Is a cerebellar deficit the underlying cause of reading disabilities?

Shahrzad Irannejad

Department of Educational and Counselling Psychology

McGill University, Montreal

A thesis submitted to McGill University in partial fulfillment of the requirements for the

Degree of Doctor of Philosophy in Educational Psychology

Major in School/Applied Child Psychology

June 2010

Abstract

This thesis critically appraised the Cerebellar Deficit Theory (CDT) which claims to provide a parsimonious explanation of the two most prominent existing cognitive deficit theories of dyslexia, namely the phonological and the speed naming deficit theories. Specifically, a mild congenital cerebellar deficit is proposed to give rise to a series of impairments that eventually lead to difficulties in rapid naming, phonological processing and reading. Conceptual problems with the theoretical model behind the CDT were first identified and discussed. The behavioral evidence related to CDT was then evaluated. Following this evaluation, four major questions related to CDT were examined: (1) Was there a relationship between word reading as measured by word identification task and (a) phonological awareness, (b) reading fluency and rapid automatized naming, and (c) purported cerebellar processing tasks?; (2) Did a subgroup of children with dyslexia selected from the sample for this thesis differ in their performance on any of the motor, cerebellar, reading, phonological, and rapid naming measures when compared to a reading-age (RA) and chronological-age (CA) match control subgroup selected from the same sample?; (3) Did any of these group differences remain when the effect of attention was controlled statistically?; and (4) Did a cerebellar deficit provide a good explanatory model at the individual level? Participants were 85 children attending mainstream English schools in Quebec. All participants completed a series of motor and cerebellar-related tasks. Their intellectual functioning, single word reading, word reading efficiency, speed naming, and phonological awareness skills were also assessed. Altogether, results did not seem to support a cerebellar deficit account of dyslexia. Specifically, findings did not reveal a significant relationship between any of the literacy measures and those related to motor and cerebellar tasks. Motor and cerebellar tasks were also not successful in differentiating between 17 participants in the dyslexia subgroup and those in the RA- and CA-match control at either a group or an individual level. This pattern persisted after attention was controlled statistically. A phonological deficit, independent from a cerebellar deficit, seemed to provide the best-supported account of reading difficulties for the dyslexia subgroup in contrast with the typical readers.

Résumé

Cette thèse a offert une appréciation critique de la théorie du déficit cérébelleux (TDC) qui prétend fournir une explication parcimonieuse des deux théories du déficit cognitif les plus répandues sur la dyslexie, à savoir la théorie du déficit phonologique et celle du déficit de dénomination rapide. Plus précisément, un déficit cérébelleux congénital léger est réputé causer un ensemble de défaillances qui, à terme, génèrent des troubles de dénomination rapide, de traitement phonologique et de lecture. La thèse a d'abord identifié et commenté les problèmes conceptuels du modèle théorique à la base de la TDC. Les données comportementales liées à la TDC ont ensuite été évaluées. À la suite de cette évaluation, la thèse s'est penchée sur quatre questions d'importance portant sur la TDC : (1) Y a-t-il un lien entre la lecture des mots telle que mesurée par l'identification des mots et (a) la conscience phonologique, (b) la fluence de lecture et la dénomination rapide automatisée, et (c) les soi-disant tâches liées au traitement cérébelleux?; (2) Est-ce qu'un sous-groupe d'enfants dyslexiques choisis parmi l'échantillon utilisé pour cette thèse s'est distingué par ses résultats à l'une ou l'autre des mesures de motricité, d'activité cérébelleuse, de lecture, d'habileté phonologique et de dénomination rapide lorsqu'on l'a comparé à un sous-groupe témoin d'âge correspondant sur le plan de ses capacités de lecture et sur le plan chronologique et choisi parmi le même échantillon?; (3) Est-ce que certaines de ces différences entre les groupes ont subsisté après que l'effet d'attention a été statistiquement géré?; et (4) Le déficit cérébelleux a-t-il offert un bon modèle explicatif au niveau individuel? Les participants furent un groupe de 85 enfants fréquentant l'école anglaise régulière au Québec. Tous ont accompli un ensemble de tâches motrices et cognitives. Leur fonctionnement intellectuel, leur habileté à lire un mot, leur efficacité en lecture, leur aptitude à la dénomination rapide et leur conscience phonologique ont aussi été évalués. Les résultats globaux n'ont pas vraiment renforcé l'idée que la dyslexie s'explique par le déficit cérébelleux. Plus précisément, les conclusions n'ont pas permis d'établir de lien étroit entre l'une ou l'autre des mesures de littératie et celles concernant les tâches motrices ou cognitives. Ces dernières tâches n'ont pas davantage permis de distinguer les 17 participants du sous-groupe de dyslexiques de ceux des groupes témoins, au niveau tant collectif qu'individuel. Cette tendance s'est maintenue après la gestion statistique de l'attention. Un déficit phonologique, indépendant de tout déficit cérébelleux, a semblé fournir l'explication la plus convaincante des difficultés en lecture du sous-groupe de dyslexiques par rapport aux lecteurs typiques.

Acknowledgments

I would like to thank the Riverside and English Montreal School Boards for their initial interest in this project. Without their permission this study would not have been possible. I am most grateful to all principals and teachers who became involved in this study and made it possible for me to complete testing directly at school sites during school time. I am also deeply thankful to all parents and their children who participated in this study.

My gratitude extends to all graduate students who helped during the initial data collection for this study. I would also like to thank Dr. Frédéric Demers for his timely and efficient French translation of my abstract and Ms. Amanda Greenman for proofreading my thesis.

I also would like to acknowledge Dr. Ronald Stringer for his support and contribution during the initial parts of this project. I am most grateful to Dr. Joan Wolforth for her wisdom and advice, which has had a great significance on the turn of events for me. I truly appreciate her unbelievable support and guidance throughout this process. I am also very thankful to Mr. Gordon Dionne for his support as well as the Ph.D. support group which helped me stay on track. I also would like to sincerely thank Gail McCoubrey, my dear friend and colleague, for her constant support, kindness, and encouragement.

I would like to extend my deepest gratitude to Dr. Robert Savage for his scholarly excellence, his precision and meticulous eye for detail, his truly outstanding and unique perspective, and his incredible supervision and advice which has truly impacted the quality of this dissertation. I am enormously grateful for having the opportunity, even though for a short while and upon my exit, to finally experience such remarkable and outstanding quality of supervision which is most rare. This day would not have been possible without him and this thesis would not have become what it is today without his significant contribution. I am also extremely grateful to my Ph.D. committee members, Dr. Ingrid Sladeczek and Dr. Steven Shaw, for their time and support and for their very helpful comments and advice that has helped improve this thesis even more.

Last but not least I would like to thank all my family and friends for believing in me. I am mostly grateful to my husband for his patience, kindness, and support during this truly long journey which at times seemed endless. Portions of this research were supported by the Social Sciences and Humanities Research Council (SSHRC) grant.

Table of Contents

List of Tables

List of Figures

CHAPTER 1

Overview

In the present study, the cerebellar deficit theory of dyslexia (Nicolson & Fawcett, 1999; Nicolson, Fawcett, & Dean, 2001) and the evidence for this theory was critically evaluated and four major questions regarding this theory were investigated. The early foundation of the cerebellar deficit theory was based on the notion of automaticity in which dyslexia is viewed as a symptom of a more general and pervasive deficit in skill acquisition (Nicolson & Fawcett, 1990). In its most recent form, automatization difficulties have been linked to a mild deficit in the cerebellum (Fawcett, Nicolson, & Maclagan, 2001; Nicolson, Daum, Schugens, Fawcett, & Schulz, 2002; Nicolson & Fawcett, 1999; Nicolson, Fawcett, & Dean, 2001). According to the authors of the cerebellar deficit theory, this theory combines both a learning- and a neurological-level viewpoint and provides a parsimonious explanation of the two existing cognitive deficit theories of dyslexia, namely the phonological and the speed naming deficit theories (Fawcett et al., 2001; Nicolson & Fawcett, 1999; Nicolson et al., 2001). The proponents of the theory argue that a mild deficit in the cerebellum gives rise to a series of impairments including difficulties in visual and motor domains, deficits in central processing speed, as well as difficulties in acquisition and automatization of elementary articulatory and auditory skills. Eventually, it is suggested, these impairments lead to deficits in writing, spelling, rapid naming, phonological processing and finally reading (Fawcett et al., 2001; Nicolson & Fawcett, 1999; Nicolson et al., 2001).

In the past few years, research investigating the cerebellar deficit theory has increased in volume. However, the evidence related to this theory has remained inconsistent across studies. An appraisal of this evidence undertaken in this thesis identifies four major issues that seem to explain the inconsistent findings. The first issue appears to be a lack of control for possible confounding factors such as attentional difficulties. The proponents of the cerebellar deficit theory suggest that cerebellar impairment in children with dyslexia seems to be independent from the presence or absence of attentional difficulties (Fawcett, Nicolson, & Dean, 1996; Fawcett et al., 2001). However, evidence from some studies seems to suggest that motor and cerebellar deficits may be associated with attentional difficulties (e.g., Raberger & Wimmer, 2003; Ramus, Pidgeon, & Frith, 2003; Wimmer, Mayringer, & Raberger, 1999). A meta-analysis of studies that have compared balance function between children with dyslexia and those of control samples confirmed that the proportion of participants who were screened for Attention Deficit Hyperactivity Disorder (ADHD) symptoms was a strong predictor of balance effects across studies (Rochelle & Talcott, 2006). The second issue that may explain inconsistent findings across cerebellar deficit studies seems to be related to the method by which motor and cerebellar tasks have been assessed. Supportive evidence derived from studies that have used Likert scales to measure motor and cerebellar tasks (e.g., Fawcett & Nicolson, 1999; Fawcett & Nicolson, 1995c; Fawcett et al., 1996) may be less reliable because such scales rely on subjective estimation by observers and provide only qualitative data (Shipley & Harley, 1971). Findings from studies that have incorporated sensitive measures (e.g., Brown, et al., 1985; Moe-Nilssen, Helbostad, Talcott, & Toennessen, 2003; Needle, Fawcett, & Nicolson, 2006) that also yield continuous quantitative data seem to provide different results, often reporting no group differences on cerebellar tasks (e.g., Brown et al., 1985).

The third issue that may underlie inconsistencies across the evidence base for the cerebellar deficit theory appears to be related to a lack of a reading-level design in cerebellar deficit studies. This lack is in spite of its importance for drawing causal interpretations. In a reading-level design study, individuals with dyslexia are matched to a group of younger readers who are reading at the same developmental level as those with dyslexia (Goswami & Bryant, 1989). Generally, a reading-level design that consists of both a reading-age match and a chronologically-age match control group has been considered optimal for reading-related studies (Bryant & Goswami, 1986; Goswami, 2006; Goswami & Bryant, 1989). Positive results obtained from this design may be more easily interpreted because the two groups (i.e., poor readers vs. the reading-age match control) are at the same reading level. Hence, any discrepancies found between the groups cannot be attributed to their differing reading achievement (Bryant & Goswami, 1986; Goswami & Bryant, 1989). Many studies investigating cerebellar deficit did not include a reading-level design, (e.g., Fawcett & Nicolson, 1995c; Fawcett & Nicolson, 1999; Moe-Nilssen et al., 2003; Nicolson & Fawcett, 1990; White et al., 2006). Those studies that did include reading-level design have methodological problems associated with the design, such as lack of information on specific tests used to match the groups, (e.g., Fawcett et al., 1996) which make interpretation of findings difficult. In the only study that has used a reading-level design appropriately (Savage et al., 2005a), the findings indicate that unlike deficits in phonological and naming speed tasks, performance on balance automaticity may relate more strongly to developmental maturation than to reading skills (Savage et al., 2005a).

Finally, the fourth issue related to incongruent findings for the cerebellar deficit theory seems to be a lack of homogeneity of samples involved in the studies. Samples of children with dyslexia in these studies are often drawn from extreme populations in clinics or schools for dyslexia and diagnosed based on the heavily criticized discrepancy-based definitions (e.g., Fawcett & Nicolson, 1995b; Fawcett & Nicolson, 1995c; Fawcett & Nicolson, 1999; Fawcett et al., 1996; Fawcett et al., 2001; Nicolson & Fawcett, 1990; Nicolson & Fawcett, 1994), whereas typical readers are sampled from mainstream schools. Thus any differences found between the groups may also reflect factors that are inherent to the samples (Jackson & Butterfield, 1989).

Overall, the evaluation of the evidence in this thesis indicates that a cerebellar deficit account of dyslexia may be less supported especially because most of the existing studies have at least one of the major aforementioned shortcomings. Consequently, an investigation of the cerebellar deficit hypothesis is needed in a study that addresses all major confounds discussed here. However, there has not been a direct attempt to address all of these methodological shortcomings in a single study. This study attempts to redress these problems.

In addition to identifying the four possible shortcomings that seem to explain the inconsistent findings across cerebellar deficit studies, I also evaluated a few studies that have explored individual differences pertaining to this theory. The studies that have explored individual differences could be put into two categories, namely those that include a non-clinical sample and those that include a clinical sample. In the first category, only two studies with non-clinical samples were available (Brookes & Stirling, 2005; Savage, et al., 2005b) and both were carried out in England. Neither study found strong evidence for a relationship between reading and related measures and cerebellar-related motor processes. In the second category, again one study with a UK clinical sample exists (White et al., 2006) in which individual performances were investigated using a multiple case design. Evidence from this study indicates that for the majority of cases, reading difficulties seem to be directly associated with deficits in phonological processing, deficits that can not be accounted for by motor impairments. Nonetheless, the study has been criticized on different accounts. Most importantly, a reading-level design was not used in this study.

An investigation of individual differences, using a reading-level design, is needed with a non-clinical sample. Findings from such an investigation can contribute to knowledge in several ways. First, using a non-clinical sample that includes only children attending mainstream schools can help resolve the problems related to homogeneity in clinical samples that were discussed above. Second, it is also important to determine if motor and cerebellar tasks can be successful in identifying mainstream children with reading difficulties. Third, given that the studies exploring individual differences were carried out in England, it is important to extend these findings in different contexts to ensure that results are not limited to a specific context or curriculum. In this sense, the extention of findings here is a contribution since the provincial curricula in Canada, while varying from school to school, are different from that in England, where pre-specified objectives for teaching are implemented nationally. These include specific prescriptions to teachers (e.g., for grammar or for phonic work) that may influence children's performance in reading and other cognitive tasks.

In short, this study is an attempt to improve the design by addressing the four major aforementioned methodological issues identified as possible explanations for inconsistent findings across cerebellar deficit studies. A larger scale non-clinical sample using a more homogeneous population was recruited. Sensitive measures, yielding quantitative data, were used to assess the main motor and cerebellar tasks. As a preliminary step, it was also investigated whether the motor and cerebellar measures could be reduced into separate clusters using statistical techniques. Reducing a larger number of variables into a smaller set is recommended in order to increase the reliability and robustness of results (Stevens, 1996). A reduction of motor and cerebellar measures via statistical methods has also not been attempted in previous studies conducted by Fawcett and Nicolson (e.g., Fawcett & Nicolson, 1999; Fawcett et al., 1996). This study also included a reading-level design consisting of both a reading-age and a chronologically-age match group to investigate group differences in tasks. Additionally, taking into consideration findings derived from White et al.'s (2006) multiple case study and its implications for clinical practice, individual differences were examined in this study in addition to group comparisons. In this study, group and individual differences were also investigated when attention was controlled statistically. Following these steps, four questions related to the cerebellar deficit theory were investigated in this project.

If as claimed by the cerebellar deficit theory, a deficit in the cerebellum is the underlying cause of phonological, rapid naming, and eventually reading difficulties, then cerebellar-related measures should be correlated to tasks that measure phonological awareness, reading fluency, rapid naming, and reading. Hence the first question in this study is:

Is there a relationship between word reading as measured by word identification and (a) phonological awareness, (b) reading fluency and rapid automatized naming, and (c) purported cerebellar processing tasks?

Additionally, assuming that a cerebellar deficit explains reading difficulties in children with dyslexia and that the motor and cerebellar tasks assess a cerebellar deficit, then according to the cerebellar deficit theory children with dyslexia should display deficits in motor- and cerebellar-related tasks. Hence, the second question in this study is:

Does a subgroup of children with dyslexia selected from the sample differ in their performance on any of the motor, cerebellar, reading, phonological, and rapid naming related measures when compared to two control subgroups that were selected from the same sample and matched to the dyslexia subgroup based on (a) their reading level, and (b) chronological age?

Furthermore, if according to the cerebellar deficit theory, cerebellar impairment in children with dyslexia is independent from the presence of attentional difficulties, then group differences in the cerebellar-related measures should persist after the effect of attention is controlled. Hence, the third question in this study is:

Do any group differences in performance on the above reading and cerebellar related measures emerge when the effect of attention is controlled statistically?

Finally, considering the predicted link between a cerebellar deficit, phonological and rapid naming processes and reading, the motor and cerebellar-related measures should be successful in distinguishing between participants in the dyslexia subgroup and the reading-age match control at an individual level. Hence, the fourth question of this study is:

Does the cerebellar deficit provide a good explanatory model at the individual level?

CHAPTER 2

Literature Review

The purpose of this chapter is to critically evaluate the literature on the cerebellar deficit theory of dyslexia (Fawcett et al., 2001; Nicolson & Fawcett, 1999; Nicolson et al., 2001) and to then explore four major questions related to this theory. To this end, this chapter is divided in three parts. Part one presents a short review of definitions of dyslexia followed by a brief discussion of some problems surrounding studies that label samples of children as having dyslexia. Part two provides a description and critique of the cerebellar deficit theory and the routes through which this deficit has been suggested to lead to reading difficulties. The early evidence leading to the development of this theory will also be critically reviewed. As a result of this critique four major research issues are identified. Part three describes the contributions of this thesis to the further exploration of the cerebellar deficit hypothesis and addresses the four major questions that are investigated in this doctoral dissertation.

Part 1: Definitions of Dyslexia

Overview

Reading disability or developmental dyslexia (the latter hereafter referred to simply as dyslexia) is one of the most common forms of learning disability (Pope & Whiteley, 2003). Dyslexia has been studied in many languages around the world (Shaywitz, Morris, & Shaywitz, 2008; Vicari, Marotta, Menghini, Molinari, & Petrosini, 2003; Walker & Norman, 2006). The disorder persists throughout adulthood and is often characterized by slow, strenuous reading and poor spelling (Fawcett & Lynch, 2000; Nicolson & Fawcett, 1999; Pennington, Van Orden, Smith, Green, & Haith, 1990;

Snowling, 1998). Despite this, there is still considerable debate surrounding the definition of dyslexia.

IQ-Based Definitions vs. Response to Intervention

The classic definition of dyslexia was based on cases published in medical journals during the late 19th and early 20th century describing children who demonstrated difficulties to read despite being intelligent (Shaywitz et al., 2008). As Shaywitz et al. (2008) have indicated, the term "congenital word blindness" (later replaced by dyslexia) was used to refer to these cases. It was defined as "a congenital defect occurring in children with otherwise normal and undamaged brains characterized by a difficulty in learning to read" (Beaton, 2004; Lachmann, 2002). Since the reading difficulties experienced by this group were unexpected, identification and diagnosis of dyslexia became based on exclusion, such that diagnosis took measures of intelligence into account and required unexpected discrepancy between intelligence and reading performance (Lundberg, 1999; Reid, 2003). This view of dyslexia has remained constant across definitions of the disorder (Shaywitz et al., 2008), even though at the time this classic definition was formulated there was still a lack of empirical knowledge about the causes underlying reading difficulties. The discrepancy-based definition is also used in many provinces across Canada (Kozey & Siegel, 2008) including Quebec. According to the Learning Disabilities Association of Quebec (LDAQ) (Association Québécoise des troubles d'apprentissage), in Quebec, children who are born in Canada are also diagnosed with dyslexia using the discrepancy-based criterion (LDAQ, personal communication, March 2, 2010). Furthermore, in many studies (including both neuroscientific and behavioral studies), dyslexia is still regarded as an unexpected reading problem occurring despite normal intelligence (e.g., Fawcett & Nicolson, 1995c, 1999; Fawcett et al., 1996;

Fawcett et al., 2001; Needle et al., 2006; Nicolson & Fawcett, 1990; Nicolson & Fawcett, 1994; Stoodley, Fawcett, Nicolson, & Stein, 2006; White, et al., 2006). Consequently, in these studies measures of intelligence are still seen as a necessary part of identifying and diagnosing dyslexia.

Nevertheless, this traditional view and diagnosis of dyslexia has been heavily criticized on the basis of longstanding empirical findings; many researchers no longer regard it as best practice (Eden & Moats, 2002; Fletcher, et al., 2002; Fletcher, Francis, Rourke, Shaywitz, & Shaywitz, 1992; Kozey & Siegel, 2008; Lopez & Gonzalez, 2000; Siegel, 1992; Siegel, 2005; Stanovich, 1996; Vellutino, Scanlon, & Lyon, 2000; Vellutino, et al., 1996). The meaningfulness of IQ-achievement discrepancy has been questioned, particularly given, the absence of a clear, precise, and agreed methodology to calculate discrepancy (Siegel, 1992) and the arbitrary cut-off scores used to determine the border between typical reading and dyslexia (Siegel, 2005). Furthermore, as a large body of evidence has suggested, measures of intelligence do not seem to be relevant in the assessment and diagnosis of reading disabilities (e.g., Fletcher et al., 2002; Lopez & Gonzalez, 2000; Siegel, 1992; Stuebing, et al., 2002; Vellutino et al., 2000; Vellutino et al., 1996). For example, they do not appear to predict word reading performance in normally achieving readers (Vellutino et al., 1996), nor do they seem to be strongly related to other important predictors of reading ability such as phonological decoding or word identification abilities (e.g., Snowling, 2006; Vellutino et al., 2000). According to Stanovich (1996; 1998), there is also no evidence to support a qualitative difference in reading errors between children with high and low intellectual functioning. Many children with low intellectual functioning have also been reported to have excellent decoding skills although they may have difficulties with reading comprehension (Snowling, 2006).

Furthermore, since language problems are one of the main difficulties among children with dyslexia, it is common for intelligence scores to be lower in these children when the intelligence measures rely on verbal skills (Fletcher et al., 1992; Reid, 1994; Reid, et al., 2001). Consequently any measurement of IQ can create a bias against children with reading disabilities because intelligence tests include some of the abilities which are deficient in these children (Siegel, 1989). This bias can also extend to younger children at risk for reading difficulties for whom IQ-achievement discrepancy definitions cannot be used because they are too young to show discrepancy (Snowling, 2006). This is because achievement tests which assess skills (e.g., reading) are likely to produce high floor effects when administered to the youngest age band (e.g., 6-0 to 6-5 years old) given that the format of some of these tests may not allow for assessment of basic reading or pre-reading skills (Strauss, Sherman, & Spreen, 2006). Most of these children might also not have been exposed to reading at the time of testing (Strauss et al., 2006).

In contrast to the discrepancy-based definition that reflects a categorical view of dyslexia, some argue that dyslexia may be the extreme end of a continuum of typical reading rather than a qualitatively distinct condition. That is to say, from this view, dyslexia reflects a quantitative trait rather than a discrete clinical disorder. In this view, IQ is not considered important to the understanding of reading and there are no clear discontinuities between IQ-discrepant and IQ-non discrepant poor readers. This is supported by evidence indicating that it can often be difficult to differentiate discrepancy defined readers with dyslexia from non-discrepant poor readers (Fletcher, et al., 1994; Gustafson & Samuelsson, 1999; Stanovich, 1993; Stuebing et al., 2002; Vellutino, Fletcher, Snowling, & Scanlon, 2004; Vellutino et al., 2000). The "continuous" view of dyslexia has also been strongly supported by findings from intervention studies, which suggest that rather than being an inherent characteristic within the individual, reading disability may be a difficulty in responding to the method of instruction that can be significantly modified by experience and the use of effective interventions (Vellutino, Fletcher, Snowling, & Scanlon, 2004; Vellutino et al., 2004; Vellutino et al., 2000; Vellutino et al., 1996; Vellutino, Scanlon, Small, & Fanuele, 2006; Weber, Marx, & Schneider, 2002).

Under the influence of research findings related to the phonological deficit theory (which will be discussed later), many professional and academic organizations have moved to exclude intellectual functioning from current re-definitions and diagnoses of dyslexia. Based on these re-definitions, dyslexia is generally characterized by difficulties in decoding single words that reflect a deficit in phonological processing abilities occurring irrespective of an individual's intellectual functioning and despite appropriate reading instruction (International Dyslexia Association [IDA], 2002; Lyon, 1995; National Association of School Psychologists [NASP], 2007; Orton Dyslexia Society, 1995; Reason, Frederickson, Hefferman, Martin, & Woods, 1999). A standard position now held by many researchers is that phonological skills seem to determine a child's decoding and reading ability and differentiate between individuals with and without reading disability regardless of the level of intelligence (Vellutino et al., 2000).

