Expression

	Norm			SLD Reading			SLD Math			SLD Writing		
Cognitive ability	Total	Direct	Indirect	Total	Direct	Indirect	Total	Direct	Indirect	Total	Direct	Indirect
					Unstandardized Effects							
g	.635	--	.635	.635	--	.635	.699	--	.699	.635	--	.635
Gc	.251	**.185**	.066	.251	**.185**	.066	.328	**.262**	.066	.251	**.185**	.066
Gf	.138	--	.138	.138	--	.138	.138	--	.138	.138	--	.138
RQ	.138	**.138**	--	.138	**.138**	--	.138	**.138**	--	.138	**.138**	--
Gsm	.236	.141	.095	.236	.141	.095	.236	.141	.095	.236	.141	.095
MS	.095	.095	--	.095	.095	--	.095	.095	--	.095	.095	--
Gs	.098	--	.098	.098	--	.098	.098	--	.098	.098	--	.098
RN	.198	**.198**	--	.198	**.198**	--	.198	**.198**	--	.198	**.198**	--
VZ	.151	.151	--	.151	.151	--	.151	.151	--	.151	.151	--
					Standardized Effects							
g	.758	--	.758	.749	--	.749	.888	--	.888	.830	--	.830
Gc	.370	**.272**	.097	.410	**.302**	.108	.450	**.360**	.091	.416	**.306**	.110
Gf	.181	--	.181	.179	--	.179	.185	--	.185	.212	--	.212
RQ	.202	**.202**	--	.207	**.207**	--	.216	**.216**	--	.245	**.245**	--
Gsm	.220	.132	.088	.252	.151	.101	.246	.147	.099	.252	.151	.101
MS	.127	.127	.186	.131	.131	.150	.137	.137	--	.139	.139	--
Gs	.094	--	.094	.089	--	.089	.107	--	.107	.101	--	.101
RN	.256	**.256**	--	.205	**.205**	--	.207	**.207**	--	.252	**.252**	--
VZ	.121	.121	--	.182	.182	--	.133	.133	--	.142	.142	--
Variance explained in WE		**79%**			**80%**			**90%**			**90%**	

Note. g = General Cognitive Ability, Gc = Comprehension/Knowledge, Gf = Fluid Reasoning, RQ = Quantitative Reasoning, Gsm = Short-term Memory, MS = Memory Span, Gs = Processing Speed, RN = Rapid Naming, VZ = Visualization, WE = Written Expression. Bold direct effects are statistically significant at the $p < .05$ level.

Table 18

Total, direct, and indirect effects for alternative Written Expression model

	Norm			SLD Reading			SLD Math			SLD Writing		
Cognitive ability	Total	Direct	Indirect	Total	Direct	Indirect	Total	Direct	Indirect	Total	Direct	Indirect
					Unstandardized Effects							
g	.625	--	.625	.625	--	.625	.691	--	.691	.625	--	.625
Gc	.251	**.180**	.071	.251	**.180**	.071	.330	**.259**	.071	.251	**.180**	.071
Gf	.163	--	.163	.163	--	.163	.163	--	.163	.163	--	.163
RQ	.163	**.163**	--	.163	**.163**	--	.163	**.163**	--	.163	**.163**	--
Gsm	.145	--	.145	.145	--	.145	.145	--	.145	.145	--	.145
MS	.145	**.145**	--	.145	**.145**	--	.145	**.145**	--	.145	**.145**	--
Gs	.107	--	.107	.107	--	.107	.107	--	.107	.107	--	.107
RN	.214	**.214**	--	.214	**.214**	--	.214	**.214**	--	.214	**.214**	--
VZ	.194	**.194**	--	.194	**.194**	--	.194	**.194**	--	.194	**.194**	--
					Standardized Effects							
g	.747	--	.747	.742	--	.742	.880	--	.880	.823	--	.823
Gc	.369	**.265**	.105	.411	**.295**	.116	.454	**.356**	.098	.417	**.299**	.118
Gf	.214	--	.214	.212	--	.212	.218	--	.218	.251	--	.251
RQ	.237	**.237**	--	.245	**.245**	--	.255	**.255**	--	.290	**.290**	--
Gsm	.136	--	.136	.158	--	.158	.150	--	.150	.157	--	.157
MS	.193	**.193**	--	.204	**.204**	--	.209	**.209**	--	.215	**.215**	--
Gs	.101	--	.101	.097	--	.097	.115	--	.115	.110	--	.110
RN	.274	**.274**	--	.222	**.222**	--	.222	**.222**	--	.272	**.272**	--
VZ	.155	**.155**	--	.233	**.232**	--	.168	**.168**	--	.183	**.183**	--
Variance explained in WE		**75%**			**80%**			**95%**			**93%**	

Note. g = General Cognitive Ability, Gc = Comprehension/Knowledge, Gf = Fluid Reasoning, RQ = Quantitative Reasoning, Gsm = Short-term Memory, MS = Memory Span, Gs = Processing Speed, RN = Rapid Naming, VZ = Visualization, WE = Written Expression. Bold direct effects are statistically significant at the $p < .05$ level.