Do Samples in Studies on Dyslexia Represent Dyslexia?

Considering the continuous view of dyslexia and the evidence related to intervention studies, an important question is whether samples involved in studies on dyslexia are representative of individuals with dyslexia. One of the limitations associated with many studies investigating dyslexia, which have also been reviewed in the present investigation, is a lack of knowledge about samples' prior educational history.

Nonetheless, information on reading instruction (e.g., whole language vs. phonologically based method of instruction) might be crucial as it can impact interpretation of the findings derived from these studies (Deault & Savage, in press). As Vellutino et al. (1996) pointed to Clay's (1987) argument in their article:

> Failure to control for the child's educational history is the major impediment to differential diagnosis of reading disability….Virtually all studies that have sought to evaluate basic process deficit explanations of reading disability are confounded by this problem and…the adverse effects of inadequate prereading experience, inadequate instruction, or both can often mask or even mimic the adverse effects of constitutionally based cognitive deficits. (p. 601)

Another limitation of many of the studies on dyslexia that have been reviewed in the present investigation (e.g., Fawcett & Nicolson, 1995b, 1995c, 1999; Fawcett et al., 1996; Fawcett et al., 2001; Moe-Nilssen et al., 2003; Nicolson & Fawcett, 1990; Nicolson & Fawcett, 1994; Raberger & Wimmer, 2003; White et al., 2006; Wimmer et al., 1999) is a lack of knowledge on how the samples involved respond to intervention. This is understandable considering the practical issues that may be associated with gathering such knowledge in a research project. Nonetheless, without this knowledge, it may not really be clear whether a given sample diagnosed with dyslexia is environmentally or genetically reading disabled. In addition to a lack of knowledge about samples' educational background and their response to intervention, many of the studies on dyslexia that have been reviewed in this thesis have ignored empirical findings surrounding response to intervention and have also continued using the out-dated and heavily criticized IQ-achievement discrepancy approach to diagnose and identify dyslexia.

Studies continue to identify and label the children in their samples as having dyslexia without acknowledging these limitations. However, in light of the evidence derived from intervention studies as well as the use of the much criticized IQ-achievement discrepancy method as the diagnostic criterion, it is certainly debatable whether the samples investigated in studies on dyslexia truly represent individuals with dyslexia. Labeling children in a sample as having dyslexia may also be problematic even when an underlying inborn neurological impairment such as cerebellar deficit (to be reviewed in part 2 of this chapter) is claimed to underlie dyslexia. This is especially problematic considering the comorbidity of dyslexia with other developmental disorders (Messaoud-Galusi & Marshall, 2010) such as attentional difficulties (Willcutt & Pennington, 2008), disorders which have not been controlled for in many of these studies (also reviewed in part 2 of this chapter). The limitations addressed here also apply to the present study. Unfortunately, it was also not possible or feasible within the scope of a PhD project exploring deficits to also collect information on samples' response to intervention. With acknowledgement of the aforementioned limitations, the term dyslexia is used in the present project to label the sample of poor readers for the sake of consistency with the other studies reviewed in this thesis.

Part 2: Cerebellar Deficit Theory

Overview

The cerebellar deficit theory combines both a learning- and a neurological- based perspective to provide a unifying explanatory framework of the existing theories of dyslexia (Fawcett et al., 2001; Nicolson & Fawcett, 1999; Nicolson et al., 2001). In this view, a mild deficit in the cerebellum, present at birth, is proposed to give rise to phonological and naming speed impairments (Nicolson et al., 2001; Reynolds, Nicolson,

& Hambly, 2003). More specifically, a series of difficulties are said to follow from a deficit in the cerebellum. These include difficulties in visual and motor domains, and in central processing speed, as well as in acquiring and automatization of elementary articulatory and auditory skills. Eventually, these impairments lead to deficits in writing, spelling, rapid naming, phonological processing and finally reading (Fawcett et al., 2001; Nicolson, Daum, Schugens, Fawcett, & Schulz, 2002; Nicolson & Fawcett, 1999; Nicolson et al., 2001).

In this part of the chapter, the cerebellar deficit theory is reviewed. Conceptual problems with the theoretical model behind the cerebellar deficit theory are first identified and discussed. Specifically, the routes through which a cerebellar deficit is proposed to lead to phonological and subsequent reading impairments are critically appraised. As a result of this appraisal, the specific route linking a cerebellar deficit to reading, which is investigated in this thesis, is determined. The critical evaluation of the theoretical model is then followed by a critical review of some of the behavioral evidence for this theory. This critical review focuses on the inconsistencies across behavioral findings associated with the theory. As a result, four major research issues that may explain this variability are identified and discussed. Even though the cerebellar deficit theory has a neurological basis, the neurological evidence related to the theory is not addressed in this thesis. As Frith (1997) argued, prior to clarifying the biological basis of dyslexia, it is vital to first be clear about the behavioral manifestations of the disorder. Thus following Frith, this thesis seeks to clarify the behavioral evidence base for dyslexia in the specific domain of motor task performance.

The Cerebellar Deficit Theory and its Role in Reading

The cerebellar deficit hypothesis is a generalized deficit theory constructed from and encompassing several existing deficit theories of dyslexia, including the phonological and speed naming deficit theories. The foundation of the cerebellar deficit theory is based on the notion of automaticity in which impairments in phonological processing and naming speed difficulties are considered to be only a few of the larger pattern of behavioral indicators of dyslexia. These include deficits in motor skills and in all domains of learning (Nicolson & Fawcett, 1995; Ramus, Rosen et al., 2003; Wolf, 1999).

The present study does not directly investigate phonological and speed naming deficits. Nevertheless, tasks tapping into these processes are included in this research since difficulties in these processes are acknowledged by the cerebellar deficit theory and claimed to be subsumed under and explained by a cerebellar deficit. Therefore, while a short explanation of the two theories (i.e., phonological and speed naming deficit theories) is warranted prior to describing the cerebellar deficit theory, the phonological and speed naming deficit theories are not the focus of this thesis and they are not critically appraised in great detail. The purpose of introducing the two theories is merely to provide the reader with a description of the theories and an overview of the key findings related to them. As indicated in the overview on pages 14 to 15, the critique is focused on the theoretical model and the behavioral evidence related to the cerebellar deficit theory.

Phonological Deficit Theory

According to phonological deficit theory, phonological awareness or the ability to reflect on basic phonemic or speech components of language (Siegel, 1993), is crucial in acquiring reading skills (Wagner & Torgesen, 1987). Developing phonological awareness enables children to make sense of the alphabetic system used in their written language

(Torgesen & Wagner, 1998). It facilitates their understanding of how words in their oral language can be represented by printed letters, and allows them to map orthography or spelling units to their corresponding phonemes or speech units (Lyon, Shaywitz, & Shaywitz, 2003; Torgesen & Wagner, 1998). This grapheme-phoneme connection can then be used to read unfamiliar words. It is often argued that once children become aware that words can be divided into their basic elements of sound, and can map these sounds to their print unit, they will be able to decode words (Torgesen & Wagner, 1998). Additionally, as words are repeatedly processed through hearing, speaking, reading or writing, their phonological representations or codes (i.e., the sounds of the letters) may become stored more effectively in working memory. In time, the processing of these codes improves, since they become adequately distinct, permanent, and accessible to a child (Torgesen & Wagner, 1998; Wagner & Torgesen, 1987).

Based on the phonological deficit theory, children with dyslexia may have an underdeveloped sensitivity to phoneme structures in words leading to difficulties in learning the phonemic skills necessary to decode words in alphabetic languages (Bowers & Newby-Clark, 2002; Liberman, 1982). This underdeveloped sensitivity may lead to difficulties attending to, accessing, and isolating these phonemes in words. An impairment in phonological awareness is, therefore, assumed to prevent individuals with dyslexia from acquiring word decoding and blending skills necessary for normal reading skill acquisition (Lyon et al., 2003; Torgesen & Wagner, 1998). More specifically, findings suggest that reading difficulty is a result of phonological processing deficits including difficulties in storage and/or retrieval of phonemes and/or rimes in words (Bradley & Bryant, 1983; Siegel, 1993; Snowling, 1995; Stanovich & Siegel, 1994; Torgesen & Wagner, 1998).

The phonological deficit theory has been supported by a large and well established body of evidence derived from a range of study designs, including reading-level, intervention and a combination of both longitudinal and intervention studies (e.g., Boada & Pennington, 2006; Bradley & Bryant, 1983; Bruck, 1992; Gillon, 2004; Hatcher, Hulme, & Ellis, 1994; Herrmann, Matyas, & Pratt, 2006; Lundberg, Frost, & Petersen, 1988; Muter, Hulme, Snowling, & Taylor, 1997; Muter & Snowling, 1998; National Reading Panel, 2000; Savage & Pompey, 2008; Snowling, 1998; van IJzendoorn & Bus, 1994). As indicated earlier, only an overview of some of the major findings are presented here. To summarize, while there is still some disagreement about the way different types of phonological skills should be understood, many researchers agree that phonological awareness may be a causal factor in reading acquisition (Savage & Carless, 2005). There is now substantial evidence supporting a link between phonological processing and reading (e.g., Rack, Hulme, Snowling, & Wightman, 1994; Savage et al., 2005b; Wagner & Torgesen, 1987). That is to say, tasks that tap into phonological processing and phonological awareness not only seem to be concurrently related to reading ability (e.g., Badian, 1993; Berninger, Thalberg, DeBruyn, & Smith, 1987; French, Opatrny, & Cochran, 2008; Savage, Carless, & Ferraro, 2007), but they also seem to predict subsequent achievement and difficulties in reading after control for literacy (e.g., Cossu, Shankweiler, Liberman, Tola, & Katz, 1988; French et al., 2008; Naslund & Schneider, 1996; Pennington & Lefly, 2001; Share, Jorm, Maclean, & Matthews, 1984; Wagner & Torgesen, 1987; Wesseling & Reitsma, 2001).

The potentially causal role of deficits in phonological processing in reading difficulties also seems to be well supported in studies that have used a reading-age match design. This evidence is clear since these difficulties have been replicated in many studies of phonological awareness and studies of non-word processing that directly assess the assembled reading processes that might be assumed to rely closely on phonological awareness (e.g., Backman, 1983; Boada & Pennington, 2006; Bowey, Cain, & Ryan, 1992; Bowey & Hansen, 1994; Bradley & Bryant, 1983; Duncan & Johnston, 1999; Felton & Wood, 1992; Gillon & Dodd, 1994; Gonzalez, 1997; Lundberg & Hoien, 1990; Savage, et al., 2005a; Thompson & Johnston, 2000). Further, a systematic review of findings from studies that have investigated non-word processing seems to indicate that the length of non-words (one syllable versus polysyllable) is important in identifying a deficit in individuals with dyslexia as compared to their reading-age match controls (Rack, Snowling, & Olson, 1992). Deficits were shown more clearly when children were asked to read bi- and polysyllabic non-words than when they read monosyllabic non-words.

Additionally, findings from intervention studies have also demonstrated that the majority of children with reading difficulties make progress in tasks such as word identification, spelling, and reading ability in response to direct instruction designed to improve phonological awareness and phonologically based decoding skills (Vellutino et al., 2004). This evidence also seems to be supported by findings from neuroimaging studies that have incorporated short or long term training for children experiencing reading difficulties (e.g., Shaywitz, et al., 2004; Simos, et al., 2002). Findings from these studies seem to confirm that reading skills in children with reading difficulties not only improve at the behavioral level, but that positive changes also occur at the neurological level as reflected in an improved activation level of brain areas related to phonological processing.

The Rapid Naming Deficit: Automaticity in Reading

The phonological deficit has generally been accepted as the most plausible explanation for reading difficulties experienced in dyslexia (Pope & Whiteley, 2003). However, some research has proposed that there are additional difficulties in dyslexia that cannot be explained by phonological processes. These difficulties are presumed to reflect an underlying automaticity deficit in reading (e.g.,Wolf & Bowers, 1999) which involve impairment in tasks that require automatic and fast retrieval process, namely rapid naming of stimuli such as objects, colors, digits, and letters (Denckla & Rudel, 1976a; Wolf, Bowers, & Biddle, 2000; Wolf & O'Brien, 2006). It is argued that difficulties experienced in these rapid retrieval processes, which can exist independently of phonological processes, lead to a reduced rate of word recognition, and thus interfere with the individual's ability to become an automatic and fluent reader (Bowers, 1995; Bowers & Wolf, 1993a, 1993b; Wolf & Bowers, 1999; Wolf et al., 2000; Wolf & O'Brien, 2006). Dysfluent word recognition is assumed to be due to deficits in the underlying processes that contribute to naming speed performance, such as attentional, visual and perceptual, memory and recognition processes, phonological, lexical, temporal, and motoric processes (Wolf, 1997; Wolf & Bowers, 1999; Wolf et al., 2000). Although rapid naming tasks also involve a phonological component along with the other underlying processes, these tasks are not viewed as phonological processing tasks and are thus not categorized as such (Wolf, 1997; Wolf & Bowers, 1999; Wolf et al., 2000). The primary emphasis is rather on the precise timing and speed of processing required within each and within all components underlying speed naming (Bowers, 1995; Wolf & Bowers, 1999; Wolf et al., 2000).

The evidence related to naming speed deficits in individuals with dyslexia has been derived from correlational, longitudinal, and cross-sectional studies, as well as some reading-level studies (e.g., Ackerman & Dykman, 1993; Bowers, Steffy, & Tate, 1988; Bowers & Swanson, 1991; Chiappe, Stringer, Siegel, & Stanovich, 2002; Denckla & Rudel, 1976b; Fawcett & Nicolson, 1994; Kirby, Desrochers, Roth, & Lai, 2008; Korhonen, 1991; Korhonen, 1995; Manis, Doi, & Bhadha, 2000; Plaza & Cohen, 2003; Rudel, Denckla, & Broman, 1978; Savage, Pillay, & Melidona, 2007; Snyder & Downey, 1995; Wolf, 1982, 1991). As indicated earlier only a brief overview of major findings is presented here. Generally, findings related to naming speed deficits seem to be more ambiguous in comparison to those related to phonological deficits. Although some findings do seem to support an independent contribution of rapid naming to reading skills (e.g., Katzir, et al., 2006; Kirby et al., 2008; Plaza & Cohen, 2003; Powell, Stainthorp, Stuart, Garwood, & Quinlan, 2007; Wolf, et al., 2002), others do not (e.g., Cardoso-Martins & Pennington, 2004; Savage et al., 2005b; Wagner, Torgesen, & Rashotte, 1994). As Savage (2004) has suggested, because rapid naming tasks appear to be too difficult for pre-reading children, participants used in studies are already early readers. Thus establishing a causal relationship between rapid naming and subsequent reading might be difficult (Savage, 2004). Consequently, interpretations of findings might be difficult as rapid naming might simply reflect early reading ability, an ability that is likely to be an excellent predictor of later reading (Savage, 2004). Similarly, establishing a causal relationship through intervention studies may also be difficult since no one has attempted to explore how training in rapid naming might influence later fluency and automaticity in reading.

The causal role of rapid naming deficits in dyslexia has not been clearly substantiated using other research designs. There are some findings from cross-sectional designs that have shown poor readers to be worse on rapid naming tasks than average readers (e.g., Ackerman & Dykman, 1993; Bowers & Swanson, 1991; Korhonen, 1991; Korhonen, 1995; Wolf, Bally, & Morris, 1986). Nonetheless, interpretation of findings related to these studies is difficult due to the lack of a reading-level design. Overall, difficulties found in rapid naming deficit in individuals with dyslexia do not seem to be as clear-cut and unambiguous compared to those deficits found in phonological processes (Pennington, Cardoso-Martins, Green, & Lefly, 2001). For example, in some studies, poor readers have demonstrated poorer performance than their same-aged peers but their performance has been equal to that of the reading-age match control (e.g., Nicolson & Fawcett, 1994). Wolf et al. (2002) have argued that the inconsistencies related to findings on naming speed deficits may be because some studies include IQ-achievement non-discrepant poor readers. That is while IQ-achievement discrepant poor readers show speed naming deficits, non-discrepant poor readers do not (Wolf & Bowers, 1999; Wolf et al., 2002). However, other empirical evidence does not seem to consistently support the suggested pattern (e.g., Metz, Marx, Weber, & Scheider, 2003; Savage, 2007). As Vukovic and Siegel (2006) have suggested, difficulties in rapid naming seem to characterize only some individuals with dyslexia, namely those who are possibly severely impaired in reading and who also demonstrate deficits in phonological processing. There are few to none with deficits only in speeded naming who also have intact phonological abilities (Vukovic & Siegel, 2006). Overall, the claim that rapid naming has a unique causal contribution to reading in English over and beyond the contribution of phonological processes still remains to be confirmed (see Savage, 2004).

Cerebellar Deficit and Links to Phonological and Rapid Naming Difficulties

The notion of automaticity was extended later to encompass the more domain-general deficits in automatic processing of stimuli that underlie both phonological and speeded naming processes. According to Nicolson and Fawcett (1990), it is more appropriate to view dyslexia as a symptom of a more general and pervasive deficit in the acquisition of skill. Based on this view, children with dyslexia are seen as having difficulties becoming automatic in any learned skill that should become automatized upon extensive practice, including both motor and cognitive skills (e.g., reading, spelling, and phonological skills) (Nicolson and Fawcett, 1990). However, the authors have argued that the difficulty in automatizing tasks may go unnoticed because children with dyslexia have the ability to consciously compensate for these problems through coping strategies, such as trying harder (e.g., by allocating extra attentional resources to the task) or by using strategies to mask their deficit.

The automatization deficit was later incorporated within a neurological level hypothesis. At a neurological level, deficits in automatization have been attributed to abnormal cerebellar-vestibular areas of the brain (Reynolds et al., 2003). The proposed cerebellar deficit theory is suggested to combine both a learning and a neurological perspective, providing a unifying explanatory framework for the existing deficit theories of dyslexia (Fawcett et al., 2001; Nicolson & Fawcett, 1999; Nicolson et al., 2001). According to Nicolson et al. (2001), they have

> provided a plausible, albeit speculative, causal analysis that explains the difficulties in reading, writing, and spelling within a consistent and coherent developmental framework. Furthermore, two of the major alternative cognitive-level explanations of dyslexia, namely the phonological deficit hypothesis and the

> double deficit hypothesis, might be integrated naturally within this framework. (p. 511)

As Nicolson et al. (2001) have proposed, cerebellar deficit is “predicted to cause, by direct and indirect means, the ‘phonological core deficit’ that has proved such a fruitful explanatory framework for many aspects of dyslexia” (p. 510). Additionally, cerebellar deficit

> provides a natural explanation of the more recent “double deficit” hypothesis. This is based on the established difficulties that dyslexic children have on “rapid automatized naming” tasks, in which the child has to name as rapidly as possible a page full of common pictures or standard colours, and suggests that dyslexia is characterized by a deficit not only in phonological skills but also in naming speed (reflecting a lower speed of processing). Naming speed difficulties are precisely those predicted by the cerebellar deficit hypothesis, given its established role in speech, inner speech, and speeded processing. Consequently, all…cognitive level hypotheses appear to be directly consistent with, and indeed, subsumed by, the cerebellar deficit hypothesis. (Nicolson et al., 2001, pp. 510-511)

According to the cerebellar deficit hypothesis, a mild congenital cerebellar damage is suggested to be the mechanism underlying the pattern of difficulties displayed in children with dyslexia (e.g., Fawcett, 2002; Fawcett et al., 2001; Nicolson & Fawcett, 1999; Nicolson et al., 2001). The cerebellum (Latin for little brain), which is a fist-sized structure located at the lower back of the brain just above the brainstem, is a motor area involved in motor skills, automatization, and adaptive learning (Nicolson & Fawcett, 1999, 2000). More specifically, this structure is suggested to be involved in maintenance of posture, visually-guided movements, motor learning, and generation of smooth movements (Ito, 1993; Ito, 2002; Stoodley, Fawcett et al., 2006). These various functions are controlled via different regions in the cerebellum. For example, while tasks such as balancing or eye movements are controlled via vestibular regions, more complicated tasks, such as planning movements that are about to occur or evaluating sensory information for action, are regulated via the neocerebellum (Kingsley, 2000). The neocerebellum comprises the more lateral regions of the cerebellar hemispheres (Kingsley, 2000).

Nicolson and Fawcett have proposed that the pattern of difficulties in children with dyslexia points strongly to the involvement of the cerebellum (Nicolson & Fawcett, 1999, 2000; Nicolson et al., 2001). The reason for linking the cerebellum to difficulties experienced in children with dyslexia was initially based on the evidence that pointed to the involvement of the cerebellum in deficits displayed in balance, automatization, and motor skills (e.g., Holmes, 1917; Ito, 1984; Ito, 1993; Lang & Bastian, 2002). Later, additional evidence that pointed to the possibility of involvement of the cerebellum in cognitive skills like language and reading (e.g., Leiner, Leiner, & Dow, 1989; Leiner, Leiner, & Dow, 1993), including findings from neuroimaging studies and studies with cerebellar patients (Fiez, Petersen, Cheney, & Raichle, 1992; Silveri, Leggio, & Molinari, 1994), led to viewing a deficit in the cerebellum as the underlying mechanism for the wide range of difficulties displayed by individuals with dyslexia.

How Does the Cerebellar Deficit Lead to Literacy Problems?

Nicolson and Fawcett (1999) and Nicolson et al. (2001) have suggested that through a set of functional pathways, as illustrated in Figure 1 on page 26, the cerebellar deficit leads to a series of impairments that eventually cause deficits in writing, spelling, rapid naming, phonological difficulties and subsequent reading problems.

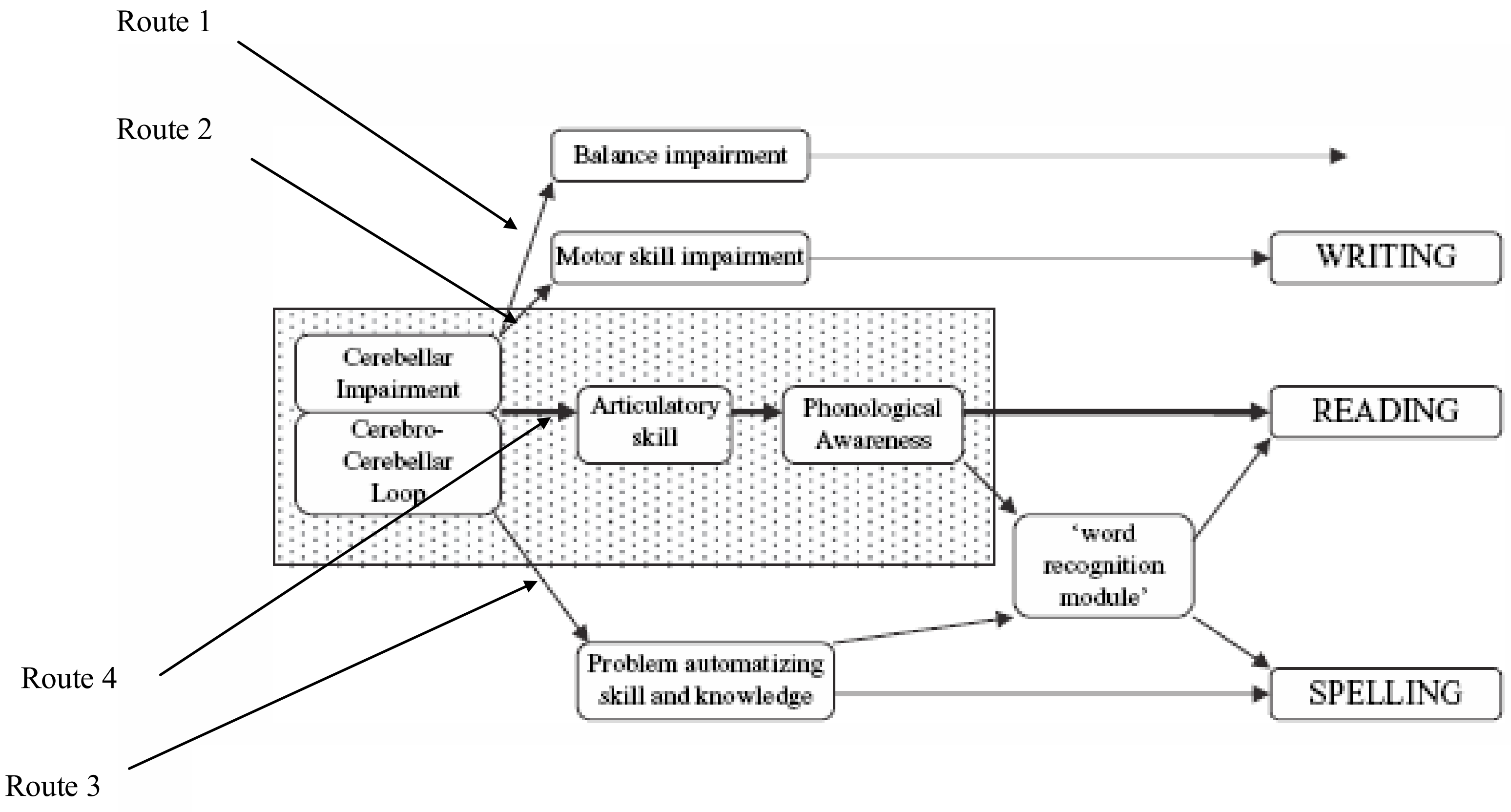


Figure 1. Nicolson and Fawcett's (1999) ontogenetic causal model of the cerebellar deficit hypothesis.

Based on this diagram, there seem to be four routes leading to problems in literacy. Of these routes, the fourth one is highlighted by the authors in this diagram as the most important. However, when each route and the related evidence is closely analyzed, it is difficult to separate the routes completely as they appear to be more or less intertwined. Keeping this in mind, these four routes and the evidence for them (when applicable) are critically evaluated.

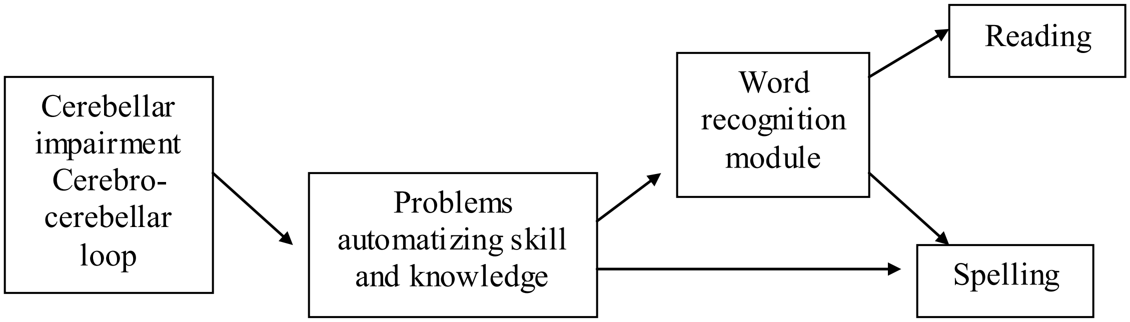
Route 1. Cerebellar Impairment ⟶ Balance Impairment ⟶

As seen in Figure 1, in the first route (taken directly from the figure) cerebellar impairment leads to deficits in balance. It is not clear from the figure what the balance impairment is leading to. According to the authors, it appears to be a general impairment independent of literacy skills. Nonetheless, it is noteworthy that balance tasks or more specifically tasks related to postural stability are included in the Dyslexia Early Screening Test and the Dyslexia Screening Test (Fawcett & Nicolson, 1995a; Fawcett & Nicolson, 1996; Nicolson & Fawcett, 1998). Balance skills are also addressed through remedial techniques, such as the Dyslexia Dyspraxia Attention Deficit Treatment (DDAT), now known as the Dore Achievement Centers (Reynolds & Nicolson, 2007) in Kenilworth, England, where physical exercises are prescribed for the treatment of dyslexia. A more detailed review and critique of the methodology and findings of cerebellar and balance intervention studies is offered in Irannejad and Savage (2009). Considering that postural stability is part of the dyslexia screening tests and dyslexia treatment, one might assume that impairments in balance illustrated in route 1 of Figure 1 may also lead to literacy problems, especially since deficits in balance have been investigated within routes 3 and 4 of Figure 1 (as will be reviewed). Nevertheless, this link does not seem to be readily clear merely by observing the figure. Consequently, the first route as it is depicted in the figure was not addressed in this study.

Route 2. Cerebellar Impairment ⟶ Motor Skill Impairment ⟶ Writing

A cerebellar deficit is proposed to provide a direct and natural explanation for difficulties in handwriting, a motor skill that requires precise timing and coordination of different muscle groups (Nicolson et al., 2001). Handwriting is suggested to be poor in children with dyslexia because of reduced motor skills that are directly caused by cerebellar deficit (Fawcett, 2002; Nicolson & Fawcett, 1999; Nicolson et al., 2001). At first glance, this route appears to be clearly separate from the other three routes. Interestingly, research investigating this specific route seems to be lacking. Searches of the literature base were conducted and no studies were found addressing this proposed link. This route was also not addressed in this study given that handwriting skills in children with dyslexia were not a focus of the present project. It should be noted that, once again, similar to balance skills, motor skills have also been investigated within routes 3 and 4 that seem to lead to literacy problems. Specific motor tasks are also part of the dyslexia screening tests as well as the DDAT dyslexia training program. Consequently, a link between cerebellar deficit, motor skills impairment and literacy problems might be inferred. Yet, this link does not seem to be well illustrated in Figure 1.

Route 3.



Route 3 seems to be intertwined with route 4. As Nicolson and Fawcett (1999) and Nicolson et al. (2001) have suggested, difficulties in spelling, reading and rapid naming (noting that rapid naming is not illustrated in their figure), arise from an articulatory-based phonological impairment (illustrated in route 4), in addition to deficits in automatization (illustrated in route 3). Nicolson and Fawcett (1999) explain that effective spelling requires both phonological skill and motor input. A decreased ability to automatize knowledge of spelling patterns as well as reduced capacity to attain implicit knowledge of orthographic regularities (or poor phonological awareness) is purported to lead to severe difficulties in spelling (Nicolson & Fawcett, 1999; Nicolson et al., 2001).

Prior to addressing the articulatory-based phonological impairment (i.e., route 4), deficits in automatization skills are addressed since they seem to be the central issue in route 3 that are suggested to be partly responsible for spelling, rapid naming and reading difficulties as suggested by the authors. Investigations of deficits in automaticity comprised the early foundation of the cerebellar deficit theory. As described earlier, according to Nicolson and Fawcett (1990), children with dyslexia have difficulty becoming automatic in any learned skill. However, this difficulty may be masked via conscious compensation used by children with dyslexia. Nicolson and Fawcett (1990) have used the “dual-task” paradigm to reveal these subtle automaticity deficits. Dual tasks were originally developed to study divided attention (Medland, Geffen, & McFarland, 2002). When completing a dual task, requirements of each task have to be held in working memory simultaneously. The assumption underlying the dual task paradigm is that the processing demands of the tasks are additive. Attentional resources should be allocated across both tasks. Hence, there should be adequate resources distributed to each task that has to be performed. However, because of the fact that coordination of both tasks requires additional resources, the resources distributed to complete each individual task decreases. The decrease in performance of dual tasks as compared to a single task is referred to as a dual task decrement (Medland et al., 2002). In the case of children with dyslexia, the dual task paradigm was used as way to reveal subtle deficits in automaticity that were thought to be masked via *conscious compensation* in children with dyslexia (Nicolson & Fawcett, 1990). In a dual-task paradigm, children with dyslexia were expected to complete two tasks simultaneously. The logic behind the paradigm was that, as a result of limited cognitive resources available to complete one process or behavior, the use of other resources will produce a deficit for other behaviors that require conscious control.

The primary task used in dual task studies by Nicolson and Fawcett (1990) and later Needle et al. (2006) included a task completely unrelated to reading such as balance tasks which included balancing on a beam, at times blindfolded, either on both feet, one foot, or when walking up and down, as well as heel-to-toe balancing (for adults). It is noteworthy to recall that balance and balance impairments investigated here were illustrated in route 1 addressed earlier which presumably led to deficits that were independent of reading and spelling. The secondary task was a novel task to divert attentional resources away from the primary task, such as choice reaction time tasks or backward counting. Using the dual task paradigm with children and adults with dyslexia who were compared to their respective control groups, Nicolson and Fawcett (1990) and Needle et al. (2006) reported finding no group differences in balance tasks when completed alone, but significant group differences in balance tasks in favor of the control groups when balancing was completed with another task.

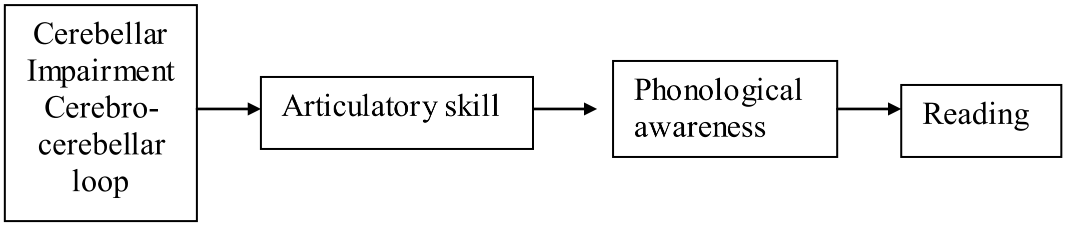
From a logical perspective, an important point made by Savage (2004) is why children with dyslexia would be capable of successfully masking their deficits in motor but not in literacy skills. In addition even if, as the automatization theory suggests, there are subtle motor deficits that can be revealed via a dual-task paradigm, it is hard to interpret these findings (e.g., Nicolson & Fawcett, 1990) in causal terms due to lack of reading-level design. Finally, the interpretation of findings for the dual task method is difficult and this paradigm has been largely criticized for a variety of reasons, such as lack of sufficient control to equate speed and accuracy of the primary and secondary tasks or the type of task included in this methodology. A more detailed review can be found in Savage (2004).

To avoid methodological complications and confounds involved in dual-task paradigms, other studies have investigated the quality of automatic skill learning without using dual-tasks. These studies have explored implicit motor sequence learning. Stoodley, Harrison, and Stein (2006), for example, reported implicit motor learning deficits in individuals with dyslexia in a serial reaction time task comparing 19 adults with dyslexia and 21 adults with no neurological or literacy problems. In this serial reaction time task, participants were asked to press the button (in a button box) corresponding to the number seen on a computer screen (i.e., 1, 2, 3, or 4). Numerical and positional cues were used to increase implicit learning and to shorten the time of the experiment. Stoodley, Harrison, et al. (2006) found that during randomly presented trials, the reaction time for adults with and without dyslexia were comparable. However, during the repeated sequence, the dyslexia group (i.e., 11 out of 19) showed less decrease in their reaction time as compared to the control group. In another study by Nicolson and Fawcett (2000), longer term learning was investigated in a group of teenagers with dyslexia (approximately 15-year-olds) and a control group who were matched to the group with dyslexia on chronological age and IQ. The study consisted of three phases. The first two phases took place over a period of 6 months. In the first phase, participants learned to navigate a computer PACMAN game in which four letter-key positions were learned in order to move a screen character. In the second phase, participants also learned an incompatible set of letter-key instructions. In the third phase which took place 1 year after the end of the second phase, longer term effects of pervious learning were investigated. The quality of performances was also tested using a dual-task design in which participants were asked to make a foot press while completing the computer game. According to the authors, while participants learned letter-key to movement associations quickly, the group with dyslexia demonstrated more errors. However, findings from longer-term follow-up studies did not seem to provide support for the automatization skills deficit, as the individuals with dyslexia did not differ in dual-task performance as compared to controls. Hence, as Savage (2004) has noted, this challenges the earlier findings that indicated individuals with dyslexia performed worse as compared to controls in computer tasks.

Evidence from other studies that have avoided the dual-task paradigm and examined implicit or motor sequence learning have further posed a challenge to the idea of skill automatization problems in dyslexia (Savage, 2004). Findings of these studies have, in fact, shown that individuals with dyslexia can learn equally as fast as average readers (Kelly, Griffiths, & Frith, 2002; Waber, et al., 2003). For example, Kelly et al. (2002) explored the role of implicit processing in their study using a choice reaction time learning task. University students with average reading skills were compared to those with dyslexia in their response times to a random sequence of pictures against response times to a complex but repeated spatial sequence. Kelly et al.’s (2002) findings indicated that students with dyslexia could learn the complex pattern at the same pace as the average readers did. Similar findings were also reported by Waber et al. (2003) who compared a large sample of children with learning problems to average reading controls.

In summary, considering (a) conceptual issues related to automaticity deficits, (b) methodological problems associated with dual-task paradigm used to investigate these deficits, and (c) lack of sufficient support for implicit learning in individuals with dyslexia, the conclusion that problems automatizing skill and knowledge can partly explain spelling, phonological awareness, rapid naming, and reading impairments remains questionable. Consequently, this route was not addressed in this study.

Route 4.



Based on Figure 1, which was presented on page 26, route 4 depicted above seems to be the central route leading to phonological awareness and subsequent reading problems. Nevertheless this becomes somewhat unclear since the explanations provided by the authors seem to involve other routes in the figure as well. According to Nicolson et al. (2001), "literacy difficulties arise from several routes. The central route is highlighted. If an infant has a cerebellar impairment, initial direct manifestations will be a mild motor difficulty – the infant might be slower to sit up and to walk."(p. 510). At this point, it is not readily clear what the connection between delayed abilities in sitting up or walking may be to later reading problems. It is possible that these are just highlighted as early signs of motor difficulties that are proposed to be present in individuals with dyslexia.

Nonetheless, as was noted earlier, motor skills have been a crucial part of the investigation of the cerebellar deficit theory and they are also part of the dyslexia screening battery and dyslexia training.

Further according to the authors, the cerebellum is also suggested to be a key structure in developing articulatory skill, a motor skill involving fine muscular control that requires timing and fluency, and is central to the development of language (Fawcett, 2002; Nicolson & Fawcett, 1999; Nicolson et al., 2001). According to Nicolson and Fawcett (1999),

> "our most complex motor skill, and that needing the finest control over muscular sequencing, is, in fact, that of articulation and co-articulation. Consequently, one would expect that the infant might be slower to start babbling and, later, talking….even after speech and walking emerge, one might expect that the skills would be less fluent, less 'dexterous', in infants with cerebellar impairment." (p. 170)

Slower babbling and talking are suggested to be direct effects of deficits in articulation resulted by a cerebellar impairment (Fawcett, 2002; Nicolson et al., 2001). What Nicolson and Fawcett (1999) and Nicolson et al. (2001) argue is that problems in articulation eventually lead to phonological core deficits and subsequent reading problems. This link, made by the authors, is based on the motor theory of speech and related evidence according to which development of phonological representations relies on speech articulation (e.g., Fowler, 1991; Liberman & Mattingly, 1985; Locke, 1983; Snowling & Hulme, 1994). According to this theory, phonetic gestures of the speaker are proposed to be represented in the brain as constant motor commands (Liberman & Mattingly, 1985, p.2). These gestural commands are suggested to be the elementary events of speech perception and production. In other words, these gestures are considered the physical manifestations that provide the basis for phonetic categories. Some examples of these gestures include lip rounding, tongue backing, or jaw raising gestures made to produce different speech sounds. These phonological units of speech are then recognized by the listener. This recognition is suggested to be based on inferences made by the listener about articulatory gestures of the speaker (Liberman & Mattingly, 1985). Considering this notion, Nicolson and Fawcett (1999) have argued that

> "very young children first perceive words as a loose bundle of articulated gestures, and in time the co-articulated gestures become grouped into the representations of phonemes…. If articulation is less fluent than normal, then it takes up more conscious resources, leaving fewer resources to process the ensuing sensory feedback. In particular, the processing of the auditory, phonemic structure of the words spoken may be less complete. There may, therefore, not be a natural sensitivity to onset, rhyme, and the phonemic structure of language – in short, one would expect early deficits in phonological awareness." (p. 170)

The above explanation has been proposed as one of the indirect effects of reduced articulation speed. Another indirect effect of less fluent articulation is suggested to be "reduced effective working memory as reflected in the phonological loop. This, in turn leads to difficulties in language acquisition" (Nicolson et al., 2001, p. 510). The phonological loop, as described by Baddeley, Gathercole, and Papagno (1998) and Baddeley, Thomson, and Buchanan (1975), is suggested to be specialized for retaining verbal information over short periods of time. In other words, the loop serves both as a storage to hold phonological codes, and as a rehearsal process to maintain weakened representations in the phonological storage.

In short, it seems that the central issue in route 4 is that the motor speech theory is used to explain the chain of events linking a cerebellar deficit to articulatory-based phonological difficulties and subsequent impairments in reading. In other words, as a consequence of a cerebellar impairment, articulation becomes less fluent. Because of this reduced fluency, the affected child needs to use more conscious effort and resources to process a given sensory feedback, such as sounds in a spoken word. Consequently, fewer resources are left to process the auditory and phonetic structure of the spoken word completely. There may, however, be an inherent logical problem in this explanation because according to Lieberman and Shankweiler (1991), "the basic phonological units that form the structure of words, the phonemes, are neither visual nor auditory. Instead, they are, to varying degrees, abstractly linguistic*"*(p. 12).

As Lieberman and Shankweiler (1991) explain, phonemes are abstract categories of language and while they are represented and expressed via sounds, they are not sounds themselves. Moreover, when a word is spoken, the phonological units in that word are not produced one at a time (i.e., the word is not spelled out). According to the authors, the speaker uses "co-articulation" of speech sounds, whereby the speech is produced at a speed that can be easily understood. For example, as Lieberman and Shankweiler note, the word "*bag"* which includes three phonemes and 3 corresponding letters in print is not spelled out in three sounds "*b*", "*a*", "*g*" when spoken, but rather in one pulse of sound whereby all three elements are merged into one sound "*bag*". Understanding and recovering the phonological structure of merged sounds by the listener is proposed to be via processes that are suggested to "be built into that aspect of our biology that makes us capable of language" (Lieberman & Shankweiler, 1991, p. 6). Considering the "co-articulation" phenomenon and the abstractness of phonemes, the phonetic structure of a spoken word cannot be extracted or understood readily by the listener merely by listening to merged sounds and attending to articulatory gestures, as assumed in the cerebellar deficit theory. In summary, a purely articulatory-based phonological awareness difficulty is unlikely in light of the fact that phonemes are neither “visual” nor “auditory” and the articulation of a word itself and the motor gestures used to convey a sound or word are distal with the underlying phonological structure and are rather inferred by the listener.

In addition to this possible conceptual problem with the articulatory-based link made between a cerebellar deficit and phonological impairment, the empirical evidence supporting this link is also not very consistent. On the one hand, at a general level, there has been some evidence supporting a possible association between the cerebellum, articulation, and speech perception. For example, Schahmann and Sherman (1998) have found that the cerebellum might play a role in impairments observed in naming and fluency of speech. Ackermann, Graber, Hertrich, and Daum (1997) also considered the cerebellum to be an “internal clock” that is responsible for processing “durational parameters of the perceived acoustic speech signal” (e.g., voice onset time). They used a series of disyllabic stimuli in their study which differed in durational parameters and found a link between cerebellar pathology and difficulties in categorical speech perception, more specifically in perception of phoneme boundaries. Ackerman, Graber, Hertrich, and Daum’s findings (1999) have also indicated a possible link between cerebellar dysfunction and difficulties in discriminating time intervals that are equally long as acoustic speech segments.

On the other hand, there are some findings indicating that cerebellar pathology may not interfere with language, speech perception, and practice effects for speech or oral movements (Fiez et al., 1992; Ivry & Gopal, 1993; Schulz, Dingwall, & Ludlow, 1999).

Additionally, there is evidence for normal phonological development, as well as normal reading and writing skills, despite severe cases of speech disorders (e.g., dysarthria, apraxia) (e.g., Cossu, 2003). Ramus, Pidgeon, et al. (2003) have also pointed to evidence indicating normal performance on some phonological tasks in children with motor speech difficulty (i.e., inborn dysarthria).

In regard to a reduced articulation speed in dyslexia, supportive findings have been reported by Fawcett and Nicolson (2002). Here two groups of children with dyslexia (ages 13 and 16) were asked to articulate a single articulatory gesture (e.g., “p”) or the sequence “putuku” repeatedly and as quickly as possible. They found that as compared to the two control groups (matched to the dyslexia groups for age and IQ), children with dyslexia had significant problems in gesture planning and single production of single articulatory gestures. Similar findings were reported by Kasselimis, Margarity, and Vlachos (2008) for a group of Greek children with dyslexia who were asked to recall over learned sequences (i.e., days of the week, 12 months, and Greek national anthem). None of these studies used a reading-level design. However, evidence supporting a link between articulation speed and problems in reading is controversial. For example, evidence that has analyzed articulation and articulation pause times during rapid automatized naming tasks has shown that articulation rates were not directly related to reading (Georgiou, Parrila, Kirby, & Stephenson, 2008; Neuhaus, Foorman, Francis, & Carlson, 2001; Neuhaus, Carlson, Jeng, Post, & Swank, 2001; Neuhaus & Swank, 2002).

Generally, researchers are more cautious in regard to the route through which the cerebellar deficit is believed to lead to articulation, phonological and subsequent language or reading difficulties. As Ramus, Rosen, et al. (2003) have noted, the link made between the cerebellum and articulation is based on outdated views. The motor speech theory, which has been used to explain the causal chain of events linking cerebellar deficit to articulatory-based phonological difficulties and subsequent reading impairment, is not consistently supported by empirical studies. Altogether, the presence of normal phonological, reading, and writing skills despite speech disorders (e.g., Cossu, 2003), along with evidence that has failed to show an association between item articulation rate and reading (e.g., Georgiou et al., 2008) or cerebellar pathology and articulation (e.g., Schulz et al., 1999), have cast some doubt on whether or not a cerebellar deficit should be accepted as the sole underlying mechanism for reading difficulties. Considering the inconsistent findings and more importantly due to the possible logical problem with an articulatory-based phonological difficulty explained earlier, a link between cerebellar impairment, articulation, and phonological awareness was not investigated in this study.

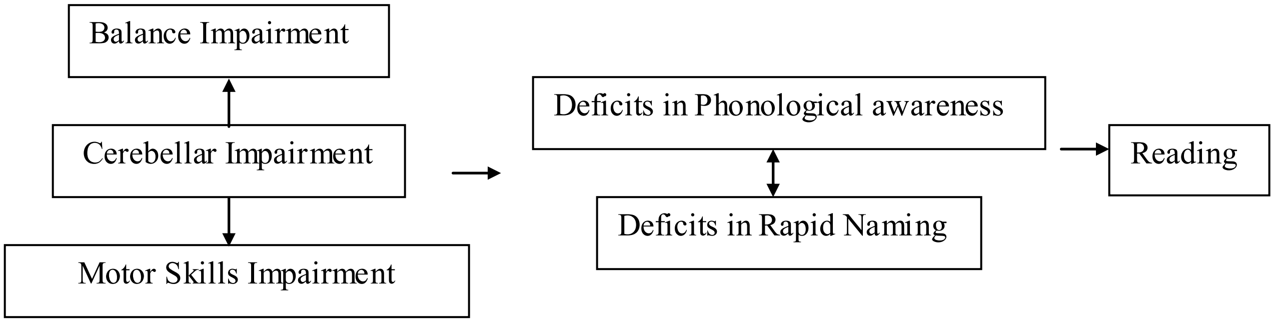
Conclusion

In summary, as was illustrated in Nicolson and Fawcett's (1999) figure on page 26, there were four seemingly intertwined routes that appeared to lead to problems in literacy. In the first route, cerebellar impairment was illustrated to lead to deficits in balance. While not readily clear from the figure, balance deficits presumably lead to impairments independent of literacy skills. In the second route, a cerebellar deficit was proposed to provide a direct explanation for difficulties in handwriting (Nicolson et al., 2001), but research investigating this specific route seems to be lacking. In route 3 and 4, deficits in automatization and articulation fluency were proposed to lead to difficulties in spelling, rapid naming, phonological awareness, and eventually reading. As discussed, findings related to deficits in automatizing skill and knowledge, illustrated in route 3, have been hard to interpret and criticized due to conceptual issues as well as methodological problems related to the dual-task paradigm that was used to investigate deficits in automaticity. Additionally, as was noted previously, the articulatory-based link made between cerebellar deficit, phonological awareness, and reading may also be questioned due to logical problems as well as inconsistencies across empirical findings.

Consequently, the question is which route leads to problems in literacy? There seemed to have been a lot of confusion regarding this matter considering that explanations provided by the authors of the theory are not always in agreement. For example, Fawcett et al. (1996) have suggested an impairment across the entire cerebellum and thus across a range of cognitive processes, as opposed to a deficit that is specific to a single region that is implicated in dyslexia. According to Nicolson et al. (2001), “the pattern of difficulties in cognitive, information processing and motor skills is predicted by the cerebellar deficit hypothesis” (p. 508). In contrast to this claim, Nicolson and Fawcett (2006) suggested that it would be incorrect to assume that cerebellar deficit is associated only with balance and/or motor difficulties because the “cerebellar deficit hypothesis…claims only that the language-related regions of the cerebellum are affected in dyslexia” (p. 261).