Table 19

Summary of cognitive—achievement relations

	Cognitive Abilities								
Academic Skill	Gc	KO	RN	PS	MS	Gsm	RQ	VZ	MA
BRS		**N R M W**		*N* **R**↑ *M W*	*N R M* **W**↑	**N R M W**	**N R M W**		*N R M W*
RC		**N R M W**	*N R M W*	*N R M W*			**N R**↑ **M W**		
MCS	*N R M W*	**N R M W**	*N R M W*	**N R M W**		**N R M W**	**N R**↑ *M*↓**W**		
AM		**N R M W**					**N R**↑ **M**↓ **W**	*N R* **M**↑ **W**↑	
WE	**N R M**↑ **W**		**N R M W**		*N R M W*	*N R M W*	**N R M W**	*N R M W*	
WE (Alt.)	**N R M**↑ **W**		**N R M W**		**N R M W**		**N R M W**	**N R M W**	

Note. Gc = Comprehension/Knowledge, K0 = Knowledge, RN = Rapid Naming, PS = Perceptual Speed, MS = Memory Span, Gsm = Short-term Memory, RQ = Quantitative Reasoning, VZ = Visualization, MA = Associative Memory, N = Norm group, R = SLD Reading group, M = SLD Math group, W = SLD Writing group, BRS = Basic Reading Skills, RC = Reading Comprehension, MCS = Math Calculation Skills, AM = Applied Math, WE = Written Expression, N=Normative sample, R = SLD Reading, M = SLD Math, W = SLD Writing. Groups which are bold indicate relations that were statistically significantly different from zero between the cognitive ability and academic skill in the final model. Groups in italics indicate paths which were included in the model but were not statistically significant. Underlined groups indicate that the magnitude of the path was statistically significantly different from the other groups in the model. The arrow next to the group indicates whether the magnitude of the path is higher or lower than the other groups.