If this is the case, then what does the ontogenetic causal model of the cerebellar deficit hypothesis illustrated in Nicolson and Fawcett’s (1999) as well as Nicolson et al.’s (2001) diagram mean? If according to that diagram, neither route 1 (i.e., impairment in balance) nor route 2 (i.e., motor skill impairments) leads to problems in literacy, then why have deficits in balance (postural stability) and motor skills been immensely explored by Nicolson and Fawcett (e.g., Fawcett & Nicolson, 1995b, 1999; Fawcett et al., 1996; Fawcett et al., 2001; Needle et al., 2006; Nicolson & Fawcett, 1990; Nicolson & Fawcett, 1994; Stoodley, Fawcett et al., 2006) as part of their investigation of the cerebellar deficit theory of dyslexia? In fact, Nicolson and Fawcett (2006) have also suggested that it is the extended difficulties in static cerebellar tasks (including balance and muscle tone) that have the power to discriminate between two groups of poor readers (i.e., discrepancy-based dyslexia versus garden variety poor readers). However, if according to Nicolson and Fawcett (2006), only the language-related regions of the cerebellum are affected in dyslexia, then the degree of usefulness of the static cerebellar tests and motor tasks as diagnostic measures may be questioned, especially since Nicolson and Fawcett's (2006) claim is more compatible with a phonological deficit model (or language-based deficits). Interestingly though, as was indicated earlier, balance and specific motor tasks are part of the battery included in the Dyslexia Early Screening Test and the Dyslexia Screening Test (Fawcett & Nicolson, 1995a; Fawcett & Nicolson, 1996). Deficits in postural stability and motor skills are also addressed through remedial techniques in the Dore Achievement Centers (Reynolds & Nicolson, 2007) via prescribed physical exercises for treatment of dyslexia.

In light of these controversies, the important question is which route should be addressed in the present study? Considering that (a) balance and motor skills have been an important part of Nicolson and Fawcett's investigation of the cerebellar deficit theory, (b) balance and motor tasks are a part of the author's dyslexia screening tests and dyslexia training, and finally (c) many other researchers who have investigated the cerebellar deficit theory of dyslexia have included tasks related to postural stability and motor skills in their studies, then a link between cerebellar deficit, motor and balance impairment, phonological awareness, rapid naming and reading may be assumed. Hence, the route investigated in this project is as follows:



In the next section, the early behavioral research investigating a link between cerebellar deficit, motor functioning, balance, and dyslexia is critically evaluated. Four major research issues are then identified as a result of this evaluation.

Cerebellar Deficit, Motor Functioning, Balance, and Dyslexia

To introduce a brief history, a possible role of vestibular dysfunction in reading disability was first proposed by Ayres (1973, 1978). Among the other early research studies pointing to a possible cerebellar deficit in dyslexia, a study by Frank and Levinson (1973) has been much-cited. Here, 17 children that were sampled from 115 slow learners showed evidence of a cerebellar-vestibular dysfunction. Later, Levinson (1988) reported that of almost 4,000 participants with a learning disability, not only did 95.5 percent have a history of “reading symptoms,” but 99.5 percent of the sample also showed one or more neurological or electronystagmographic signs of cerebellar-vestibular dysfunction. The reliability of Levinson’s findings have, however, been questioned due to the use of an unclear criteria for defining dyslexia and the lack of a control group (Beaton, 2002).

In later investigations, Nicolson et al. (1995) attempted to reveal time estimation deficits in children with dyslexia, which were originally found by Ivry and Keele (1989) in cerebellar patients. Time estimation, which is a non-motor task, required participants to decide between both the length and loudness of two auditorily presented intervals. Nicolson et al.’s (1995) results indicated that similar to cerebellar patients in Ivry and Keeles’ study, children with dyslexia, as compared to reading-age controls in their study, displayed deficits in temporal, but not in loudness estimation. While Nicolson et al.’s (1995) findings revealed similar time estimation deficits between children with dyslexia and cerebellar patients, some researchers have suggested that this particular deficit and its relation to a cerebellar deficit in individuals with dyslexia may have been overrated (Beaton, 2002). The cerebellum may not necessarily be the underlying cause of a time estimation deficit, as temporal discrimination in normal individuals has also been related to the superior left cerebral hemisphere, rather than the cerebellum (e.g., Nicholls, 1996). Moreover, performance on a time estimation task is not necessarily a direct index of cerebellar function, and making a direct link between problems in this task and in reading impairment is difficult (Fawcett et al., 1996).

Following the research on skill automatization (i.e., dual task studies which were briefly described) and time estimation, a series of studies were conducted which were mainly related to the cerebellar deficit theory. These studies are discussed in detail since they pertain directly to the present research.

In this body of research, Fawcett et al. (1996) and Fawcett and Nicolson (1999) investigated whether individuals with dyslexia as compared to normal readers displayed noticeable deficits in traditional signs related to cerebellar impairment similar to those shown by patients with gross damage to their cerebellum (Fawcett & Nicolson, 1999; Fawcett et al., 1996). As Holmes (1939) has described, focal or neurodegenerative damage to the cerebellar cortex and its nuclei usually leads to profound motor abnormalities which are seen as fundamental components of cerebellar dysfunction, and can be explained in terms of four core deficits.

The first motor abnormality is hypotonia, or reduced muscle tone, in which an individual shows reduced resistance during an external disturbance, or passive movement of an extremity. As a result, limbs can be easily displaced with little force, and arm excursion during walking is increased. The second motor problem is asthenia, or fatigability, and it is described as a condition in which the individual becomes unusually tired after repeated movement. The third motor abnormality includes disorders of associated movements, in which the individual displays problems such as unconscious swinging of the arms during normal locomotion. Finally, the last motor problem is ataxia, or disorders of voluntary movement in which the individual has problems with accuracy and organization of voluntary muscle actions, resulting in uncoordinated movements involving the trunk (e.g., disturbance in posture, gait) and limbs, speech, or eye movements (e.g., Diener, et al., 1992; Diener, Hore, Ivry, & Dichgans, 1993; Holmes, 1917, 1939; Palliyath, Hallett, Thomas, & Lebiedowska, 1998).

In their studies that included children with dyslexia, Fawcett and Nicolson (1999) and Fawcett et al. (1996) used clinical cerebellar tests that were described in a classic text by Dow and Morozzi (1958) to replicate problems with muscle tone, posture, gait or movements of the extremities seen in cerebellar patients with gross damage to their cerebellum. Following other studies, these tests will be referred to as cerebellar tests or measures in this thesis. Nonetheless, it is acknowledged that these tests do not directly measure cerebellar functioning but are rather indirect measures of possible behavioral manifestations of cerebellar impairment. Three different categories of tests were used either fully or partially in these studies.

In the first category, two tests were included to assess ability to maintain posture and muscle tone in standing position (i.e., balance time) and in response to active displacement of station (i.e., postural stability in response to a gentle push to the participant's back). The second category included seven tests to assess hypotonia of the upper limbs in the standing and sitting position, and in response to active or passive displacement of the limbs (e.g., arm displacement, static tremor, arm shake, muscle tone). Finally, in the third category, five tests were used to assess ability to initiate and maintain a complex voluntary movement (e.g., repeated finger to finger pointing or toe tap speed).

Using all of the aforementioned cerebellar tests, and including participants from their previous panel of research, Fawcett et al. (1996) compared three groups of children with dyslexia (i.e., 12 children with mean age of 10.7, nine with mean age of 14.4, eight with mean age of 18.6) with three groups of chronologically-age match control groups with average reading skills. The groups were also matched on intellectual level. The number of children included in the control groups were 8, 11, and 7 respectively. Altogether the study included 55 participants. Dyslexia was identified using a discrepancy-based definition and the groups with dyslexia were reported to have reading ages of at least 18 months below their chronological age. It was also reported in the study that two of the age-matched control groups (i.e., the groups with mean age of 14.6 and 10.9) were used as reading-age match controls for the two older dyslexia groups (i.e., the group with mean age of 18.6 and 14.4). Finally, the control group with mean age of 10.9 was also used as a reading-age match control for the dyslexia group with mean age of 18.6.

According to Fawcett et al. (1996), none of the children in their study showed evidence for Attention Deficit Hyperactivity Disorder (ADHD) as measured by the DSM-

IIIR scales (American Psychiatric Association [APA], 1987). It was also reported that children with dyslexia had already participated in a range of other experiments conducted by Fawcett and Nicolson (e.g., Fawcett & Nicolson, 1995b, 1995c; Nicolson & Fawcett, 1994; Nicolson et al., 1995) and they were initially recruited through a local Dyslexia Institute or the local branch of the British Dyslexia Association (Fawcett et al., 1996).

Fawcett et al. (1996) reported significant effects of age for some of the motor and cerebellar tasks including balance time and muscle tone. Based on chronologically-age match comparisons using effect size ratios, Fawcett et al. (1996) reported deficits for all cerebellar tasks in children with dyslexia. Reading-age match comparisons using effect size ratios revealed deficits for the majority of tasks (i.e., 11 out of 14). For most tasks, the magnitude of impairment was reported to be greater than that for impairment in reading. Due to higher incidence rates for impairments in children with dyslexia in this study, as opposed to those cited for cerebellar patients, it was concluded that deficits found in children with dyslexia may represent an impairment across the entire cerebellum, as opposed to a deficit that is specific to a single region (Fawcett & Nicolson, 1999; Fawcett et al., 1996). Using three tasks that were suggested to be representative of the three category of tests described earlier, as well as additional reading tests, Fawcett and Nicolson (1999) attempted to replicate their findings with a different and larger sample that included 126 children with dyslexia and normal readers who were divided and matched in four different age groups (8-9 year olds, 10-11 year olds, 12-13 year olds, and 14-16 year olds). The children with dyslexia were recruited from special schools for dyslexia or local schools with special units for dyslexia. For the purposes of this study (i.e., Fawcett & Nicolson, 1999), the children were screened to ensure they met a discrepancy-based definition for dyslexia. Children in the control group were chosen by their teachers on the criterion that their reading age was not below average. For the purposes of the study, they were also administered a short-form of an intelligence test and a test of reading and spelling.

Significantly worse performance was reported in children with dyslexia, as compared to chronologically-age match controls, in all three tests of cerebellar function (i.e., in postural stability, arm shake, and toe-tapping speed). Performances were reported to be exceptionally poor in all four dyslexia groups on postural stability and limb shake. Additionally, children with dyslexia, as compared to chronologically-age match controls, displayed more difficulties on tests of segmentation and nonsense word repetition. The same pattern of results was also reported for a subgroup of children with dyslexia and controls matched for IQ (Fawcett & Nicolson, 1999).

A Critique of the Cerebellar Deficit Theory of Dyslexia

Evidence related to cerebellar signs in individuals with dyslexia has been reported in many studies. Findings, nevertheless, do not seem to be consistent across studies that have investigated the cerebellar deficit theory. According to Nicolson and Fawcett (2006), however, the extended difficulties in static cerebellar tasks (including balance and muscle tone) can address the cause of heterogeneity of data as they have the diagnostic power to discriminate between two groups of poor readers (i.e., discrepancy-based dyslexia versus garden variety poor readers). That is, while the incidence of phonological, speed naming, and motor deficits are high in both types of poor readers, only the children classified using discrepancy-based dyslexia demonstrate deficits on the static cerebellar tasks, namely balance and muscle tone. This claim was based on a study in which balance and muscle tone deficits were reported to be unique to the individuals identified with dyslexia using the discrepancy-based definition (Fawcett et al., 2001).

Apart from aforementioned critics related to the discrepancy-based definition of dyslexia, the above claim has not been replicated by other researchers. Savage (2007), for instance, included two groups of children, one consisting of 25 children with a formal diagnosis of developmental dyslexia and another consisting of 18 children with a formal diagnosis of intellectual disabilities (ages in sample ranging from 11 to 14). The two groups only differed on verbal and nonverbal cognitive abilities, thus providing a good opportunity to test for the Fawcett et al.'s (2001) claim with regard to cerebellar deficit. Savage's (2007) findings revealed no significant differences between the two groups in spelling, word reading, phonological processing or basic verbal response speed. Furthermore, no difference was found on the measure of postural stability as assessed by the Dyslexia Screening Test (DST) between the two groups. Indeed, children with developmental dyslexia showed small non-significant advantages on postural stability and even larger significant advantages on the motor task of bead threading. Additionally, no deficit on postural stability measure was revealed, when postural stability mean scores obtained by the two groups of children who participated in this study were compared to those of typical children in the DST test standardized sample. In fact, based on the DST criterions, both groups were identified as being at "no risk" on this measure.

Considering the above evidence, Nicolson and Fawcett's (2006) claim that the static cerebellar tests can explain the heterogeneity of results for the cerebellar deficit does not seem to be well supported. However, as will be reviewed, in evaluating the evidence the inconsistency across findings appears to be due to four major issues. One issue seems to be a lack of control for possible confounding factors. This, for example, seemed to be the case in the study conducted by Fawcett et al. (1996) described earlier. While Fawcett et al. (1996) did seem to screen their sample for attentional difficulties to some extent, they did not report controlling for it statistically. Attention was also not controlled for in some of their other studies (e.g., Fawcett & Nicolson, 1995c; Fawcett & Nicolson, 1999; Nicolson & Fawcett, 1990). As will be reviewed, controlling for confounding factors such as presence of attention problems is important especially because difficulties in motor and cerebellar tasks may be associated with attention problems (Rochelle & Talcott, 2006).

The second issue that seems to explain inconsistent findings across the cerebellar deficit studies may be the method by which motor and cerebellar tasks have been assessed. For example, in the early body of research described previously (e.g., Fawcett & Nicolson, 1999; Fawcett & Nicolson, 1995c; Fawcett et al., 1996), cerebellar deficits were investigated using Likert scales. However, a problem with this type of scale is that they are not sensitive enough, and they provide subjective estimation by observers which is expressed in qualitative rather than quantitative units (Shipley & Harley, 1971). As will be reviewed, some studies that have incorporated more sensitive measures, such as accelerometer sensors to measure postural stability, seem to provide different results, some of which are not supportive of the cerebellar deficit theory.

The third issue for incongruent findings across cerebellar deficit studies appears to be a lack of a reading-level design. As will be discussed, there is a scarcity of research for cerebellar deficit theory of dyslexia that includes a reading-level design, a design that has been regarded as crucial for making causal interpretations (Goswami, 2006; Goswami & Bryant, 1989). The one study by Savage et al. (2005a), which will be reviewed, that has used this design has not provided support for a cerebellar deficit account of dyslexia.

Finally, the fourth issue that may explain heterogeneous data for cerebellar deficit theory seems to be a lack of homogeneity in samples involved in studies. In many studies on dyslexia that will be reviewed, the method of sampling has led to a lack of homogeneity. This in turn can affect the interpretation of findings since any differences found between dyslexia and control groups may be due to differences inherent to sampling. In the next four sections, each of the aforementioned issues along with relevant evidence is considered in more detail.

(1) Lack of Control For Confounding Factors

One reason for the discrepancies in findings related to cerebellar deficit may be a lack of control for possible confounding factors such as presence of attention difficulties. While some studies in the literature have found poor balance in children with dyslexia, results seem to suggest that poor balance during dual tasks as well as poor postural stability may only be specific to the group of children with dyslexia who also exhibit attention difficulties (Savage, 2004). Wimmer et al. (1999), for example, compared a group of German children with dyslexia with their chronologically-age match controls on performance on balance tasks used by Nicolson and Fawcett (1990), as well as naming speed and phonological skills. In this study, parents of the children were interviewed and asked to complete the Conners Attention Rating Questionnaire, in order to rule out the possibility that performance on motor tasks might be affected by difficulties in attention. Additionally, teachers were asked to rate children on attention skills. Attention scores were then controlled statistically and used as covariate. Using this procedure, findings indicated that differences between poor and good readers were solely explained by difficulties in naming speed and phonological awareness. Similar results have also been reported by other researchers (e.g., Denckla, Rudel, Chapman, & Krieger, 1985; Raberger & Wimmer, 1999; Raberger & Wimmer, 2003; Ramus, Pidgeon et al., 2003).

Rochelle and Talcott (2006) have also sought to address inconsistencies found in studies on balance impairment (including difficulties in dual-task balance as well as in postural stability) in dyslexia in their meta-analysis. They investigated the magnitude and consistency of balance difficulties in population with dyslexia. They also addressed sampling and stimulus characteristics that seemed to significantly modulate the magnitude of effects that were obtained for balance measures across studies. Rochelle and Talcott (2006) reviewed all studies which were published between 1985 and 2004 and compared participants with and without dyslexia on behavioral measures of balance.

Their final sample included 17 published studies. Fifteen of these 17 studies were reported to include adequate statistical information for their inclusion in the meta-analysis. They indicated that the potential variables which seemed to modulate the magnitude of the effect size found between groups included age of participants, proportion of the sample for which a diagnosis of ADHD could be excluded, intelligence and reading scores, type of balance measures administered (i.e., single versus dual task paradigms). Finally, screening for co-occurring Developmental Coordination Disorder (DCD) was also reported to be among the modulating factors. However, this latter effect was only present in one of the studies in the sample. Rochelle and Talcott could not perform an evaluation of phonological skills due to a considerable non-uniformity of measures used across the studies.

Considering all these factors, Rochelle and Talcott (2006) found that the relationship between balance deficits (i.e., deficits in dual-tasks balance and postural stability) and dyslexia seemed to be indirect. According to Rochelle and Talcott, the link between impaired balance and dyslexia was most strongly influenced by variables other than reading skills. More specifically, they reported that studies frequently failed to discriminate adequately between co-occurring developmental disorders. In fact, according to the authors, a strong predictor of balance effects across studies was the proportion of participants who were screened for ADHD symptoms. Unfortunately, they could not draw firm conclusions about the impact of including participants with co-occurring DCD given that only one of the studies had screened for this disorder. Nonetheless, Rochelle and Talcott predicted that this disorder also strongly modulated balance effect size, especially because of reported co-occurrences of dyslexia with both ADHD and DCD.

(2) Method by Which Motor and Cerebellar Tasks Have Been Assessed

Another factor that may explain incongruent findings across cerebellar deficit studies is possibly related to the method by which motor and cerebellar tasks have been assessed. Findings from studies that have used Likert scales to assess motor and cerebellar tasks (e.g., Fawcett & Nicolson, 1999; Fawcett & Nicolson, 1995c; Fawcett et al., 1996) may be less reliable since these scales provide qualitative data and more subjective estimation by observers (Shipley & Harley, 1971). To avoid problems related to such scales, a few studies have attempted to explore cerebellar deficits using measures that seem to be less subjective and more sensitive while also providing continuous quantitative data. In an interesting study by Brown et al. (1985), the role of vestibular dysfunction in dyslexia was investigated by comparing postural stability of 15 boys with dyslexia to 23 boys who were reported to be normal readers. All participants in the study came from Caucasian middle class families. They were between the ages 10 to 12, right handed with normal vision and hearing and no history of emotional problems, hyperactivity, or birth stress. All participants had a full scale IQ above 88, as measured by the Wechsler Intelligence Scale for Children-Revised and difference between their verbal and nonverbal skills was reported to be less than 30 points. Reading in all children was assessed via a silent reading comprehension as well as an oral reading test. Diagnosis of dyslexia was based on discrepancies between reading and mental age.

A “dual axis torsion rod system” was used to measure the amount of sway in participants (Shipley & Harley, 1971) in postural stability tasks. This system consists of four components. It includes a platform assembly upon which participants stand and an electromechanical system which is used to sense and provide a read out in arbitrary units to indicate the amount of sway. The system also includes a vibrator which is attached to the platform. This vibrator can impart a horizontal, non-rotating, circular motion to the platform. In other word, it allows the platform a limited angular displacement from the horizontal of approximately one to two degrees in any direction when the person’s center of gravity shifts. Finally, the system consists of an automatic timer that can be set for any time period for a given test (Shipley & Harley, 1971). Brown et al. (1985) adapted this system for their purposes in order to obtain a continuous output proportional to platform displacement.

Postural stability in the sample was measured with participants’ feet either side by side or heel-to-toe and with their eyes open and closed in both conditions. They had to stand barefoot on the platform in each of these positions and maintain their balance. This was first completed for a test period and then for the actual study. Each position was adopted for nearly 12 seconds and data obtained by the sensor was analyzed in the computer. Brown et al. (1985) found no differences between the groups under any of the conditions tested. They also reported no difference in how participants used visual information to maintain their posture.

In a similar vein, Needle et al. (2006) also found no overall differences in heel-to-toe balance measures (i.e., postural stability) obtained from 17 adults with dyslexia and

20 adult normal readers when they used a motion tracking system to measure degree of wobble. For the balancing task, participants were required to maintain a heel-to-toe position for a minute at a time, moving as little as possible (i.e., standing with one foot in front of the other with the heel touching the toe in a straight line, with arms outstretched). To measure the degree of wobble, a motion tracking system was used that tracked movement in three dimensions at a rate of 120 HZ. The recorded data was then analyzed in the computer. The sensors were attached to the forefinger of each outstretched hand. Needle et al. (2006) reported no differences between the group with dyslexia and control group in the heel-to-toe balancing condition.

In a more recent study, in which a triaxial accelerometer sensor was used to investigate both postural stability and gait, Moe-Nilssen et al. (2003) did find some support for cerebellar deficit theory. The sample included 22 children with dyslexia and 18 normal readers from Norwegian schools (ages 10 to 12). Accelerometer sensors are devices used for measuring acceleration, vibration, and gravity induced reaction forces. Their sensitivity can vary depending on their sampling rate ability. Depending on the sensor used, the accelerations can be measured in two-dimensional or three-dimensional planes. In instances where body sway or wobble is being measured, usually portable sensors that can be attached to the body (e.g., via a Velcro belt) are used. The amount of sway is then measured via the sensor and the data gathered can be analyzed either directly in the recorder or in an interface software program in the computer.

In Moe-Nilssen et al.'s (2003) study, the accelerometer was attached to participants' backs using a fixation belt and data was recorded at 128 HZ. Moe-Nilssen et al.'s study also differed in that they did not push participants in the back as in previous studies (e.g., Fawcett et al., 1996; Fawcett & Nicolson, 1999) to challenge balance.

Participants in this study had to stand with their eyes closed and their feet 0.1 m apart opposed by a constant horizontal drag, equivalent to 5% of their body mass, from a rope which was fastened to a weight. To challenge participants' balance, the weight was lifted without their knowledge. This caused them to fall backwards until a reactive movement was made to regain balance. Moe-Nilssen et al. (2003) measured the difference in time between the release of the drag and the response of the subject using two synchronised accelerometers (i.e., one attached to the participant's back and another to the weight that was fastened to the rope).

Postural stability (i.e., degree of sway) was also measured in quiet standing positions with feet together and eyes both closed and open. The surface that participants stood on was either firm or thick, consisting of a compliant mat or compliant pillow. Degree of sway was also assessed in the quiet heel-to-toe standing position with eyes open on a firm surface. In all positions where participants' eyes were open, they were asked to fixate on a visual target that was positioned on the wall at their eye level. In all positions, participants had to stand as still as possible for 30 seconds. Moe-Nilssen et al. (2003) also assessed participants' gait in their study. Participants had to walk barefoot on both a flat and an uneven surface. Walking speed, length of step, and number of steps per minute were used as measures for four different walking conditions that included walking slow, preferred, fast, or very fast.

Moe-Nilssen et al. (2003) reported group differences in mean walking speed which seemed to be faster on both flat and uneven surfaces for the control group as compared to the group with dyslexia. With the effect of gender controlled, very fast walking on an uneven surface also seemed to correctly discriminate 77.5% of participants in their respective groups. Among the balance tasks, children with dyslexia were reported to show impaired balance as compared to controls only on tests of undisturbed balance with eyes open and fixated on a visual target on the wall, but not when eyes were closed. Moe-Nilssen et al. speculated that this finding might be a result of the fact that children with dyslexia were not able to take adequate advantage of the visual cue positioned on the wall, possibly because of some impaired readers' inability to maintain steady fixation.

Although Moe-Nilssen et al.'s (2003) findings are interesting especially because of their use of more sensitive measures, there are some factors related to their sample and methodology which might affect interpretation of their results even if some group differences have been found. For example, the children in Moe-Nilssen et al.'s study were randomly selected from a teacher selected sample which included the top and bottom 5% readers from 200 pupils. Once the sample for study was selected, the poor readers were administered different reading as well as intelligence measures. This procedure, however, was not repeated for normal readers due to time constraints in the study. Nevertheless, lack of information on reading skills for normal readers, who were described "as the top 5% readers" can make the interpretation of results difficult. Furthermore, Moe-Nilssen et al. used the out-dated and empirically criticized discrepancy-based definition to identify individuals with dyslexia in their sample. As they reported, based on the literacy achievement and cognitive measures administered, 84% of the poor readers in their sample satisfied a discrepancy-based definition of dyslexia. Finally, an important missing element in the methodology used in Moe-Nilssen et al.'s study, as well as many other studies reviewed in this thesis which have investigated the cerebellar deficit, is the lack of a reading-level design. This issue will be addressed separately in the next few sections.

(3) Lack of a Reading-Level Design

One of the important issues in research on dyslexia is the degree to which the evidence derived from these studies can truly establish causes underlying this disorder, especially considering its complexity. Some explorations of methodology in reading research have identified that convergent findings from a combination of both longitudinal and intervention methods might be best-place for researchers to identify causes (e.g., Bradley & Bryant, 1983). As Goswami and Bryant (1989) have pointed out, it is the combination of these two types of evidence that is required to arrive at firm causal conclusions. Nonetheless, the expense and difficulty associated with constructing both longitudinal and intervention studies alongside the practical difficulties of identifying the influence of the candidate variable to the exclusion of other related variables often means that research of this kind is undertaken on candidate theories that have reached a certain degree of intellectual and empirical maturity in other designs.

In the first instance, therefore, researchers have used other, simpler designs to identify variables that might at least potentially be candidate causal explanations of reading difficulties. As Backman, Mamen, and Ferguson (1984) have argued, in many studies on dyslexia, such as some of those reviewed here (e.g., Fawcett & Nicolson, 1995c, 1999; Fawcett et al., 2001; Moe-Nilssen et al., 2003; Nicolson & Fawcett, 1990; Raberger & Wimmer, 2003; Wimmer et al., 1999; White et al., 2006), normal readers have been used as a control group and matched to a dyslexia group on chronological age and sometimes also on overall level of intellectual functioning. Differences found between the groups on the measures in question have been then assumed to reflect the underlying causes that explain difficulties in reading (Backman et al., 1984). Nonetheless, the positive results obtained from these studies (i.e., overall poorer performances of children with dyslexia as compared to the control group) are difficult to interpret and may not really reflect causes of reading difficulties (Bryant & Goswami, 1986), due to the other possible between-group differences that have not been controlled (Backman et al., 1984). One major difference, as Bryant and Goswami (1986) have pointed out, is the fact that the groups compared in these studies differ on their reading level, such as in Moe-Nilssen et al.'s (2003) study described earlier or some of the other studies reviewed. Consequently, any differences found between the groups may be explained just as easily by the fact that they differ on their reading level (Bryant & Goswami, 1986).

In a reading-level design study, individuals with dyslexia are matched to a group of younger readers who are reading at the same developmental level as those with dyslexia (Goswami & Bryant, 1989). Positive results obtained from this design are more easily interpreted since the two groups are at the same reading level and any discrepancy found between the groups cannot be attributed to their differing reading achievement or simply to reading "delay" (Bryant & Goswami, 1986; Goswami & Bryant, 1989). Consequently, if the group with reading difficulties is worse on the particular measure in question, it is considered acceptable to conclude that skills underlying the specific measure examined may reflect causes of difficulties in reading (Bryant & Goswami, 1986). To interpret findings related to the variable in question in causal terms, the best design is considered one that uses both a reading-level and an age-level control group and makes comparisons between all of these groups (Bryant & Goswami, 1986; Goswami, 2006; Goswami & Bryant, 1989), as was applied in this study.

Unfortunately, many studies investigating cerebellar deficit have either not included a reading-level design (e.g., Fawcett & Nicolson, 1995c; Fawcett & Nicolson, 1999; Fawcett et al., 2001; Moe-Nilssen et al., 2003; Nicolson & Fawcett, 1990; Raberger

& Wimmer, 2003; Wimmer et al., 1999; White et al., 2006) or if they have there are methodological problems associated with the design, thus making interpretation of findings difficult. In the study by Fawcett et al. (1996), for example, which was described earlier, the authors reported using two of the age-match control groups (i.e., the groups with mean age of 14.6 and 10.9) as reading-age match controls for the two older dyslexia groups in their study (i.e., the group with mean age of 18.6 and 14.4). Their control group with mean age of 10.9 was also used as a reading-age match control for the dyslexia group with mean age of 18.6. While this was a positive point in the study, details related to matching were not provided by the authors. For example, the choice of reading test used to match the groups was not evident. Reading is generally considered to be a complex phenomenon and performance on even a simple task of accurate word reading can reflect complex cognitive central processes such as the activation of spelling rules for decoding or word-specific information as well as possibly more peripheral processes such as attention (Backman et al., 1984; Jackson & Butterfield, 1989). The processes underlying reading can become even more complex when tasks such as text level reading or reading comprehension are used for matching the groups (Jackson & Butterfield, 1989). The choice of measures used to match groups on reading level and reporting them in a study is thus crucial.

Additionally, Fawcett et al. (1996) did not report any statistical tests to explore whether the groups actually matched on their reading age. This would have been especially important since closer observations of average reading ages that were presented in a table for all groups seemed to suggest that the reading age for two groups that were used as reading-age match controls did not match the reading age of the older dyslexia groups they were matched to. In one of the cases, the reading age of the control group was one year higher than that of the dyslexia group. In the other case, the reading age of the dyslexia group was two years higher than the control group. Considering that the groups did not seem to be appropriately matched on their reading level, the causal interpretation of results obtained in the study may be questioned.

In fact, one study by Savage et al. (2005a) that has used a reading-level design appropriately has provided an alternative explanation for postural stability differences found among individuals with dyslexia and normal readers. Savage et al. (2005a) compared performances on phonological, rapid naming, and reading tasks as well as postural stability/balance in nine 10-year-old poor readers to 9 age-matched, and 9 younger reading-age match controls. Poor readers were found to perform significantly poorer than their chronologically-age-match peers on naming speed and some of the phonological tasks (e.g., spoonerisms and nonsense word reading). They were also found to perform significantly poorer than their reading-age match controls in nonsense word reading. Interestingly, however, they performed significantly *better* in a task measuring postural stability compared to their younger reading-age match peers.

Savage et al.'s (2005a) findings are intriguing mainly because both reading-age and chronologically-age match control groups were used in their study. Considering that samples in Savage et al.'s (2005a) study were matched on their reading ability, interpretations of findings become easier. In other words, poorer performance found on phonological tasks in poor readers as compared to the reading-age group cannot be due to differing reading levels but may rather be due to underlying skills necessary related to phonological and naming speed tasks. On the other hand, given that poor readers who were also older performed significantly better than the younger reading-age match group on measures of postural stability is intriguing because as Savage et al. (2005a) suggest this finding indicates that performance on balance automaticity may be more related to developmental maturation rather than to reading.

This view has also been supported by a more recent study in which Stoodley, Fawcett, et al. (2006) investigated the possibility of whether cerebellar difficulties were a result of developmental delay or if they continued through adulthood. Their sample included 28 adults with dyslexia and 26 adult controls. Adults with dyslexia were diagnosed by an educational psychologist prior to participating in the study. The diagnosis of dyslexia was reported to be based on discrepancies found between performance on standardized cognitive tests and participants’ achievement on reading and spelling tests. The sample with dyslexia was reported to have high cognitive ability with literacy skills that were poor in context of their cognitive ability, but still at or around their age level. The control group was reported to have no previous or current literacy difficulties. Comorbid conditions such as attention difficulties or developmental coordination disorders were not assessed in the study but the authors indicated the absence of these problems based on participants’ oral reports.

A balancing task (i.e., postural stability) was among the cerebellar measures reported in this study. The task was measured using a motion-tracking system that tracked movement in three dimensions at a rate of 120 HZ. The recorded data was then analyzed in the computer (Stoodley, Fawcett et al., 2006). In the balance task, participants were asked to balance alternatively on their right and left foot both with open and closed eyes, with each condition repeated three times. The amount of disturbance in balance was measured using head movements. To record head movements, participants had to wear a cap on their head, which was attached to the motion-tracking device using a Velcro strip. They were then asked to balance as best as they could during 10 seconds of recording time. Any dropping of the raised foot was also noted and reflected in the degree of wobble. Stoodley, Fawcett, et al. (2006) reported no significant difference in balancing tasks between adults with dyslexia and controls. Correlations found between the balancing tasks and literacy measures were also reported to be weak. The authors suggested that the reason no group differences were found in balance tasks might have been either due to the task being too easy for adults or merely because adults do not show any signs of balance difficulties.

(4) Lack of Homogeneity in Samples Involved in Studies

An important issue to consider in studies on dyslexia is the fact that we are not dealing with a true experimental study. As Jackson and Butterfield (1989) have stated, in a true experiment a crucial control factor is that participants are drawn from a single population and assigned randomly to different experimental treatments. Assuming large enough samples, group differences obtained could therefore be causally related to the treatment in question since the treatment factor is the only source of systematic variance between the groups and the effect of extraneous variables is controlled through the random allocation procedure.

However, this does not seem to be the case in many studies on dyslexia. In a typical study on dyslexia, the sample representing individuals with dyslexia is usually drawn from an extreme group, such as one diagnosed based on certain clinical criteria, whereas the comparison or control group is drawn from another source such as a representative sample of public schools (Jackson & Butterfield (1989). Despite the fact that this critic dates back even to the 1960s (e.g., Campbell & Stanley, 1963), the problem noted still seems to exist in dyslexia research. For example, this was the case in Fawcett et al. (1996) and many of their other investigations in which samples seemed to be drawn from different populations (i.e., possibly with differing educational systems or backgrounds). In these and many other studies, the origin of control groups is either unknown or the group is recruited from public schools. On the other hand, participants characterized as having dyslexia are usually drawn from clinics or schools for dyslexia and diagnosed based on the out-dated and empirically criticized discrepancy-based definition (e.g., Fawcett & Nicolson, 1995b; Fawcett & Nicolson, 1995c; Fawcett & Nicolson, 1999; Fawcett et al., 1996; Fawcett et al., 2001; Nicolson & Fawcett, 1990; Nicolson & Fawcett, 1994; Moe-Nilssen et al., 2003; Needle et al., 2006; Stoodley, Fawcett et al., 2006; White et al., 2006). Even if a reading-level design is used, this type of sampling still affects the degree to which the groups match in their reading level (Jackson & Butterfield, 1989). Any differences found between the groups is thus likely to be confounded by sampling procedures since the groups that are being compared may inherently be different due to the fact that by definition they have been drawn from different populations (Jackson & Butterfield, 1989).

Conclusion

In summary, the available findings for the cerebellar deficit theory seem to be mixed. The ambiguity in results appear to reflect methodological shortcomings including: (a) lack of control for possible confounding factors, (b) method by which motor and cerebellar tasks have been assessed, (c) lack of a reading-level design, and (d) lack of homogeneity in samples involved in studies. There has not been an attempt to address most of these methodological shortcomings in a single study and there is a need for such research. This was undertaken in the present investigation.

A larger scale sample using a more homogeneous population was recruited. In order to minimize the impact of methodological weaknesses associated with sample homogeneity, a solution is to draw samples from the same population as suggested by Jackson and Butterfield (1989). For example, both groups can be drawn from the same school or the same school system, as Savage et al. (2005a) did in their study. Additionally, it is also important to keep all sampling criteria constant except for the chronological age (i.e., older poor readers vs. younger normal readers). It might be noted that while using comparable sampling procedures is very important, this is quite rare in research in this domain. This was undertaken here. The study also used a reading-level design including both a reading-age and a chronologically-age match group. Group differences in performance were also investigated when the effect of attention was controlled statistically. Sensitive measures, yielding quantitative data, were utilized to assess the main cerebellar tasks. Besides the aforementioned factors that were taken into account in this research to improve the methodology, another approach undertaken here was exploring individual differences. To this end, evidence related to this approach is also evaluated in the next section.

Exploring Individual Differences

In addition to the cerebellar deficit studies considered in the previous four sections in which group comparisons were made between dyslexia and control groups, there are also some other studies that have explored individual differences to investigate this theory. These studies can be categorized into two subtypes depending on the sample involved. This includes (a) studies that have used typical readers, namely non-clinical samples; and (b) studies that have used atypical readers, namely clinical samples. Results from such studies can have important clinical implications. These studies are described below.

Studies using non-clinical samples. There are two studies in this category which have used a non-clinical sample. One was a correlational study by Brookes and Stirling (2005) who investigated the relationships between cerebellar deficits and dyslexia in 27 grade 4 students (ages 8 to 9). All children were reported to have average or above average intelligence. None had a diagnosis of dyslexia. They were all administered a battery of cognitive skills (i.e., word reading, digit span forward and backward, picture arrangement, and knowledge of common sequences - e.g., days of the week). A reading age was assessed for participants to calculate the discrepancies between their actual age and their reading ability. The authors reported obtaining a broad range of scores across cognitive tasks such as reading abilities ranging from 20 months behind chronological age to 38 months ahead. A cerebellar index was also obtained for each child based on measures of postural stability, muscle tone, and complex movements which were used in previous studies by Fawcett and Nicolson (e.g., Fawcett et al., 1996, Fawcett & Nicolson, 1999). Three additional indexes were also calculated for each of the domains assessed (i.e., postural stability, muscle tone, and complex movements). Brookes and Stirling (2005) reported significant correlations between reading age discrepancy, knowledge of common sequences, picture naming and cerebellar index, but no significant correlation was found between digit span measures and the cerebellar index.

Although this study is interesting in light of the fact that it uses a non-clinical sample, caution should be advised in how findings are interpreted since the pattern of results did not seem to support the cerebellar deficit theory. Indeed a closer observation of the correlation matrix for balance and reading age discrepancy presented in the article indicated that all participants including the 12 with reading ages below their chronological age had balance scores within the average range. It is also surprising that some reading measures that are used in many studies to assess phonological processing were not used in this study. This would have been useful especially since the cerebellar deficit theory suggests that phonological deficits are a result of a cerebellar impairment.

The second study in this category was conducted by Savage et al. (2005b) who also examined the relationship between motor balance automaticity and other literacy measures including rapid naming, phonological awareness, nonsense word reading, and rapid perception among 61 grade 3 and grade 5 children with below average, average, and above average reading and spelling abilities. Savage et al. (2005b) found a moderate relationship between postural stability and non-word reading in their study. Their findings also indicated that postural stability loaded modestly on the same factor as phonological processing and rapid naming. Nonetheless, postural stability was not found to be a significant predictor of word reading, nor did this measure reliably distinguish between the average and the below, and above average readers in their sample.

Altogether, the pattern of results from these two studies does not seem to provide strong support for the existence of a relationship between motor and cerebellar measures and reading. It is possible that relationships found between these measures are explained by other possibly motor components or processes that are shared between them. For example, Savage et al. (2005b) have suggested the correlations found in their study may not have been because balance automaticity underlies phonological processes but rather a result of common variance possibly in motor aspects of response production shared by both postural stability tasks and phonological processing tasks.

In summary, only two studies have explored the relationship between cerebellar deficit and dyslexia in a non-clinical sample and both were conducted in England. An investigation of this relationship is needed with a non-clinical sample as findings can have important implications. More specifically, using a non-clinical sample that includes only children attending mainstream schools can help resolve the problems related to sample homogeneity in clinical samples that were discussed earlier in this chapter. It is also important to determine if motor and cerebellar tasks can be successful in identifying mainstream children with reading difficulties. Finally, given that the studies exploring individual differences were conducted in England, it is important to extend these findings in different contexts to ensure that results are not limited to a specific context or curriculum. In this sense, extention of findings is a contribution since the provincial curricula in Canadian schools are different from that in England. In particular, the provincial curricula vary from school to school in Canada. In England, however, pre-specified objectives for teaching are implemented nationally. These involve specific prescriptions to teachers (e.g., for grammar, or for phonic work) that may influence children's performance on reading and other cognitive tasks.

Further, the study by Brookes and Stirling (2005) included a wide range of cerebellar tasks which were reduced into a total cerebellar index as well as three indexes for each of the domains assessed (i.e., postural stability, muscle tone, and complex movements). This reduction was not performed using statistical methods but was rather based on measures that were used in previous studies by Fawcett and Nicolson (e.g., Fawcett et al., 1996, Fawcett & Nicolson, 1999) under the category of postural stability, muscle tone, and complex movements. Reduction of motor and cerebellar measures via statistical methods has also not been attempted in previous studies conducted by Fawcett and Nicolson (e.g., Fawcett & Nicolson, 1999; Fawcett et al., 1996). Nevertheless, reducing a larger number of variables into a smaller set is recommended in order to increase the reliability and robustness of the results, as the inter-correlation between variables is a threat to homogeneity of variance. This has been undertaken in this study.

Studies using clinical samples. There is only one study using a UK sample that fits under this category. White et al. (2006) conducted a multiple case study to examine whether a range of sensorimotor deficits underlie difficulties in reading in dyslexia. Since White et al.'s multiple case study approach was followed in the present investigation, their study is described in detail. Participants in White et al.'s (2006) study included 23 children with dyslexia and 22 normal readers (ages 8 to 12) who were matched to those with dyslexia on age, non-verbal intelligence, and gender. All children were reported to have non-verbal IQ of at least 85. The control group was reported to consist of children from a larger sample who were screened first and then selected to match to the group with dyslexia on the aforementioned measures.

Children with dyslexia were all previously diagnosed by a chartered educational psychologist. Most of the children with dyslexia had received a neuropsychological assessment which did not include any direct measures of phonological awareness or sensory measures. Based on this assessment, children were classified depending on the severity of their symptoms using a 6-point scale, ranging from "not dyslexic" to "very severe dyslexic." The classification system as reported by White et al. (2006) was based on intellectual functioning and performance on reading, spelling, and other diagnostic tests such as speed of information processing and working memory (i.e., digit span). White et al.'s samples were reported to be classified on the highest three points of the scale (i.e., moderate, severe, or very severe dyslexia). Four children in the group with dyslexia had an additional diagnosis of dyspraxia, one child had an additional diagnosis of ADHD and another child had a diagnosis of both dyspraxia and ADHD. Children who were suspected of having broader language disorders were excluded from the sample.

All participants in White et al.'s (2006) study completed literacy tests which included a standardized assessment of their reading and spelling using the Wide Range Achievement Test-Third Edition, as well as measures to assess phonological awareness, short-term memory, and rapid automatized naming. These tasks included: (a) rhyme - identifying which two words out of three end with the same sound, (b) spoonerisms – e.g., replacing the first sound of a word with a new sound or exchanging initial sounds of two words, (c) non-word reading, (d) naming speed, and (e) fluency – saying as many words as possible in a given category.

All participants were also reported to complete a range of visual, auditory, and motor tasks to cover as much empirical and theoretical ground as possible. For the purposes of the present study, only tasks related to the cerebellar deficit theory are addressed here. These included: (a) bead threading – number of beads (out of 15) threaded onto a string as quickly as possible; (b) finger to thumb – thumbs and index fingers of opposite hands joined, lower thumb-finger is released, hands are rotated in the opposite direction to join again at top (time to complete 10 times as quickly as possible); (c) stork balance – time spent standing on one foot without moving the other foot from the supporting knee or the hands from the hips recorded for up to 20 seconds; and (d) heel-to-toe – number of steps achieved while walking along a line, placing heel of one foot against the toe of the other for up to 15 steps.

White et al.'s (2006) findings indicated that children with dyslexia performed significantly poorer than normal readers on measures of reading, spelling, phonological tasks (including rhyme, spoonerism, picture and digit speed naming, and rhyme fluency).

No group differences were reported in motor tasks of bead threading and finger to thumb, but group differences were found on tasks of stork balance and heel-to-toe walking.

To investigate individual differences on measures, White et al. (2006) first removed participants in their control group who had "abnormally low performance" on any of the measures. These were identified as children whose performance on measures was more than 1.65 standard deviations below the mean of the control group. As a next step, White et al. created summary factors which accounted for all tasks in a given modality in their study. The summary factors relevant to this thesis included: (a) a literacy factor (i.e., combined performance on reading, spelling, and non-word reading); (b) a phonology factor (i.e., combined six phonological tasks including spoonerisms, picture and digit naming, rhyme fluency, rhyme, and non-word reading); and (c) a motor factor (i.e., combined all motor tasks). White et al. reported paritialling out the effect of non-verbal intelligence from all of their factors because of significant correlations they found between non-verbal IQ and literacy ability. Age was also partialled out from all their sensorimotor factors. In addition to group comparisons, individual performances were graphed for each of their factors.

For the literacy factor, significant group differences were reported, indicating overall poorer performance in children with dyslexia as compared to the control group. Based on the individual analysis, all 23 children with dyslexia were reported to be outliers on the literacy factor compared with only one child in the control group. This is illustrated in Figure 2 on page 71, in which values on the y-axis are *z*-scores for the literacy factor with the mean of zero, and the outliers are those below the -1.65 cut-off point shown by a broken line in the figure.

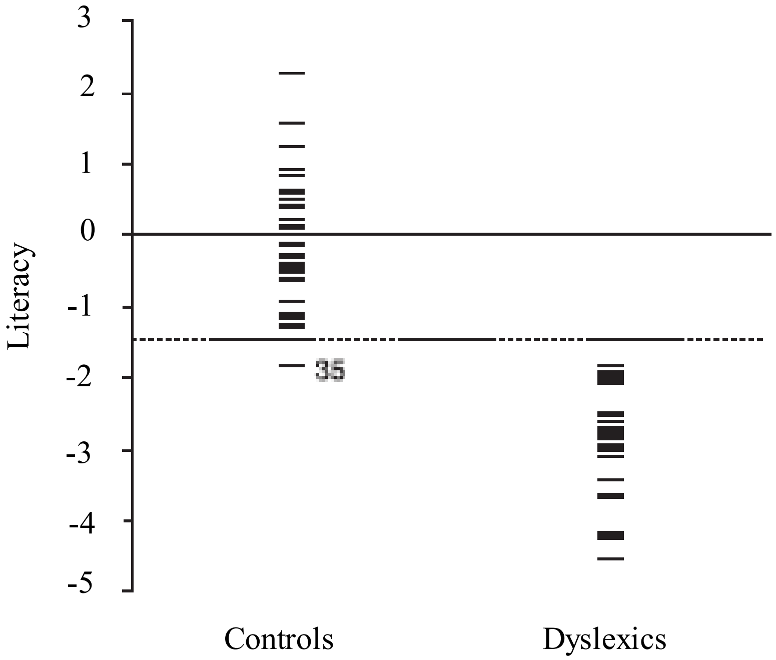


Figure 2. White et al.'s (2006) graph of individual performance for literacy summary factor.

The outlier in the control group was reported to be removed from all other analyses. Therefore, based on author's findings, literacy factor seemed to successfully distinguish between children with dyslexia and normal readers not only at a group but also at an individual level. For the phonology factor, the group with dyslexia was reported to show poorer performance than the control group. As illustrated in Figure 3 on page 72, following individual analyses, 12 of 23 children with dyslexia were reported as outliers on the phonology factor as compared to none of the children in the control group.

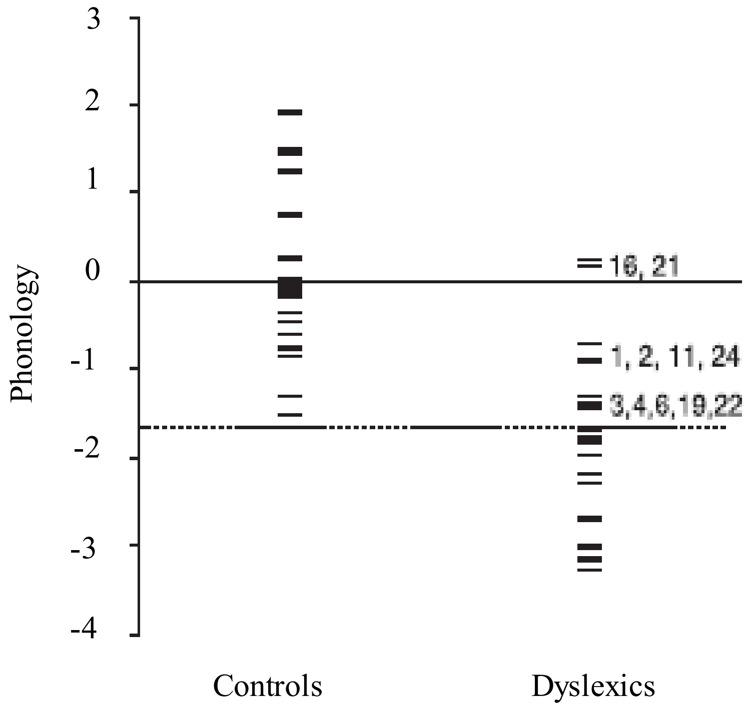


Figure 3. White et al.’s (2006) graph of individual performance for phonology summary factor.

Again, the outliers are cases falling below the -1.65 cut-off point that is shown by a broken line. Therefore, similar to the literacy factor, the phonology factors used in White et al.’s study also successfully distinguished between children with dyslexia and normal readers at both the group and individual level.