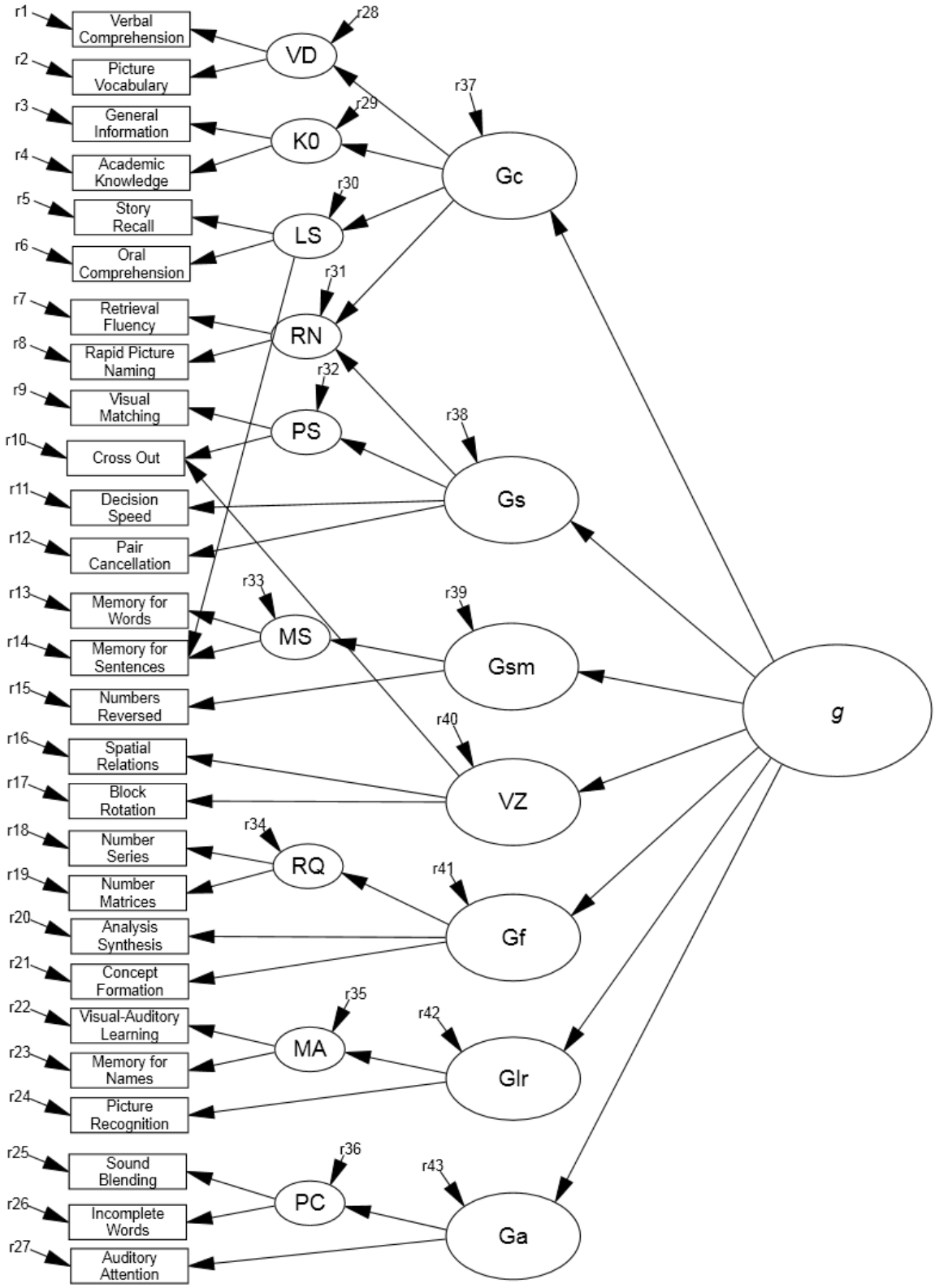

Figure 1. Hypothesized CHC Model used for identification. See Table 3 for definitions of latent variables.

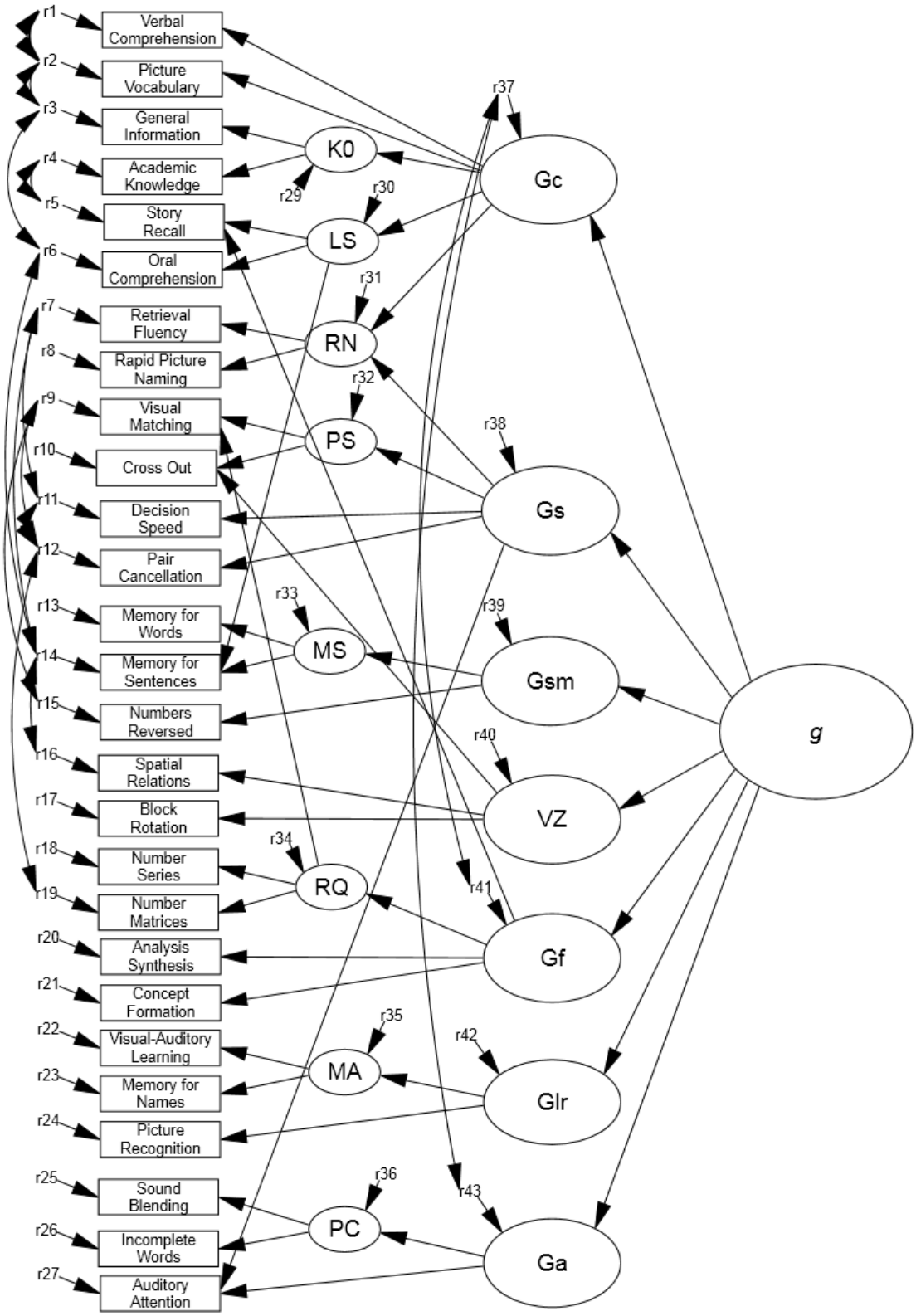

Figure 2. Final CHC Model used for analysis. See Table 3 for definitions of latent variables.

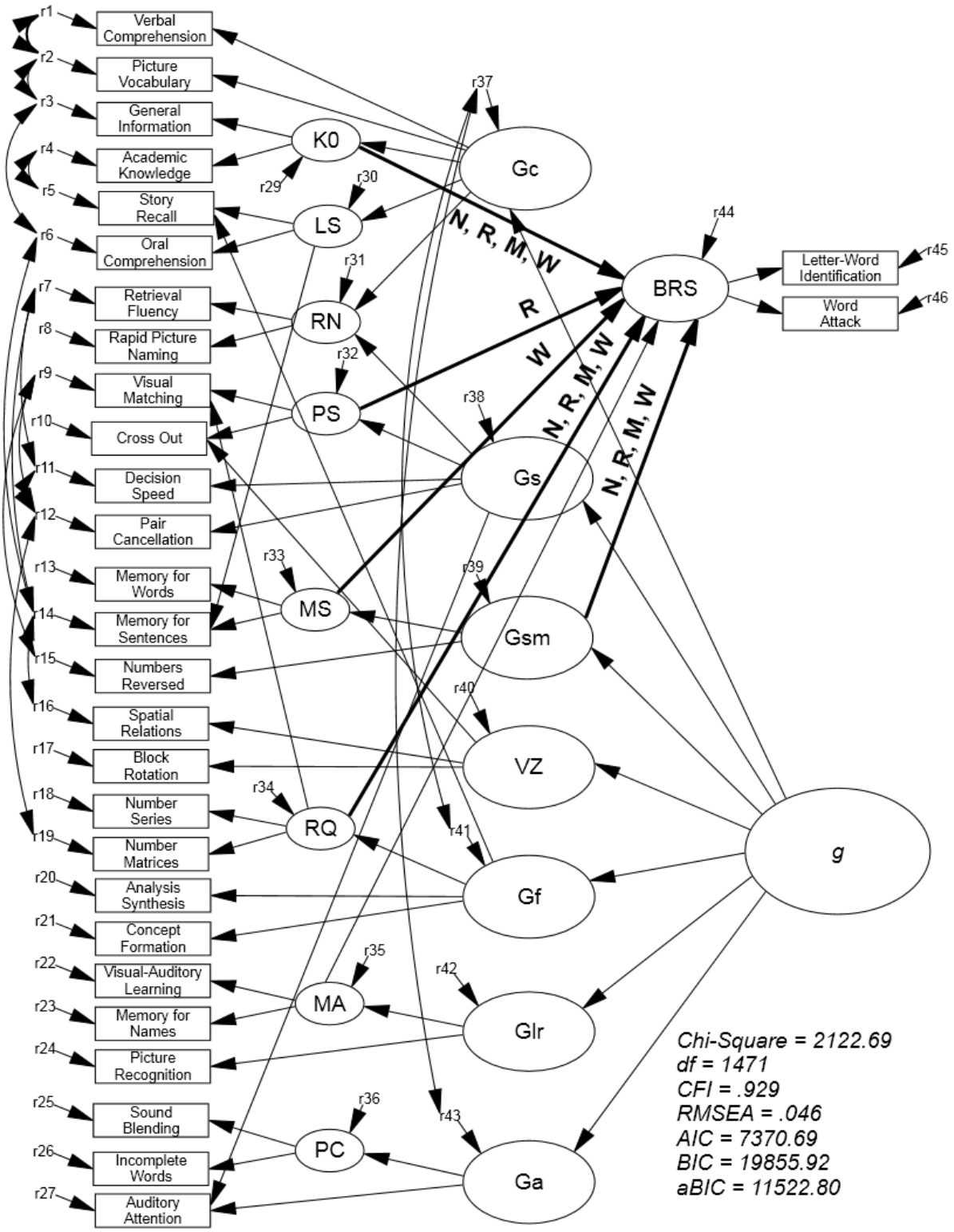

Figure 3. Structural model for Basic Reading Skills. See Table 3 for definitions of latent variables. N = Norm, R = SLD Reading, M = SLD Math, W = SLD Writing. Structural paths in bold indicate a statistically significant relation between the cognitive ability and academic skill. Arrows pointing up next to the group name indicate a statistically significant larger magnitude; arrows pointing down indicate a statistically significant smaller magnitude.