Finally, for the motor factor, group comparisons did not reveal significant results. White et al.’s individual analysis of the motor factor, which is illustrated in Figure 4 on page 73, indicated that only 5 of 23 children with dyslexia were outliers on this factor (i.e., cases below the broken line).

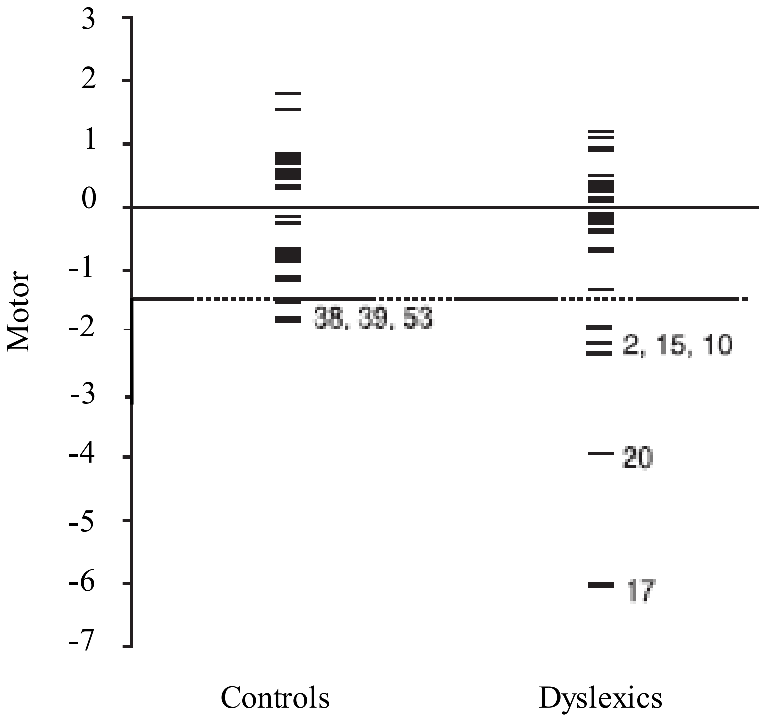


Figure 4. White et al.'s (2006) graph of individual performance for Motor summary factor.

Three children in the control group were also reported to be outliers on the motor factor. Hence, unlike literacy and phonology factors, the motor factor in White et al.'s (2006) study did not successfully distinguish between children with dyslexia and normal readers at a group level or at an individual level. Overall, these results can have important clinical implications. For example, based on these findings one could argue that tasks that assess phonological skills as opposed to those assessing motor skills would be more appropriate for assessment, diagnosis and identification of individuals with dyslexia.

Based on the overall pattern of individual analysis, White et al. (2006) reported having 14 children with dyslexia who had a sensorimotor impairment in addition to phonology and literacy impairments. Five of these 14 showed poor performance on the motor factor. White et al. compared all 14 children exhibiting sensorimotor problems in the group with dyslexia with the other children in the same group who did not exhibit such difficulties on the sensorimotor factors. No statistically significant differences were found between these two sub-groups on literacy and phonological awareness measures.

White et al. (2006) also explored the extent to which the different deficits predicted literacy ability. Results from their correlation analyses indicated that phonology was a good predictor of literacy skill for the whole sample as well as within each group. This was reflected in strong correlations obtained between the literacy and phonology factor for the entire sample ($r = .76$), for the dyslexia group ($r = .60$), and moderately strong correlation obtained for the control group ($r = .45$). White et al. also reported moderate correlations for the whole sample between the motor factor and both the literacy and phonology factors ($r = .38$ for both). However, as White et al. noted, these correlations seemed to be due to two children with dyslexia who had very poor performance on all three domains (i.e., motor, literacy, and phonology). After removing these two outliers, the correlations between motor factor and both literacy and phonology factors were no longer significant. Additionally, White et al. conducted a multiple linear step-wise regression to investigate which factors predicted literacy performance. Their findings indicated that the phonology factor accounted for 60% of the variance on literacy performance. As reported by the authors, none of the other factors (including motor, visual, or auditory) were found to be significant predictors of literacy.

According to White et al.'s (2006) investigation of individual performances, the majority of cases of reading difficulties seemed to be directly caused by deficits in phonological processing which could not be accounted for by auditory or motor impairments. According to White et al., while there seemed to be an association between sensorimotor syndrome (i.e., referring to the 14 cases that had shown motor, visual and/or auditory deficits in addition to phonology and literacy) which might be an indication of some common underlying biological factor, this syndrome did not directly explain reading difficulties.

White et al.'s (2006) study is intriguing in that it has attempted to examine a range of sensorimotor difficulties in dyslexia. Additionally, it is innovative in its attempts to investigate differences on measures not only at a group but also at an individual level. As Ramus, White, and Frith (2006) have noted, "individual data are important because they assess the extent to which children who are diagnosed as dyslexic have similar sensorimotor and cognitive profiles, and whether a single deficit could underlie the reading problems" (p. 266). Moreover, results obtained from individual analysis can have important clinical implications for practitioners. According to Bishop (2006), findings of White et al.'s (2006) study pose a challenge for the sensorimotor account of dyslexia because they demonstrate that different impairments are only weakly related to dyslexia and they don't seem to form a coherent pattern. On the other hand, difficulties in dyslexia seem to be directly related to a linguistic-phonological impairment. While Bishop (2006) advise caution against concluding that impairments in phonology are causally related to poor reading, she agrees that phonological tasks in White et al.'s study were more successful than the sensorimotor tasks in differentiating children with dyslexia from those in the control group.

Critical Evaluations of the White et al.'s Study and the Response by White et al.

White et al.'s (2006) study has attracted the attention of other researchers. Some have criticized White et al.'s work on a range of different grounds. A total of four independent critical commentaries on the White et al.'s target article (Bishop, 2006; Goswami, 2006; Nicolson & Fawcett, 2006; Tallal, 2006) were published in the same issue of the journal of *Developmental Science* as the original White et al. paper. In addition, White et al. provided a response to these commentaries in the same issue of that journal. Some of the relevant criticisms of the White et al. paper are discussed here in a thematic form, as they pertain to the present study, along with White et al.'s research team's response to these specific criticisms (Ramus et al., 2006). In addition, an evaluation of the relative strength of these two sets of arguments (commentaries and subsequent defense) is provided.

Critique of the sampling of the group with dyslexia. According to Goswami (2006), Nicolson and Fawcett (2006) and Tallal (2006), a problem identified in the White et al. study was that the children who were diagnosed with dyslexia in the study may have been "garden-variety" poor readers. Generally, the critiques suggested that these children did not seem to meet the discrepancy-based diagnosis of dyslexia as they did not always show very poor reading ability, and sometimes had quite typical phonological ability (e.g., standard scores around 100 on rhyme and alliteration tests). In addition, the "typical readers" often had somewhat above-average reading ability (e.g., a non-word reading test standard score of 114.95; Tallal, 2006).

In response to these critiques, White et al. point out the issue that most of the children with dyslexia had an independent formal assessment and had been documented by the Dyslexia Institute (a center that undertakes formal assessment of dyslexia) along with a history of dyslexia. White et al. also note that many of the norms for reading and phonology tests may be outdated, as the widespread use of phonic approaches to reading in the National Literacy Strategy (NLS) in England may have led to higher overall national standards on these tasks and that at least some tasks they used were normed in the United States so standard scores might not be representative of U.K. norms. Finally

White et al. also note that the children with dyslexia and the control group children differed significantly overall on these measures and met formal regression–based criteria for dyslexia.

In evaluating these comments, it is worth noting that it might be desirable in future research in this domain to use and report psychometric tests standardized in the same countries as the research is undertaken to confirm that typical readers were indeed typical (that is, with reading test and phonological awareness test standard score means around 100) and poor readers showing reading and phonological deficits (e.g., mean standard scores around 85 on these measures). Given the purported role of a particular curriculum (the NLS), studies in other geographical contexts would also help in interpreting whether observed patterns partly reflected the particular pedagogical approaches used in England currently, as White and colleagues suggest.

Manipulation of the control group. Tallal noted the fact that the group differences found in White et al.'s study might have been intensified because of the fact that outliers in the control group were detected and removed. In response to this critique, White et al. noted that this procedure was only used to detect outliers in their control group prior to the individual analyses. Otherwise, all scores reported in figures and tables were unaltered without any data point excluded. According to White et al., removing outliers from the control group only affected the deviance threshold for each variable (i.e., outliers in each domain in individual analyses). The authors suggested this was necessary to prevent occasionally inattentive participants in the control group from falsely affecting the normal range of performance and thus reduce the possibility of detecting outliers in the group with dyslexia.

While assumptions drawn about the degree of attentiveness in participants' performance may be correct, they may also be biased or less objective. Additionally, the possibility that attentional issues might have affected the performance of some individuals in the control group can just as easily apply to participants in the dyslexia group. As reviewed in some studies discussed earlier, attention has been associated with performance on motor and balance tasks and controlling for it is important. Hence, a more objective way to control for possible effects of attention in both groups would have been to use a more standardized measure to screen individuals and to statically control for it in *both* groups.

Lack of a reading-age control. Turning back to critiques of the White et al.'s study once again, Goswami (2006) pointed out the fact that White et al.'s study did not include a reading-age control group. According to Goswami (2006), including both a reading- and chronologically-age match control group has been considered optimal in studies of dyslexia. The combination of results from the two designs is considered useful. A chronologically-age level control can serve as a complement to the reading-level control since an age-level control can identify processes or skills on which poor readers and their same aged peers don't differ and thus may not be potential causes of difficulties in reading (Jackson & Butterfield, 1989).

However, White et al. refrained from using a reading-level design because they argue that,

> "reading-age matched controls, being younger children, would inevitably have worse, or equal (but certainly not better) performance on the sensorimotor tasks than the age-matched controls. This would make the dyslexic group even *less*

> deviant (probably indistinguishable) on sensorimotor measures with respect to that reading-age control group." (White et al., 2006, p.252)

They further state, "if we were to perform a comparison with a reading-age control group, this could only reinforce our conclusion (perhaps spuriously) that sensorimotor deficits affect only a small subset of this group of dyslexic children" (p. 252).

In evaluating these claims, the point made by White et al. (2006) that the sensorimotor hypotheses fared very poorly even using a chronologically-age match control is important and arguably represents a *minimum* requirement of any theory of dyslexia. From this view, a reading-age match design is thus likely to be even more conservative. In addition, if one returns briefly to the results of the review of the methodological techniques of reading-age match and chronologically-age match studies earlier in this thesis, it might be worth recalling that Goswami was a strong advocate for the view that null findings from the chronologically-age match but not reading-age match studies *are* interpretable as evidence against a causal role for a given variable (Bryant & Goswami, 1986; Goswami & Bryant, 1989). That is, only positive results in the reading-level control and negative results in the chronologically-age level control are interpretable (Goswami & Bryant, 1989). In a chronologically-age match control, positive results obtained are not interpretable because of the differing reading levels in individuals with dyslexia and the age-matched control (Bryant & Goswami, 1986; Goswami & Bryant, 1989). On the other hand, according to Goswami and Bryant (1989), negative results derived from an age-level match control where no differences are found between the dyslexia and control groups is interpretable in causal terms. That is, if the two groups who are at the same age and cognitive ability are also equal in the variable being measured, then negative results indicates that the variable in question may not be causally related to reading difficulties. Others were more cautious about interpreting null findings in the absence of full experimental control through randomization (e.g., Jackson & Butterfield, 1989).

Notwithstanding these points, using both a chronologically- and reading-age match control group can shed further light on findings depending on the patterns found. For example, if as compared to a younger reading-age match control group, the older poor readers and their age-match controls perform better and equally well, then it may be that the reading-age match group is poorer due to being younger and less mature. But more importantly, as Goswami (2006) has noted, the reading-level control group would have been important for arguments about developmental causation since results would have demonstrated whether the children with dyslexia performed similarly or worse than a younger reading match control group on sensorimotor tasks. Without a reading-level control group, interpretation of findings becomes difficult.

Critiques of the power of the study. White et al.'s study was also critiqued for factors such as the wide age range and small sample size which might have affected the power of the study to reveal group differences (Goswami, 2006; Nicolson & Fawcett, 2006). As Goswami (2006) noted, children can be quite different in their performance on sensorimotor and cognitive tasks depending on their developmental stage. Hence, the study could have compared groups at different points in development.

In response to these critiques, White et al. (2006) agreed that chronological age was related to sensorimotor tasks in their study and that this was controlled for by partialling out age from all sensorimotor scores that were entered in statistical and individual analyses. However, they pointed out that in regard to sample size, their study was similar to many other studies that have investigated sensorimotor tasks. Overall, according to White et al. there was no reason to believe that their study was "underpowered". As they noted, the main point of their study was not whether there were significant group differences but it was rather to analyze patterns of deficits within each individual. As they emphasized, "individual data are important because they assess the extent to which children who are diagnosed as dyslexic have similar sensorimotor and cognitive profiles, and whether a single deficit could underlie the reading problems" (Ramus et al., 2006, p. 266). Consequently, White et al. (2006) suggested that testing 44 cases in depth seemed sufficient for their purposes, as their study was a "multiple case study," which is rare. One might also note that such studies are relevant to the practical utility of theories of dyslexia to professionals for the assessment of individual children with reading difficulties.

Conclusion

In summary, White et al.'s (2006) multiple case study provides findings that can be valuable in clinical practice. An investigation similar to White et al. is much needed using a non-clinical sample outside of England to ensure that results are not limited to a specific context or curriculum. As noted earlier, using a non-clinical sample can help resolve issues related to sample homogeneity in clinical samples. Moreover, an extention of White et al.'s findings is a contribution especially given that the provincial curricula in Canada are different from that in England. It is also important to determine whether motor and cerebellar tasks can be successful in identifying mainstream children with reading difficulties. Consequently, an approach similar to White et al. was undertaken in this project with some changes to improve the design considering some of the critiques that were made to White et al.'s study. Considering the importance of a reading-level design for interpretation of findings, this study included a reading-age match group. Unlike White et al., who removed outliers from their control group prior to exploring individual differences, in this study this step was not followed. White et al. did not statistically control for attention in their investigation of group and individual differences. This was undertaken in this project. Finally, a non-clinical as opposed to a clinical sample was used in this study.

Summary and Implications for the Present Investigation

Recently, apparent automatization deficits in children with dyslexia have been linked to a congenital mild deficit in the cerebellum. The cerebellar deficit theory is suggested to combine both learning and neurological perspectives to provide a unifying explanatory framework of the existing theories of dyslexia (Nicolson et al., 2001). According to this theory, a mild inborn deficit in the cerebellum is proposed to be the underlying cause of impairments in phonological processing and naming speed (Nicolson et al., 2001). More specifically, this mild deficit is suggested to cause a cascading series of impediments in motor and visual skill areas, as well as difficulties in central processing speed and in acquisition and automatization of elementary auditory skills. These deficits will, it is claimed, then lead to difficulties in writing, spelling, rapid naming, phonological processing, and consequently reading will be impaired (Nicolson et al., 2001).

Research investigating the cerebellar deficit seems to be growing. Nonetheless, evidence remains inconsistent and the validity of the theory has been questioned on several accounts. Behavioral evidence for the cerebellar deficit (i.e., motor difficulties and poor postural stability) has also not been consistent across studies. Evaluating the body of evidence indicates that the inconsistencies may possibly be due to methodological shortcomings including (a) lack of control for possible confounding factors, (b) method by which motor and cerebellar tasks have been assessed, (c) lack of a reading-level design, and (d) lack of homogeneity in samples involved in studies.

In summary, the behavioral evidence related to cerebellar deficit is limited as most existing studies have at least one of the enumerated major shortcomings. So far there hasn't been an attempt to address most of these methodological shortcomings in a single study. This was undertaken in the present investigation. A larger scale sample using a more homogeneous population was recruited. Instead of using the much criticized discrepancy-based criterion, performance on a word reading task was used to identify dyslexia. The study also used a reading-level design including both a reading- and a chronologically-age match group. Sensitive measures were utilized to assess the main motor and cerebellar tasks. Additionally, in light of the evidence derived from White et al.'s (2006) multiple case study and its implications for clinical practice, individual analyses were also undertaken in this study in addition to group comparisons. Group and individual differences in performance were also investigated when the effect of attention was controlled statistically. Finally, as a preliminary step, in order to increase reliability of results, data reduction techniques were used to reduce the motor and cerebellar variables. These improvements and contributions of the present study are described in more detail in the next part.

Part 3: The Present Study

Strengths and Contributions of the Present Study

(1) Control for Attention Difficulties

Since evidence seems to indicate that a strong predictor of balance effects across studies is the proportion of participants who are screened for ADHD symptoms (Rochelle & Talcott, 2006), all children in the present study are also screened for possible attention problems. Attentional problems have also been shown to co-occur with reading difficulties (Rochelle & Talcott, 2006). Hence, screening for the presence of ADHD and controlling for it statistically allowed the investigation of a possible impact of attentional problems on group differences found in both reading and cerebellar measures.

(2) Use of Sensitive Measures for Motor and Cerebellar Tasks

This study includes a wider range of motor and cerebellar measures as compared to some other studies. The static cerebellar tests, which are suggested to have a unique diagnostic power by Nicolson and Fawcett (2006) are also both included in this study. As opposed to more subjective Likert scale measurements, careful scientific measurement of motor skills (in postural stability and muscle tone) using an accelerometer sensor and a goniometer (explained in detail in Method) also add to the reliability of findings in this study. Additionally, statistical data reduction techniques are also employed to obtain extra, more reliable clusters of motor and cerebellar measures.

(3) Use of Reading-Level Design

Overall, the use of reading-level design and including both a reading-age and a chronologically-age match control improves the chances for interpretation of findings at a causal level (Goswami, 2006). There is only one study by Savage et al. (2005a) that directly contrasts phonological, RAN, and motor deficit accounts using an RA-match design. The present RA-match study is a much-needed contribution to this literature. As indicated by Jackson and Butterfield (1989) and discussed earlier, the strength of a reading-level design can be maximized by careful consideration of some quite basic methodological and statistical factors that are often ignored in some reading-level designs. Therefore, careful measures are undertaken in this study that follow Jackson and

Butterfield's (1989) suggestions for increasing the strength of reading match studies. These are:

1. As recommended, the sample in the present study is drawn not from different populations but recruited from the same school system.
2. Important details about the sample, such as criteria used for sampling, details about consent, attrition rate in the sample, which are often missing in other studies, are reported in this project. I have followed the guidance on best practice in clinical trials advocated by the CONSORT team criteria. Thus, following Moher, Schulz, and Altman's (2001) recommended best practices from medical guidelines for sampling and recruitment, a flow chart is included in the Method section that displays details about sampling procedure and attrition rates. In some cases where information was missing for the sample, I have investigated the possible impact of this factor on the findings.
3. The present study also explores whether the three subgroups selected from the sample in the study are homogenous on measures of IQ and first language spoken by participants to ensure group differences are less likely to be explained by other extraneous factors.
4. Again as Jackson and Butterfield recommended, the questions addressed in this project are explored using various techniques mentioned earlier. All findings (including null findings) are also explored using a range of appropriate convergent techniques such as effect size analysis to ensure that statistical significance is also balanced with considerations of the practical significance of findings.

5. Additionally, as recommended by Goswami and Bryant (2006) the reading-level design in this study includes both a reading-level and an age-matched control group which have been considered optimal. Including both of these groups is very important for making causal interpretations.

(4) Increasing Sample Homogeneity

As was indicated earlier, in most studies the sample representing individuals with dyslexia has usually been drawn from an extreme group (e.g., one diagnosed based on certain clinical criteria such as the outdated discrepancy-based definition), whereas the comparison or control group has been drawn from another source such as a representative sample of public schools. Consequently, any differences found between the groups is likely to be confounded by sampling procedures since the groups that are being compared may be inherently different due to the fact that by definition they have been drawn from different populations (Jackson & Butterfield, 1989).

In order to address these problems, the present study is the first study in Canada that has attempted to investigate cerebellar deficits and a range of reading measures in a non-clinical sample of mainstream elementary students. As noted earlier in evaluating the critiques of, and responses by, White et al. (2006), one concern was that a particular pedagogical approach (namely The National Literacy Strategy in England), might have affected the performance of the children with dyslexia and the typical readers in a way that might not be evident elsewhere. Certainly there is evidence that following such intervention scores on phonological awareness and non-word decoding tasks are somewhat higher than in test norming locations in North America. An extention of the White et al.'s study in Canada is therefore a useful contribution to knowledge in and of itself as it addresses the important issue of the generalizability of findings. The reason for using a mainstream sample is to reduce possible sources of sampling bias associated with clinical population in some studies that investigated cerebellar deficits. As Jackson and Butterfield (1989) suggested, drawing samples from the same population (e.g., the same school or school system) minimizes the impact of such methodological weaknesses. A larger sample is also used here in order to ensure obtaining a wide range of reading abilities including poor, average and good readers.

Exploring Individual Differences

This study is also a contribution in the general sense that it attempts to extend White et al.'s (2006) findings. This work is among the very few studies that have investigated phonological *and* cerebellar deficits outside the UK in Canada. There are some important practical additional advantages that arise from this. In particular the results are, by definition, not affected by the particular curricular approach in England (The National Literacy Strategy) which White et al. have argued led to substantially higher reading ability in their poor and average reader groups, and which also might conceivably have influenced *the way* children read. That is to say, as all children in the White et al. study had at least average phonological skills and good decoding skills, this might have meant they relied on these skills in a way that children in other contexts might not do. Hence, a extension of White et al.'s findings in a distinctly different context would suggest that a purely curricular explanation of results seems unlikely.

Analysis at an individual level in addition to group level comparison also adds to the strength of findings in this study. As indicated earlier, this is important in terms of better testing models suggested by Jackson and Butterfield. Additionally, as Ramus et al. (2006) have also argued: "it is no use to keep inundating the literature with study after study showing group differences on sensory or motor tasks: this will not do" (p. 268), and that future studies: "will be convincing only to the extent that they provide reliable individual data and demonstrate much greater prevalence and predictive power of sensorimotor deficits in the dyslexic population than has been observed so far" (p. 268). Additionally, the pattern of findings derived from the present study has important implications for school psychology and practice as discussed later in this thesis.

The present study differs in some respects from White et al.'s (2006) as explained below:

Selection of subgroups. In White et al.'s (2006) study, dyslexia was defined based on IQ-achievement discrepancies. Participants were also required to have non-verbal IQ scores above 85. In contrast to White et al., this study includes an estimate of intellectual functioning that comprise of both a verbal and non-verbal scale. However, IQ is not used as a selection criterion and participants are not excluded based on IQ scores. Nonetheless, once the subgroups are selected, it is explored whether they matched on their intellectual functioning.

Cut-off point for individual analysis. The present study also differs from White et al.'s (2006) in the cut off point used to determine outliers (or poor performance) in the subgroups selected from the sample. As indicated earlier, prior to individual analysis White et al. removed outliers from their control group with scores that were more than 1.65 standard deviations below the control mean. They then recalculated the control mean and poor readers whose performance on tasks was more than 1.65 standard deviations below the new recalculated control group were considered outliers.

The present study differs in the sense that White et al.'s (2006) preliminary step of removing outliers from the RA-match control are not followed. White et al.'s cut-off point is also not used to detect outliers on tasks partly because of the fact that this study includes mainstream children for whom a prior diagnosis is not known. Struggling readers in the present sample are identified based on word reading performance that is below the 25th percentile (i.e., more than .67 standard deviation below the mean). This cut-off point has been used in other studies as well (Fletcher et al., 1994; Juel, 1988; Shaywitz et al., 2004). In order to maintain consistency across the study, this cut-off point is also used for individual analyses.

Type of sample. White et al. (2006) investigated individual data in a group of children that were previously diagnosed with dyslexia (using the discrepancy-based definition) compared to a control group that was matched to the group with dyslexia on gender, age, and non-verbal intelligence. This is not the case in the present study as knowledge on whether the sample had a prior diagnosis of dyslexia is unknown. The sample in this study includes mainstream children. Most of the children in the dyslexia subgroup seemed to match the Quebec definition of dyslexia which also uses the discrepancy-based criterion. However, the discrepancy-based criterion as such is not used in this study to identify children with dyslexia. Instead, performance on a word reading task is used. Once subgroups are selected, whether they match on IQ and first language spoken is explored.

Control groups. In contrast to White et al.'s (2006) study, the present project includes both reading-age and chronologically-age match controls. Individual differences in this study are investigated between the dyslexia subgroup and the reading-age match control. Since these groups do not differ in their reading level, the chances of drawing causal conclusions are increased.

Control for attention. In contrast to White et al.'s (2006) sample, the sample in this study is screened for attentional difficulties. This allows for further investigation of individual differences in reading with the impact of attention controlled.