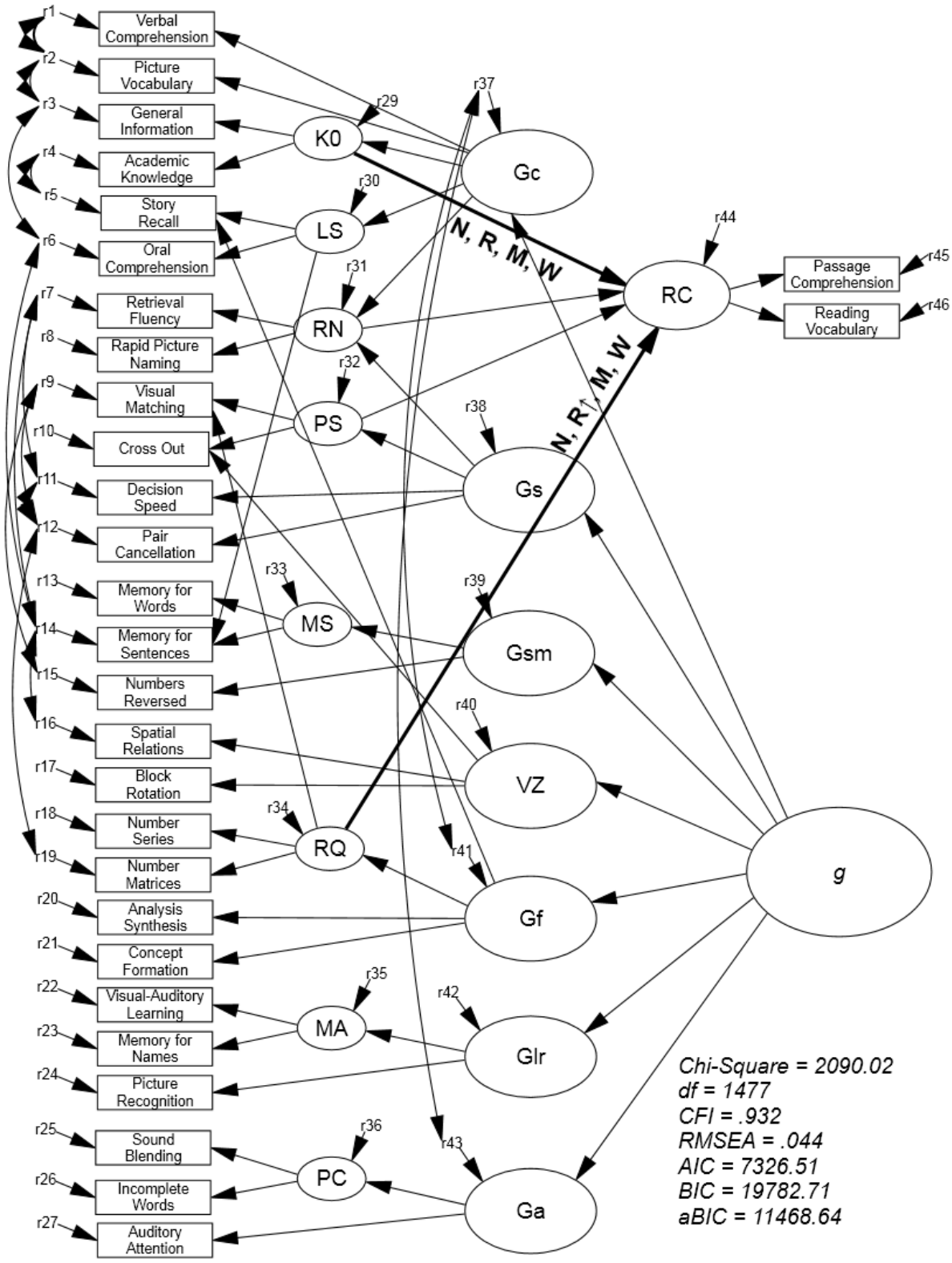

Figure 4. Structural model for Reading Comprehension. See Table 3 for definitions of latent variables. N = Norm, R = SLD Reading, M = SLD Math, W = SLD Writing. Structural paths in bold indicate a statistically significant relation between the cognitive ability and academic skill. Arrows pointing up next to the group name indicate a statistically significant larger magnitude; arrows pointing down indicate a statistically significant smaller magnitude.

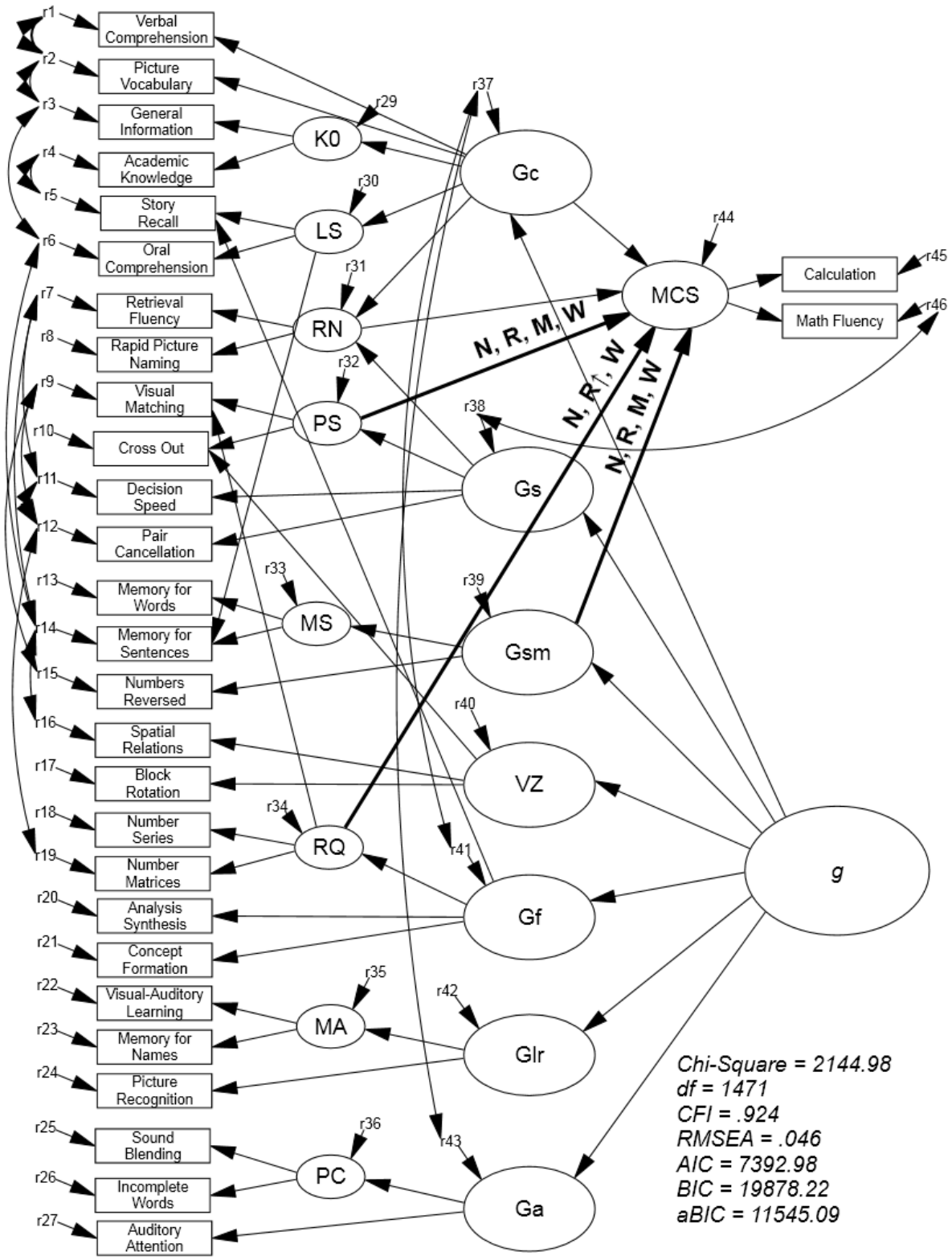

Figure 5. Structural model for Math calculation Skills. See Table 3 for definitions of latent variables. N = Norm, R = SLD Reading, M = SLD Math, W = SLD Writing. Structural paths in bold indicate a statistically significant relation between the cognitive ability and academic skill. Arrows pointing up next to the group name indicate a statistically significant larger magnitude; arrows pointing down indicate a statistically significant smaller magnitude.