Summary factors. The summary factors for the present study differ from those in White et al.'s (2006) study. First and foremost, group comparisons made in a reading-level design are based on using raw scores. Consequently, the individual analysis in this study is also based on raw scores for reading measures as opposed to standard scores, which were used in White et al.'s study. Second, in contrast to White et al.'s study, which included a literacy factor based on measures of word reading, non-word reading and spelling, in this study a literacy factor is not created. This study does not include a spelling measure and the non-word reading task is an outcome measure on which the groups are being compared. Additionally, since a reading-level design is used in this study, participants in the dyslexia subgroup and the reading-age match subgroup are matched on their reading level using the word reading task. Consequently, a separate figure is displayed to illustrate the reading ages on the word reading task for the dyslexia subgroup and the reading-age match control.

Third, White et al. (2006) created summary factors for tasks in their study that belonged to one modality. The factors created for the present study correspond to the tasks that are used in this project. Unlike White et al. who combined naming speed along with other phonologically based measures to create a phonology factor, these measures are not combined here. Individual differences in the two phonologically based measures (i.e., non-word reading and phonological awareness) used in this study are each investigated separately. I have also investigated individual differences in speed naming measures separately. The aim here in constructing two separate variables is to evaluate the possibly-independent effects of rapid naming and phonological awareness. This study also includes measures related to fluency in word and non-word reading. Individual differences in these measures are also investigated separately. Finally, all motor and cerebellar measures are not combined together as White et al. did. Instead, the clustering of the inter-relationships between postural stability, toe tapping, motor speed, and muscle tone measures is first investigated empirically. The latent motor and postural stability variables are then used to investigate possible associations with literacy.

Research Questions

The present study intends to investigate four main questions regarding the involvement of cerebellar processing in reading acquisition. Prior to addressing the main questions a preliminary step is undertaken to investigate the following question:

Can the motor and cerebellar measures used in this study be reduced into separate clusters using statistical data reduction techniques?

As noted earlier, this step is taken to increase the reliability of results. While a statistical reduction of motor and cerebellar measures has not been attempted in previous studies conducted by Fawcett and Nicolson (e.g., Fawcett & Nicolson, 1995c, 1999; Fawcett et al., 1996), the authors considered the tasks of bead threading and peg moving as pure motor measures, whereas the three tasks assessing postural stability, muscle tone, and speed in toe tapping were considered as cerebellar tasks. By using a principal component analysis, it is investigated whether the motor and cerebellar measures form clusters that would confirm the pattern suggested by Fawcett and Nicolson (1995c, 1999) or if the analysis would reveal different components.

Question 1:

Is there a relationship between word reading as measured by word identification and (a) phonological awareness, (b) reading fluency and rapid automatized naming, and (c) purported cerebellar processing tasks?

Considering evidence from convergent studies that have shown strong relationships between reading and phonological awareness in dyslexia, it is expected to find significant relationships between word reading measures in this study and tasks that measure phonological processing. Additionally, if as claimed by the cerebellar deficit theory, a cerebellar deficit is the underlying cause of phonological, rapid naming, and eventually reading difficulties, then motor- and cerebellar related measures should be correlated to tasks that measure phonological awareness, reading fluency, rapid naming, and reading.

Question 2:

Does a subgroup of children with dyslexia selected from the sample differ in their performance on any of the motor, cerebellar, reading, phonological, and rapid naming related measures when compared to two control subgroups that were selected from the same sample and matched to the dyslexia subgroup based on (a) their reading level, and (b) chronological age?

This question is addressed using 10 one-way analyses of variance. All the analyses include one between participants factor with three levels, namely group (i.e., dyslexia subgroup versus reading-age match versus chronologically-age match). Evidence supporting this assumption has been inconsistent due to methodological and sampling issues. Additionally, Savage et al. (2005a) who used a reading-level design, found group differences only in naming speed and some of the phonological tasks. They also found group differences in the postural stability task in favor of the older poor readers. Nonetheless, assuming that a cerebellar deficit explains reading difficulties in children with dyslexia and that the motor and cerebellar tasks assess a cerebellar deficit, then according to the cerebellar deficit theory children with dyslexia should display deficits in motor- and cerebellar-related tasks in relation to the reading- and chronologically-age match controls.

Question 3:

Do any group differences in performance on the above reading and cerebellar related measures emerge when the effect of attention is controlled statistically?

This question is addressed using 12 one-way analyses of covariance. As indicated in the literature review, some evidence has shown that while phonological and naming difficulties are associated with reading difficulties, performance on motor tasks is rather associated with attentional difficulties (e.g., Denckla et al., 1985; Raberger & Wimmer, 1999, 2003; Ramus, Pidgeon et al., 2003; Wimmer et al., 1999). Additionally, as discussed earlier Rochelle and Talcott (2006) who addressed inconsistencies found in studies on balance impairment in dyslexia in their meta-analysis, a strong predictor of balance effects across studies was the proportion of participants who were screened for ADHD symptoms. Nonetheless, following a few studies in which Fawcett and Nicolson have excluded children with possible attentional difficulties (e.g., Fawcett et al., 1996; Fawcett et al., 2001), the authors have suggested that cerebellar impairment in children with dyslexia seems to be independent from the presence or absence of ADHD. Consequently, if this assumption is correct, then the group differences in cerebellar-related measures should survive the attention covariance in the present investigation.

Question 4:

Does the cerebellar deficit provide a good explanatory model at the individual level?

To investigate individual differences in performances in reading and cerebellar measures in the dyslexia subgroup and the reading-age match control, the multiple case study procedure used by White et al. (2006) is broadly followed. Additionally, individual differences in task performance are investigated when the effect of attention is controlled. White et al. (2006) found in their study that compared to cerebellar measures, phonological awareness tasks were more successful in differentiating between participants with dyslexia and their same-aged peers who were normal readers. In the present study, a reading-level design is used. Considering the alleged link between a cerebellar deficit, phonological processing, rapid naming, and reading, the motor and cerebellar measures should be as successful as phonological and rapid naming measures (if not more) in distinguishing between the participants in the dyslexia subgroup and those in the reading-age match control subgroup in this study.

CHAPTER 3

Research Methodology

This chapter is separated into four different parts. Part one provides information on the participant recruitment process for this study followed by part two which describes the procedure for testing participants. Part three of the method section presents all the demographic information that was collected from parents about the participants. Finally, the last part of the method section lists and describes all measures, including behavioral, cognitive, reading, phonological, rapid naming, motor and cerebellar tasks.

Part 1: Participant Recruitment Process

Prior to providing an overview of the recruitment process and the final pool of participants, a few points that impacted participant recruitment and their enrollment in the present project are explained.

Limitations Affecting Participant Recruitment

Limitation related to school selection. Recruitment was limited to only schools in which English was the leading language during the early elementary years. There were only a few schools that met this condition. This is because the official language of the province of Quebec is French. In Quebec, all students must attend French-language public schools under the Charter of the French Language that was passed in 1977, unless certain conditions apply (Office Québécoise de la Langue Française, 2008). It is only under specific conditions that a "Certificate of Eligibility" is granted to permit a child to attend an English-language school. According to the "Minister of Education in Quebec/ Ministère de l'Éducation" (2009), under the Charter of the French Language, in order to qualify or be entitled to receive instruction in English, the child must be first and foremost a permanent resident of Quebec. Children who live in Quebec temporarily may also qualify for a temporary authorization to receive instruction in English. Among these children, a certificate of eligibility for attending an English-language school is granted only to those (a) who did most of their elementary or secondary studies in English in Canada; or (b) whose siblings (i.e., brother or sister) or parents (i.e., mother or father) did most of their elementary studies in English in Canada; or (c) whose father or mother attended school in Québec after August 26, 1977 and could have been declared eligible for instruction in English at that time. In short, these conditions guarantee English exposure at home meaning that for the participants of this study (or at least for most of them), English was the first language (L1) or among one of the first languages at home. Therefore, the sample seemed to generally represent genuine English-dominant bilinguals who were exposed concurrently to English and another language.

In addition to the above rules that apply to children attending English-speaking schools in this province, it is also important to understand the attempts made to specifically target Anglophones or English-dominant bilinguals. While English school boards in Montreal are responsible for Anglophone public schools, there are only a few schools with English as the leading or the only language of instruction. Language immersion programs are by law used as a form of bilingual education in schools. In a language immersion program, a second language, (i.e., French) referred to as the target language, is used as a teaching tool. The target language is the language in which children have had no prior training (which is the case for most children who have been allowed to attend English-speaking schools). Activities in and outside of the classroom (e.g., social studies, math, history as well as meals, or every day tasks) are all conducted in the target language. In an early immersion program, students begin the second or target language from the age of 5 or 6 (i.e., kindergarten or grade 1), whereas in a middle immersion program, the immersion of the target language begins at the age of 9 or 10 (i.e., late elementary or grade 4). Finally, in a late immersion program, students begin the second language between the ages of 11 to 14 (i.e., grade 6 onward) (Office Québécoise de la Langue Française, 2008). Considering the fact that the testing in this study was completed in English and it was crucial that children attending elementary schools were fluent not only in speaking and understanding but most importantly in reading in the English language, many of the elementary schools with immersion programs were not suitable for the purposes of this study. For example, in many schools on the lists provided by the two English school boards, the immersion program offered included 85% French instruction and only 15% English instruction during early elementary years. Additionally, in schools with early French immersion (starting in Kindergarten or grade 1), students have done all school work in French, except in English language arts, usually starting between grades 2 and 4. Some of the schools that were contacted also advised against testing in their schools, as children had not begun reading in English. In his review of studies completed on immersion programs and immersion language learners in Canada, Baker (1993) has also confirmed that early immersion students lag behind their monolingual peers in literacy (i.e., reading, spelling, and punctuation) for the first few years. As a result, the selection of schools for this project was limited only to those in which English was the leading (i.e., target) language of instruction during the elementary school years.

School closures. The closure of schools associated with English School Boards was another factor that affected participant recruitment, as this involved two of the schools associated with the two school boards that were contacted. This is not an uncommon event in Quebec. Families with children moving out of the English school board territory is often the reason for the permanent closure of many schools in the English school boards.

Participant attrition. Lastly, another independent study was conducted in conjunction to this research project and a few issues related to this study seemed to affect the recruitment of participants and also led to participant attrition, as explained below. The purpose of the independent study was to explore eye movement during reading, thus requiring eye movement recording equipment and desktop computers. The demands of this independent study were added to those of the present research project (including goals, procedures, and all measures) while permission for recruiting participants was obtained from the university, school boards, schools and parents. Including measures for both studies increased testing time required and the use of proposed equipment in the independent study also required space in schools. These factors may have affected the schools' decision to participate in the present project.

Additionally, data collection for the present study (which occurred in conjunction with the eye movement project) had to be postponed for some time because of technical difficulties related to the eye movement recording equipment. This led to a loss of 62 participants (recruited from two schools) who were ready for data collection but could not be enrolled for testing. The attrition was due to the fact that testing times could no longer be rescheduled in the two schools. This was because in one school teachers were not open to rescheduling since they were involved in different projects later in the year as well as in the next school year. The principal of the other school also informed the examiner that testing could not be rescheduled, as the school was closing permanently. As a result of the fact that the technical difficulties were unable to be resolved easily, a decision was made to pursue the present study independently from the other research project to avoid the loss of more participants who were in the process of being recruited from other schools. All schools and parents involved in this research study were informed that the additional measures related to the eye movement study would be collected from participants at a later point.

Recruitment Process

Participants in this study included elementary school students from grade 1 through grade 5. The required steps that were followed prior to recruiting students included obtaining permission from university Research Ethics Committees, local school boards and individual schools associated with the boards. In the majority of schools that granted authorization, obtaining permission from elementary school teachers was also necessary. Parental consent was the final step required to recruit and enroll participants in this study.

Following these requirements, all five English school boards in the Greater Montreal area were contacted and two provided authorization to conduct the study in their schools. Altogether, 23 schools associated with the two school boards that met the English language criteria were contacted though e-mail or phone. Only one school associated with the first board and four schools associated with the second board provided permission for conducting the study. In sum, the quantity of potential elementary school students whose teachers agreed to participate in the research process was 427 (i.e., 120, 50, 142, 84, and 75 in the five schools, consecutively). The parents of all 427 students were asked for consent to participate in the study. Out of 427 parents in the five schools, 145 provided consent for their children to participate in the study. Out of the 145 children, 62 from two different schools associated with the two school boards could never be enrolled for testing in the study because of technical difficulties related to the equipment utilized for the eye movement project, as explained earlier. From the remaining 83 students from three different schools, all 83 were enrolled for testing but data on only 80 could be used in the study. Data from one student was excluded because of the student's difficulties in understanding and following directions, resulting in the student not completing the entire battery of tests. Two other students could not be tested because of absence from school during the scheduled testing times.

In an attempt to recruit additional participants, the co-director of a summer camp associated with an English Canadian university in Montreal was contacted, and with her permission, all 44 parents of elementary school students who were attending the camp that summer were asked for their consent. Out of these 44 parents, only three provided permission for their children to participate in this research project. All three children were tested, but one child's data could not be used because she could not complete the entire battery of tests required for the study as a result of difficulties in scheduling structured times for testing during the camp period. Further attempts were made to recruit participants through placement of advertisements in two local newspapers that served the Greater Montreal area. The advertisements were placed once in each newspaper. The advertisement placed in the first newspaper did not initiate any response. The advertisement placed in the second newspaper initiated two responses (i.e., one parent with two children and another parent with one child). All three children were enrolled and tested for the study.

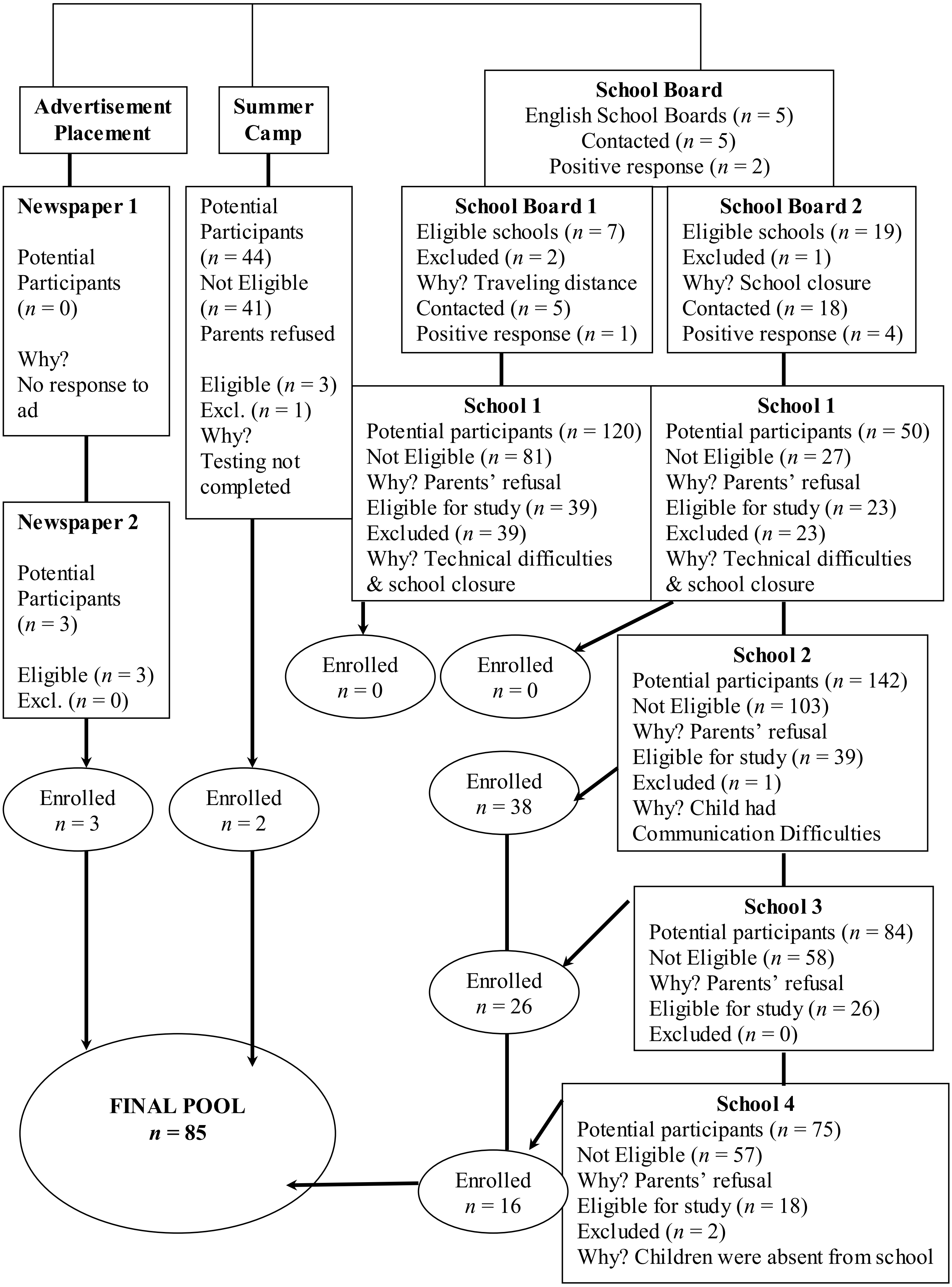
A flow chart illustration of the recruitment process is presented on page 102-103. This flow chart was included following Moher et al.'s (2001) recommended best practices from medical guidelines for sampling and recruitment. Some of the terms used in the flow chart are described in Table 1.

Table 1

Explanation of Terms in Flow Chart of Participants' Recruitment Process

Term	Explanation
Eligible schools	Refers to schools that met the language criteria (i.e., using English as the leading language in early elementary years)
Potential participants	Refers to all elementary students from grades 1 through 5 (in schools that granted authorization for conducting this study) whose teachers had agreed to be involved in the research process
Traveling Distance	Refers to the two schools associated with the first school board that were not contacted due to the fact that they were located in townships which were extremely far and difficult to reach with public transportation
Technical difficulties	Refers to technical difficulties in the eye movement equipment that was utilized in the independent eye movement project that was being conducted in conjunction to the present research during the beginning of recruitment process
School closure	Refers to the two schools associated with the two boards that were closing permanently for reasons presented earlier (e.g., drop of student population as a result of families moving out of the school board's territory)

Figure 5. Flow chart of participants' recruitment process.


Advertisement Placement
Summer Camp
School Board
English School Boards (n = 5)
Contacted (n = 5)
Positive response (n = 2)
Newspaper 1
Potential Participants (n = 0)
Why? No response to ad
Potential Participants (n = 44)
Not Eligible (n = 41)
Parents refused
Eligible (n = 3)
Excl. (n = 1)
Why? Testing not completed
School Board 1
Eligible schools (n = 7)
Excluded (n = 2)
Why? Traveling distance
Contacted (n = 5)
Positive response (n = 1)
School Board 2
Eligible schools (n = 19)
Excluded (n = 1)
Why? School closure
Contacted (n = 18)
Positive response (n = 4)
School 1
Potential participants (n = 120)
Not Eligible (n = 81)
Why? Parents' refusal
Eligible for study (n = 39)
Excluded (n = 39)
Why? Technical difficulties & school closure
School 1
Potential participants (n = 50)
Not Eligible (n = 27)
Why? Parents' refusal
Eligible for study (n = 23)
Excluded (n = 23)
Why? Technical difficulties & school closure
Newspaper 2
Potential Participants (n = 3)
Eligible (n = 3)
Excl. (n = 0)
Enrolled n = 0
Enrolled n = 0
School 2
Potential participants (n = 142)
Not Eligible (n = 103)
Why? Parents' refusal
Eligible for study (n = 39)
Excluded (n = 1)
Why? Child had Communication Difficulties
Enrolled n = 3
Enrolled n = 2
Enrolled n = 38
School 3
Potential participants (n = 84)
Not Eligible (n = 58)
Why? Parents' refusal
Eligible for study (n = 26)
Excluded (n = 0)
Enrolled n = 26
FINAL POOL n = 85
School 4
Potential participants (n = 75)
Not Eligible (n = 57)
Why? Parents' refusal
Eligible for study (n = 18)
Excluded (n = 2)
Why? Children were absent from school
Enrolled n = 16

Final Pool of Participants

In all, the final pool of participants included 85 elementary school students between the ages of 6 and 13 ($M = 7.80$, $SD = 1.40$) attending grades 1 through 5 (28% in grade 1, 42% in grade 2, 20% in grade 3, 4% in grade 4, and 6% in grade 5). Including a wide age range is known to be appropriate for the purposes of a reading-age match study because the sample should include younger average readers as well as older poor and average readers. Thus, a sample drawn from various elementary school classes was appropriate, as it allowed for the anticipated construction of matched subgroups consisting of both a younger reading-age and an older chronologically-age control subgroup matched to older subgroup with reading difficulties. Additionally, it was intended to recruit from a large number of elementary school students in order to increase the power in analysis as has been recommended by Tabachnik and Fidell (2007). Prior to describing the demographic information and all measures collected for this study, the procedure for the collection of measures is explained in the next section of this chapter.

Part 2: Testing Procedure

All participating children completed a brief measure of cognitive functioning and a series of reading, phonological, and rapid naming tasks. Furthermore, all participants engaged in a series of motor performance and cerebellar tasks that could be performed fairly quickly and were not stressful or demanding. The whole procedure took approximately 70 to 90 minutes per child. The principal researcher and examiners who administered the tests were unaware of the children's reading and motor skills.

For participants recruited from individual schools, testing was completed in the school setting in rooms provided by the school. Testing sessions were arranged with teachers and it was agreed that participants would complete the tasks in two sessions. The brief measure of cognitive functioning and reading, phonological and rapid naming tasks were completed in one session that lasted between 45 and 60 minutes. All motor and cerebellar tasks were completed in another session that lasted approximately 30 minutes. The order of tasks completed was counterbalanced both within each session and between sessions. The order between sessions could not always be controlled as time allotted to testing depended on students' schedule at school and their availability. For the few participants recruited from the summer camp, testing was completed at the school in which the camp activities took place where an empty room was available. As a result of the difficulty in scheduling structured testing times, testing could only be completed during the short breaks (i.e., 15- or 30-minute breaks) between camp activities. Finally, for the few participants recruited through the advertisement placement, testing was completed in one 70 to 90 minute session in the university's research laboratory.

Prior to the administration of tasks, an assent form was introduced to all participating children to document their interest in taking part in the study. The content of the assent form was explained to the children in simple terms. It was explained that their participation would help find answers to questions about how children learn to read. They were informed about the different types of activities they were to complete and that they could refuse to participate in any of the activities at any time. The confidentiality of information was also explained. They were also informed that a short report on their reading performance would be provided for their parents or for the school. It was explained that they were not graded on any of the tests they did and the only reason parents or their school would receive these reports was to help children read better. Finally, prior to documenting their interest, all participants were informed that they would receive "a little something" for their help that would be given in the envelope containing the report that their parents receive. It was not explained what the gift was and most did not inquire further about what they would receive. Upon completing the activities, children were also given a few stickers that they selected. They were not told about the stickers prior to completing the tasks.

A brief written performance report was provided to the parents of all of the children who completed the cognitive functioning measure, reading, phonological and rapid naming tasks. For the participants who were recruited from individual schools, the school received a copy of this report only if parents provided their consent. For children who demonstrated reading difficulties, parents also received handouts suggesting different intervention strategies that addressed the specific difficulties. These handouts were also given to schools, if they were interested. All participating children, including those who completed only part of the tasks, were compensated for their time in the form of a gift certificate in the amount of $10.00 that could be exchanged for movie passes.

Part 3: Demographic Information

Overview

The following demographic information was collected for all participants from their parents: (a) ethnic origin, (b) generational status, (c) mother tongue (i.e., first language spoken by the child), (d) first language child used for writing, and (e) bi/multilingualism. Information on the first language spoken and written by the child was also collected from children during testing sessions. Where appropriate, demographic information collected for the sample was compared to corresponding available data from Statistics Canada. Since the year for available data in the Census of Canada did not always correspond to the years in which data collection took place for this study, a constant year was chosen for comparison when statistics data was always available from the Census of Canada.

Ethnic Origin and Generational Status

Ethnic origin. Information on ethnic origin was collected from participants' parents. Table 2 presents the proportion (in percentage) of ethnic origins included in the sample. The ethnic origin for 11% of the participants was unknown, as parents did not provide the information.

Table 2

Proportion of Ethnic Origins in Sample

Ethnic Origin	Percentage
Canadian	41%
Greek	6%
Italian	4%
Portuguese	1%
Jewish	1%
Visible Minorities	
South/South East Asia	22.5%
African American	6%
Hispanics	2.5%
West Asian	3%
Chinese	1%
Mixed Etnicithy[a]	1%
Unknown Ethnic Origin	11%

[a] Specified by parents as mixture of Asian with white parents.