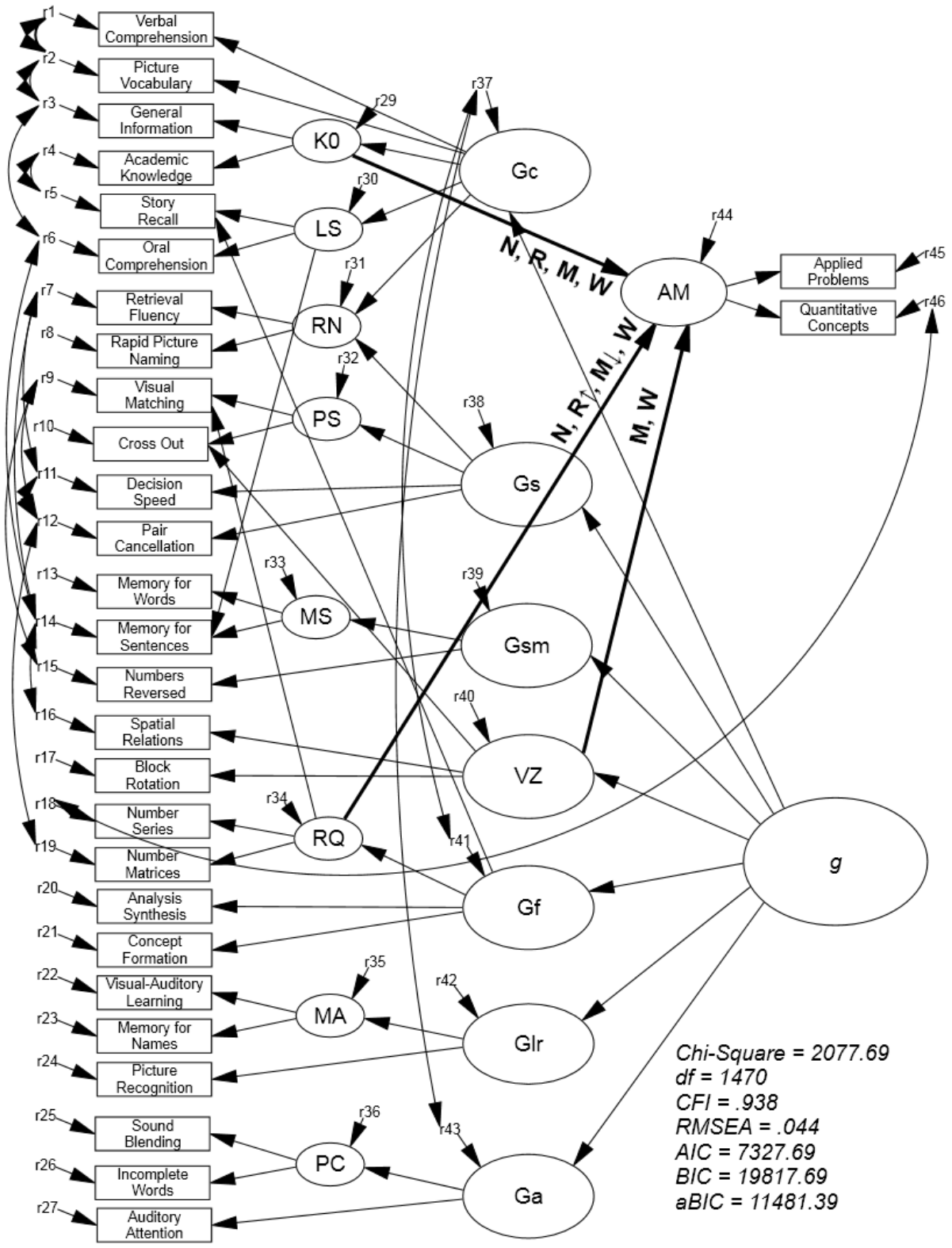

Figure 6. Structural model for Applied Math. See Table 3 for definitions of latent variables N = Norm, R = SLD Reading, M = SLD Math, W = SLD Writing. Structural paths in bold indicate a statistically significant relation between the cognitive ability and academic skill. Arrows pointing up next to the group name indicate a statistically significant larger magnitude; arrows pointing down indicate a statistically significant smaller magnitude.

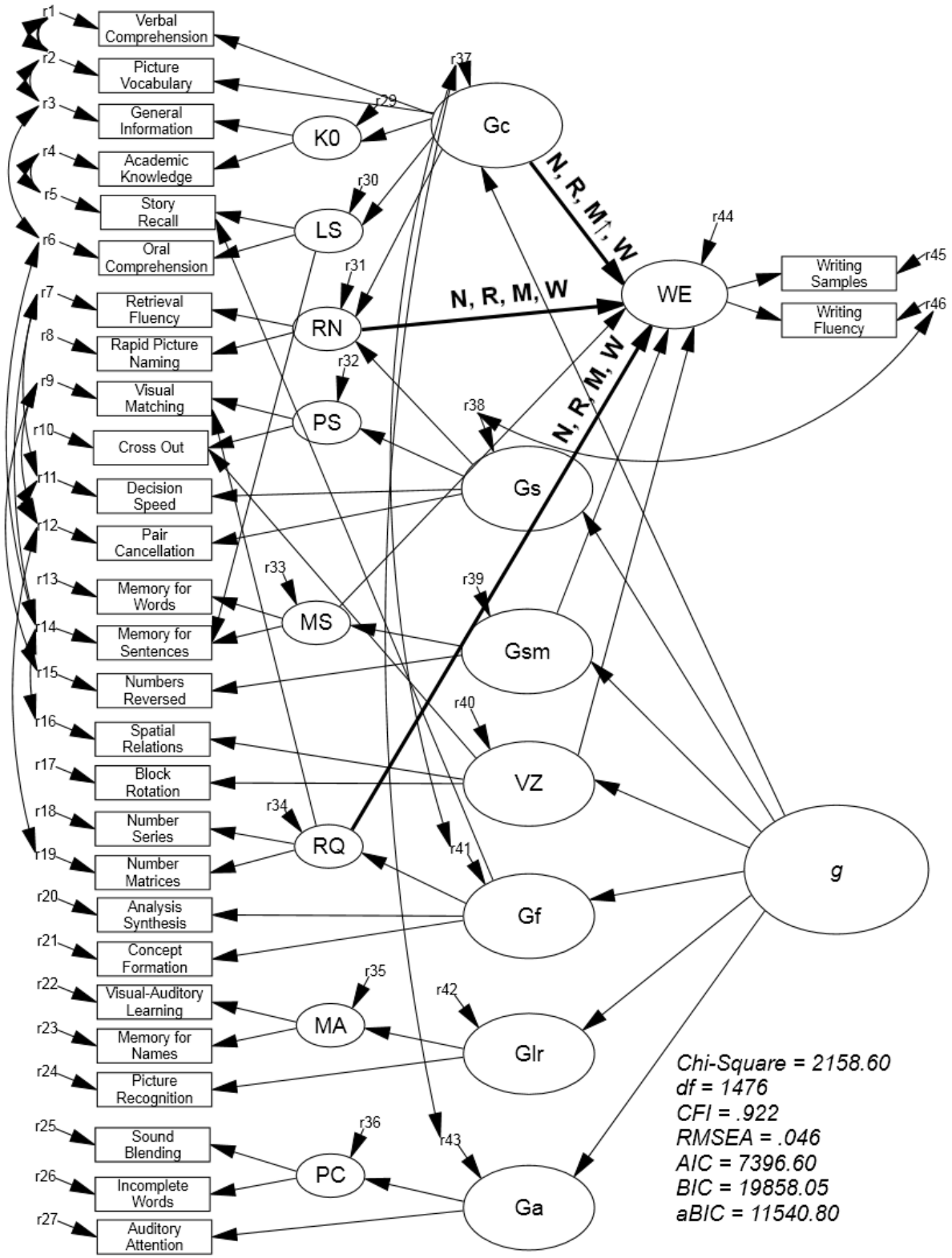

Figure 7. Final model for Written Expression. See Table 3 for definitions of latent variables. N = Norm, R = SLD Reading, M = SLD Math, W = SLD Writing. Structural paths in bold indicate a statistically significant relation between the cognitive ability and academic skill. Arrows pointing up next to the group name indicate a statistically significant larger magnitude; arrows pointing down indicate a statistically significant smaller magnitude.

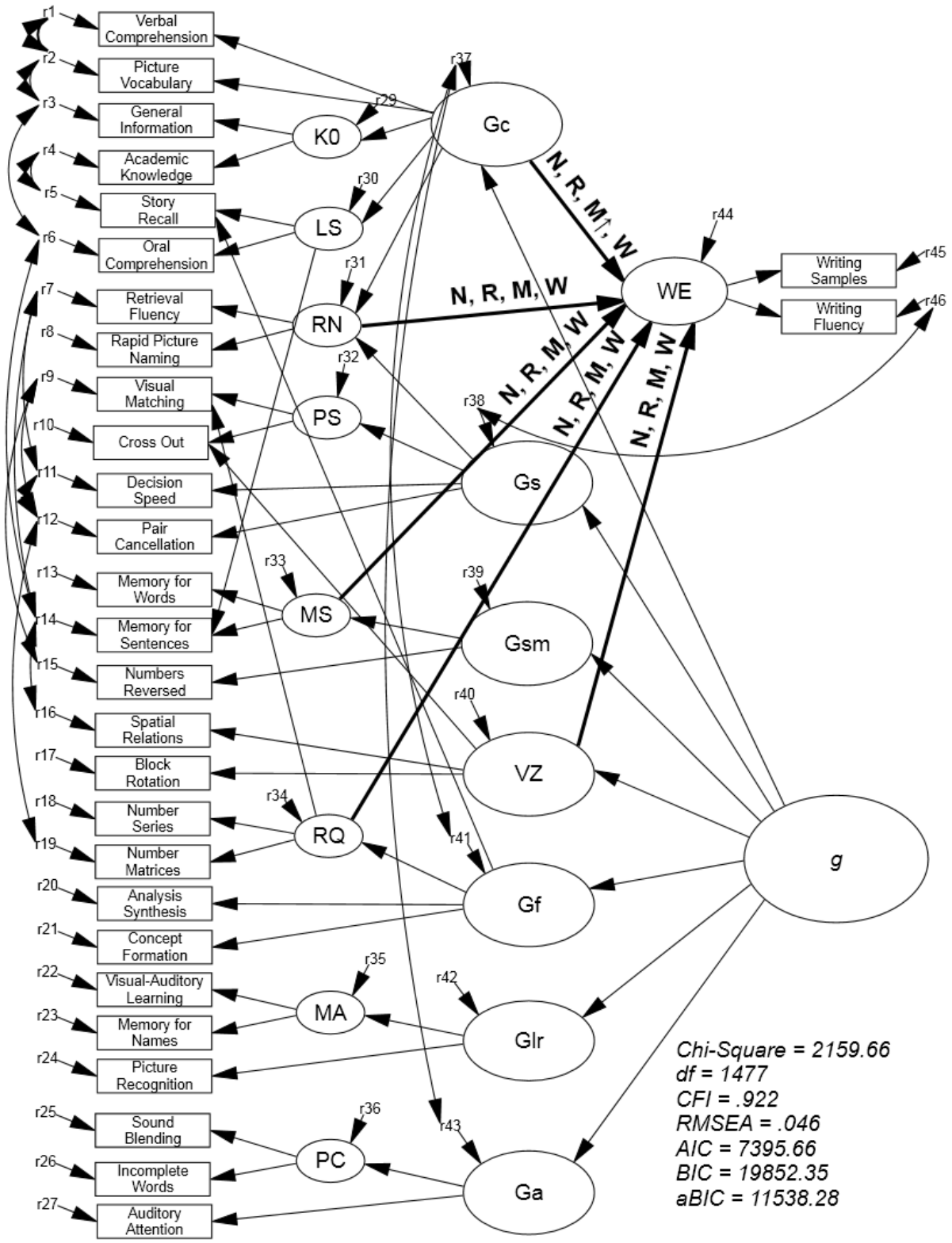

Figure 8. Alternative model for Written Expression. See Table 3 for definitions of latent variables. N = Norm, R = SLD Reading, M = SLD Math, W = SLD Writing. Structural paths in bold indicate a statistically significant relation between the cognitive ability and academic skill. Arrows pointing up next to the group name indicate a statistically significant larger magnitude; arrows pointing down indicate a statistically significant smaller magnitude.