The ratio of ethnicities reported in the sample was compared to the ratios reported in Statistics Canada 2006 for the Greater Montreal Area, using two chi-square analyses. The first chi-square analysis comparing the proportions of Canadian, Greek, Italian, Portuguese, and Jewish ethnicities in the sample with data from Statistics Canada was significant, $\chi^2 (4, N = 85) = 11.30, p < .05$. This indicated that the ratio of all ethnicities, except for Greek and Italian, were comparable to those reported in the Statistics Canada 2006 for the Greater Montreal area. In contrast to Census of Canada, which reported more Italians than Greeks, the sample in this study included more Greeks than Italians. The second χ^2 analysis compared the proportion of visible minorities (i.e., South/South East Asians, African American, Hispanic, West Asian, and Chinese) in this sample to those reported by Statistics Canada 2006 for the Greater Montreal area. This χ^2 was also significant, $\chi^2 (4, N = 85) = 25.74, p < .001$, indicating that the proportion of African Americans, Hispanics, West Asians, and Chinese in this sample were comparable to those reported in Statistics Canada 2006. However, the sample in the present study included considerably more participants with South and South East Asian origin than the reported proportions in Statistics Canada (i.e., 22.5% vs. 3%).

Generational status. Information on generational status is presented in Table 3.

Table 3

Proportion of Generational Status in Sample

Generational Status	Percentage
Participants not born in Canada	5%
Participants born in Canada	95%
First Generation[a]	35%
Second Generation[b]	17%
Third Generation[c]	42%

[a] Child was born in Canada. [b] One or both parents were also born in Canada. [c] One or both parents and/or grandparents were also born in Canada.

Languages Use

Mother tongue or first language spoken by child. Information on the first language spoken by each child as reported by parents and by the child is presented in Table 4. Five participants' parents did not report information on first language.

Table 4

Number of Anglophones, Francophones, and Allophones in Sample

	Parent Report	Child Report
Anglophone	53	54
Francophone	7	7
Allophone[a]	20	24

[a] Allophone means a person whose first language is neither of Canada's official languages of English and French. Reported languages in allophones included Hindi, Gujarati, Punjabi, Tamil, Bangle, Bengali, Twi, Farsi, Urdu, Chinese, Greek, and Spanish.

The proportion of first languages spoken (as reported by parents) was compared to those available from Canada Statistics 2006 for Montreal, using chi square statistics. Among the reported languages for the allophones in the sample, only some (i.e., Punjabi, Chinese, Greek and Spanish) could be entered in the chi square analysis because statistics on other languages were not included in the 2006 Montreal Census. The available languages were entered together as a group in the analysis along with English and French. The chi square was significant, $\chi^2(2, N = 85) = 240.29$, $p < .000$, indicating that the proportion of allophones in the sample was comparable to ratios reported in Canada Statistics for Montreal. However, this sample included more Anglophones or *English-dominant bilinguals* than the reported proportion in Montreal (i.e., 62% vs. 12%) and fewer Francophones or *French-dominant bilinguals* (i.e., 8% vs. 65%). This was expected in the sample, since the present study was intended to target the Anglophone population in Montreal.

Bi/multilingualism. According to parents' reports, 17 of 53 Anglophones or English-dominant bilingual in the sample spoke English as the only language. The remaining Anglophones were reported to be bilingual or in some cases trilingual. That is, in addition to English they either spoke French, a language other than French, or both. All 7 Francophones or French-dominant bilinguals in the sample were reported to be either bi- or trilingual. They all spoke English and a few spoke a language other than English and French. Finally, all allophones in the sample were reported to be bi- or trilingual. They all spoke English and a few spoke both English and French in addition to their mother tongue. English-French bilingualism in Anglophone, francophone and allophones in the sample was compared to the proportions reported in Canada Statistics 2006 for the province of Quebec since data was not available for Montreal. The chi square statistic was not significant (χ^2 (2, N = 85) = 2.90, $p > .05$) indicating that proportion of bilingualism in the sample for this study was comparable to those reported for Quebec.

First language used for writing by child. Information on the first language used for writing as reported by parents and children is presented in Table 5. Information on first language used for writing was not reported by the parents of 4 participants. The reported proportions for dominant language used for writing in the sample could not be compared to Census Canada since no statistics on languages used for writing were available.

Table 5

First Language Used for Writing by Participants

	Number of Children	
	Parent Report	Child Report
English	75	70
French	1	4
Other	5	10

Part 4: Measures

Overview

Data collected from participants in this study included a behavioral rating for attention and hyperactivity, as well as estimates of cognitive functioning, word reading ability, word reading fluency, phonological processing, and rapid naming skills, using standardized and widely used measures. All participants also completed a series of motor and cerebellar tasks. Whenever possible, an estimate of reliability was calculated for measures included in this study to compare to published reliabilities if they were available. In the next three sections, all measures used are described. The first section introduces the behavior rating scale. The second section describes the cognitive, word reading, reading fluency, rapid naming, and phonological processing measures. The final section includes the motor and cerebellar tasks.

Behavioral Measures

To obtain a behavioral measure related to inattention and hyperactivity, a behavior rating scale adapted from the short form of the Conners' Parent Rating Scale – Revised (S) and the short form of the Conners' Teacher Rating Scale – Revised (S) (Conners, 1997) was used. Prior to describing how the Conners' Parent and Teacher Rating Scales (S) were adapted for this study, the original forms are described. The reason for adapting these forms for this study will also be explained. Finally, the distribution of scores that were derived for the sample is presented.

Conners' Parent and Teacher Rating Scales – Revised (S). The short form of the Conner's Parent and Teacher Rating Scales-Revised (Conners, 1997) are screening instruments for assessment of ADHD and related behavioral problems in children and adolescents ranging from 3 to 17 years of age. The child's behavior in the past month is rated on a 4-point Likert scale ranging from “Not True at All (Never, Seldom)” to “Just A Little True (Occasionally), “Pretty Much True (Often, Quiet a bit)” and “Very Much True (Very Often, Frequently). For both of these rating scales, the norms derived (i.e., raw and *T*-scores) are gender and age dependent. The directions in both scales are easy to follow and the items are generally clearly written and easy to understand for anyone at or beyond the tenth grade level (Conners, 1997).

The Conner’s Parent Rating Scale – Short Form includes 27 items and the Conners' Teacher Rating Scale – Short Form includes 28 items. Four subscales can be calculated for both Parent and Teacher Rating Scales which include: (a) Oppositional subscale (b) Cognitive Problems/Inattention subscale, (c) Hyperactivity subscale, and (d) ADHD Index. In both Parent and Teacher Rating Scales, *T* scores derived for each of the four subscales can range from 40 to 90. T scores falling below 45, while slightly atypical, are considered low scores that should not raise concern. *T* scores between 45 and 55 (i.e., 32nd – 73rd percentile) are considered average. *T* scores between 56 and 60 (i.e., 74th – 85th percentile) are considered slightly atypical or borderline which should raise concern. *T* scores between 61 and 65 (i.e., 86th – 94th percentile) are considered mildly atypical reflecting possible significant problems. *T* scores between 66 and 70 (i.e., 95th – 98th percentile) are considered moderately atypical indicating significant problems. Finally, *T* scores above 70 are considered markedly atypical indicating significant problems.

Adapted form used in this study. It is generally recommended to use parent and teacher rating scales in conjunction with one another to obtain a more accurate picture of a child’s behavior in different settings. However, this was not possible in the sample for the present project. In order to avoid making time demands on teachers, for the purposes of this study one rating scale was adapted by integrating the short forms of the Conner’s

Parent and Teacher Rating Scales. This rating scale was intended to be completed by parents. The short versions were used so that parents could easily complete the rating scale in a short period of time (i.e., 5 to 10 minutes). The reported correlations between the short forms of the Conner’s Parent and Teacher Rating Scales have revealed considerable variability, possibly because of some of the differences in the two scales or actual differences in behavior observed at school and home (Conners, 1997). A more accurate picture of the child’s behavior was obtained by integrating the two forms and including items related to both school and home settings.

For the purposes of this study, only items reflecting Cognitive problems/Inattention subscale, Hyperactivity subscale, and ADHD index from both rating scales were integrated to form the adapted Parent Rating Scale. Items reflecting the “oppositional” subscale were not used in order to keep the integrated rating scale as short as the original forms and because the primary focus of this research was on attention. The items reflecting this scale also did not serve the purpose of the study as inattention and hyperactivity subscales and the ADHD index. The adapted form included 28 items, 21 of which were all the items that reflected the Cognitive/Inattention subscale, Hyperactivity subscale, and ADHD index of the original Conners’ Parent Rating Scale-Revised (S). Eleven of these twenty-one items were also items in the original Conner’s Teacher Rating Scale-Revised (S) form. The original Conner’s Teacher Rating Scale – Revised (S) included seven additional items on behavior in group and school settings that also reflected the Cognitive/Inattention subscale, Hyperactivity subscale, and ADHD Index. These items were also added to the adapted form. The instructions for completing the adapted Conner’s Parent Rating Scale was the same one as used in the original Conner’s Parent Rating Scale – Revised (S).

The adapted form was intended to be completed by parents of participants. However, in the course of participant recruitment from one of the local schools, the principal recommended that teachers, instead of parents, complete the child's behavior rating scale. Thus for 28 of 85 participants in the sample, teachers instead of parents completed the behavior rating scale. Any potential bias that might be related to this approach was investigated formally and reported in the Results chapter.

Distribution of scores derived for sample. When computing *T*-scores and percentile ranks for the adapted Conner's Parent Rating Scale, if parents completed the form, scoring males and females in the sample was based on parental norms using only the 21 items that were part of the original Conner's Parent Rating Scale – Revised (S). When the 27 forms completed by teachers were computed, scoring males and females was based on teachers' norms using only 18 items that were part of the original Conner's Teacher Rating Scale – Revised (S). As indicated earlier, eleven of these 18 items were the same as the items in the Conner's Parent Rating Scale – Revised (S).

A Cognitive/Inattention subscale, a Hyperactivity subscale, and an ADHD index score was derived for all participants who were rated by their parents. However, for those participants rated by their teachers, only a Hyperactivity subscale and an ADHD index could be derived. A Cognitive/Inattention score could not be obtained for these participants because of the fact that the items related to this scale were taken from the original Conners' Parent Rating Scale – Revised (S) and thus could not be used in scoring the forms rated by teachers.

Taking into consideration that deriving *T* scores for the different subscales did not include all the corresponding items in scale and depended on whether parents or teachers rated the scale, only the ADHD index raw scores were used in the analyses conducted for this study in order to keep a measure that was consistent across the sample. The reasons for choosing the ADHD index as the target measure was because this index has been reported as an effective screening measure to identify children and adolescents meeting the ADHD diagnostic criteria (Conners, 1997). The index has also been used in other studies that have investigated reading and motor-cerebellar related measures in children (e.g., Raberger & Wimmer, 2003; Wimmer et al., 1999). The distribution of the ADHD index raw scores obtained for the sample in the present study as they corresponded to *T* scores is presented in Table 6.

Table 6

Distribution of ADHD Index Raw Scores in Sample

	Percentage
Below Average to Average (40-55)	60%
Slightly Atypical (56-60)	12%
Mildly Atypical (61-65)	10%
Moderately Significant (66-70)	8%
Markedly Significant (> 70)	10%

Reliability estimates for the adapted rating scale. The Spearman-Brown odd-even reliability coefficient calculated for the sample in this study was .85 for the Hyperactivity subscale, .81 for the Cognitive/Inattention subscale, and .91 for the ADHD index. The internal consistency coefficient alpha for the sample was .92 for Hyperactivity, .90 for

Cognitive/Inattention, and .98 for ADHD index. Spearman-Brown odd-even reliability and internal consistency coefficient alphas were also calculated separately for Hyperactivity scale and ADHD index for ratings that were completed by parents and those completed by teachers. As presented in Table 7, the reliabilities were similarly high for both ratings. Overall, the reliability estimates obtained for the sample and those obtained separately for forms completed by parents and teachers were similar to the published reliability estimate ranges reported for the Conners' Parent Rating Scale-Revised (S) (i.e., .86-.94) and Conners' Teacher Rating Scale-Revised (S) (i.e., .88-.95).

Table 7

Spearman Brown Odd-Even Reliability and Internal Consistency Coefficient Alphas for Ratings Completed by Parents and Teachers

	Odd-Even		Cronbach Alpha	
	Hyperactivity	ADHD index	Hyperactivity	ADHD Index
Forms completed by Parents	.82	.88	.90	.94
Form completed by Teachers	.93	.97	.94	.98

Cognitive, Reading, Fluency, Rapid Naming and Phonological Measures

Cognitive functioning. Including an IQ measure is usually recommended in studies that involve reading assessment, mainly for methodological reasons and not necessarily to support the discrepancy definitions of dyslexia (McPhillips, 2003). The two-subtest form of the Wechsler Abbreviated Scale of Intelligence (WASI)

(Psychological Corp, 1999) was used to obtain an estimate of participants' cognitive functioning. The WASI is a brief measure of cognitive functioning designed for ages ranging from 6 to 89 and it consists of four subtests, two verbal and two non-verbal tests. The two verbal subtests (Vocabulary and Similarity) yield an estimate of verbal intelligence, and the two nonverbal subtests (Matrix Reasoning and Block Design) yield an estimate of performance or nonverbal intelligence. A full scale can be derived from all four subtests or from only two subtests (i.e., Vocabulary and Matrix Reasoning), which can be administered in approximately 15 minutes. The *Vocabulary* subtest of the WASI is a test of expressive vocabulary that includes four pictures and 37 words, measuring verbal knowledge, memory, learning ability, as well as crystallized and general intelligence. Participants were required to provide definitions of words pronounced by the experimenter. The *Matrix Reasoning* subtest, consisting of 35 items, is a measure of non-verbal fluid reasoning and general intelligence. In this test, children selected an option from five choices that best completed a visual pattern.

While the WASI does not provide a comprehensive cognitive assessment, it is considered adequate to use for cognitive screening (Psychological Corp, 1999). The correlations between this form of the WASI and the comprehensive test of cognitive functioning in the Wechsler series, namely the Wechsler Intelligence Scale for Children-Third Edition (WISC-III) are reported to be .82 for Verbal IQ, .76 for Performance IQ, .82 for the 2-subtest Full Scale IQ, and .87 for the 4-subtest Full Scale IQ (Saklofske, Caravan, & Schwartz, 2000). The published average reliability estimates reported for the WASI standardization sample (i.e., 1100 children between the ages of 6 to16) are .89 for the Vocabulary subtest, .92 for the Matrix Reasoning subtest, .93 for Verbal Scale, .94 for nonverbal scale, and .93 for the 2-subtest Full Scale Intelligence (Psychological Corp, 1999).

For the sample in this study, the Spearman-Brown odd-even reliability coefficient for the Vocabulary subtest was .78. The internal consistency coefficient alpha was .88, which is similar to the published reliability estimate of .89 reported for the standardization sample. The Spearman-Brown odd-even reliability coefficient calculated for the Matrix Reasoning subtest for the sample was .85. The reliability coefficient calculated, using Cronbach alpha, was .92, which is the same as the published reliability reported for the standardization sample.

Reading Measures

Word reading. The *Word Identification* subtest of the Wide Range Achievement Test-Third edition (Wilkinson, 1993), which screens basic reading skills in examinees ages 6 to 75, was used. The test consists of 15 letters of the alphabet and 42 individual words out of context ordered in decreasing fluency and increasing complexity. Depending on age, an examinee is asked to either begin with pronouncing the letters of the alphabet or the list of words. The published median reliability estimates reported for the WRAT-3 reading subtests is reported to range from .90 to .95 for all ages used in the 4433 standardization sample (with 100 individuals in each age band). For the sample in this investigation, the Spearman-Brown odd-even reliability coefficient for the Word Identification subtest was .91 and the reliability coefficient, using the internal consistency coefficient alpha, was .95, which is similar to the published reliability reported for the standardization sample.

Reading accuracy and fluency. As mentioned by Torgesen (2002), after reading instruction begins in first grade, direct assessment of fluency and accuracy in word reading is recommended as the best way to identify children who have fallen behind in these skills. To measure word reading accuracy and fluency, the Test of Word Reading Efficiency (TOWRE) (Wagner, Torgesen, & Rashotte, 1999) was used. This test is designed for administration to ages ranging from 6 to 25 years. It consists of two subtests: (a) the *Sight Word Reading*, which includes 104 real words ordered in decreasing word frequency and increasing orthographic complexity; and (b) *Phonemic Decoding* that includes 63 non-words or pseudowords ordered by increasing orthographic complexity and length. The two subtests measure the number of real words and non-words read accurately by the examinee within 45 seconds. Examinees read a short list of words and non-words for practice prior to beginning each subtest.

The published test-re-test reliability reported for the TOWRE ranges from .83 to .93. The published average reliability estimate reported for the TOWRE exceeds .90 for the standardization sample, which includes more than 1500 individuals with ages ranging from 6 to 25. A test-retest reliability estimate for TOWRE tasks could not be obtained for the sample in this study because participants were only tested once.

Phonological Measures

Phonological awareness. The *Elision* subtest of the Comprehensive Test of Phonological Processing (CTOPP) (Wagner et al., 1999), which consists of 20 words, was used to obtain an estimate of participants' phonological awareness skills. The test can be administered to children aged 7 to 24. The CTOPP also includes a Blending subtest as part of the phonological awareness cluster, which was not used in this study. Between the two tasks, the Elision subtest has stood out as a key element of phonological awareness and a consistently powerful measure to identify children with reading difficulties (French et al., 2008). This task has been suggested to be a powerful tool to identify children with reading difficulties and thus an important element in screening for reading success (French et al., 2008). Between the two tasks of Blending and Elision, Elision may also be less contaminated since Elision, unlike blending, is not a generally taught skill (Savage, Abrami, Hipps, & Deault, 2009). Considering these factors and given the length of the testing battery used in this study and the limited testing time available in school settings, only the Elision and not the Blending subtest was included in the battery.

For the Elision subtest, the examinee was asked to repeat the word, minus a single sound, which could be in the initial or middle position (e.g., say *cup* without the *k* sound or say *powder* without the *d* sound). Responses were scored for accuracy. Practice items were provided prior to administering the test items. The published reliability estimate reported for the Elision subtest of the CTOPP is .93. The standardization sample for the CTOPP includes 1,656 individuals, with 76 to 155 students included in each age range (with greater age representation in the youngest age ranges). For the sample in this study, a similarly high reliability was obtained with the Spearman-Brown odd-even reliability coefficient of .91 and the internal consistency coefficient alpha of .95.

Phonological recoding (non-word reading). The *Word Attack* subtest of the Woodcock Reading Mastery Tests-Revised (Woodcock, 1987; Woodcock & Johnson, 1989) was administered to all participants. The test can be administered to ages 5 to 75 and it measures the ability to apply structural and phonetic skills to pronounce unfamiliar nonsense words, similar to a situation in which one would encounter unfamiliar real words. It consists of a list of 45 nonsense words or words with a very low frequency of occurrence in the English language, which increase in order of complexity. Prior to administering test items, two practice items are provided. Of the 45 items in the Word Attack test, only 11 are polysyllable non-words (i.e., two-, three-, and four-syllable stimuli), while the rest are monosyllables, meaning that not all participants had a chance to read every polysyllable non-word depending on when the ceiling was reached. Despite this fact, group and individual differences on polysyllables were also investigated, taking into consideration that the systemetatic review of findings related to non-word reading indicated that performance on more complex non-words may be more likely in identifying a deficit in individuals with dyslexia (Rack et al., 1992).

The published reliability estimate reported for the WRMT-R subtests, including the Word Attack subtest, is above .90 for over 3000 individuals included in the standardization sample. Similar to the published reliability estimate, the Spearman-Brown odd-even reliability coefficient calculated for the Word Attack subtest for the sample in this study was .94 and the reliability estimate using a consistency coefficient alpha was .97.

Rapid Naming Measures

Three of the four rapid naming subtests of the Comprehensive Test of Phonological Processing (CTOPP) (Wagner et al., 1999) were used to obtain an estimate of participants' rapid naming skills, which included the *Rapid Digit, Letter, and Object Naming* tests. These tests can be administered to ages 5 through 24. The rapid naming tests are related to efficiency in activating name codes from memory of verbal material and measure the fluid access to and efficient retrieval of these verbal names (i.e., names of digits, letters, or objects in this case) in isolation or as a part of a series (Wagner et al., 1999). Generally, the alphanumeric rapid naming tasks have been more often related to reading difficulties (Bowers et al., 1988; Savage, Pillay et al., 2007; Snyder & Downey, 1995; Wolf, 1982). However, following some other studies a non-alphanumeric task, namely object naming, was included in measures for contrast (e.g., Catts, Gillispie, Leonard, Kail, & Miller, 2002; de Jong & Share, 2007). The CTOPP also includes the *rapid color naming* test which was not used in this study. One reason for excluding this task was to avoid some of the possible confounds associated with it, such as difficulties in the perception of color (e.g., Roessner, et al., 2008; Tannock, Banaschewski, & Gold, 2006).

For the rapid number, letter and object naming tasks which were used in this study, examinees were presented with two pages (one page at a time) that contained four rows and nine columns of either six randomly arranged digits (i.e., 2, 3, 4, 5, 7, and 8), letters (i.e., a, c, k, n, s, t), or objects (i.e., pencil, star, fish, chair, boat, key). Examinees were asked to name the stimuli (i.e., numbers, letters, and objects) as quickly as possible, starting on the top row, from left to right, and moving to the next row, and so on, until all numbers were named. In each of the rapid naming tasks, the total number of seconds taken to name all of the digits, letters, and objects on both pages presented was measured. The published test-retest reliability estimates for CTOPP subtests are reported to range from .74 to .97. A test-retest reliability estimate for rapid naming tasks could not be obtained for the sample in the present study because participants were only tested once.

Motor and Cerebellar Tasks

Motor Tasks

Two tasks, *peg moving* and *bead threading*, were used to measure fine motor skills as suggested by Fawcett and Nicolson (1995c).

Peg moving. The *peg moving* task, known to be sensitive to cerebellar damage (Haggard, et al., 1995; Miall & Christensen, 2004), was taken from Fawcett and Nicolson's (1995c) study, in which they used a commercially available children's pegboard resembling that used in Annett's study (1985 cited in Fawcett & Nicolson, 1995c). Similar to Fawcett and Nicolson's (1995c) study, the pegboard used in this study consisted of 10 rows of 10 holes, but was somewhat larger (10 x 10 in). Prior to beginning the task, the experimenter filled the top row of the board with pegs. Children were asked to hold the board steady with their non-preferred hand and move the pegs with their preferred hand (i.e., the hand they used for writing) as quickly as possible, jumping over the empty row into the third row of holes. They were then asked to move the pegs further 2 rows down the board, and so on until finally 5 rows were completed. Prior to beginning the task, moving pegs (as instructed above) was first demonstrated for children by the experimenter.

For this task, the experimenter instructed the children to pick up only one peg at a time. The trial was restarted if a child picked up more pegs. Children were also asked to ignore pegs that fell off the board. Following the procedure used by Savage and Frederickson (2006), for each row the number of pegs placed as well as the time from which the child touched the first peg until he/she released the last peg was recorded with a stopwatch in order to measure possible speed-accuracy trade off. The dependent variable was the speed with which the task was completed over five trials (i.e., mean time to complete five trials). Cerebellar dysfunction was expected to lead to a longer time required to complete the task (Fawcett & Nicolson, 1995c).

The published test-re-test reliability reported for the peg moving task using 53 adult students (i.e., correlation between mean peg moving times measured on two occasions separated by an interval of 6 to 18 months) is .69 (Annett, Hudson, & Turner, 1974). It should be noted, however, that the purpose of the peg moving task in Annett's studies has been different than the purpose in Fawcett and Nicolson's study (1995c). For example, Annett has used the difference in peg moving time for the two hands to look at handedness correlates of intellectual functioning and disability.

For the sample in this research, a test-re-test reliability could not be obtained since participants completed the pegboard task once. However, the average correlation obtained among mean peg moving times for the five trials was .69, which is the same as the published test-retest reliability reported by Annett et al. (1974) for 53 adults. Using the last two trials of the peg moving test, an internal consistency alpha coefficient was also calculated for the sample, $r = .81$.

Bead threading. The *bead threading* task was taken from the Dyslexia Screening Test (Fawcett & Nicolson, 1996). Children were required to hold an 85 cm string (3 mm in diameter) vertically in the hand they use for writing while standing up. They were then asked to take beads (with their other hand), one at a time, from a basket containing 12 round wooden beads (each 4 cm in diameter with a hole of approximately 0.5 cm) and thread them on the string (3 mm in diameter) as quickly as possible. Prior to beginning the task, the experimenter demonstrated this for participants by threading one bead and explaining the steps involved. Participants were then given a chance to practice threading two beads on the string. The two beads threaded by participants were taken off of the string and put back in the container before beginning the task. If a bead was dropped during the task, participants were asked not to pick it up and to continue threading with another bead from the container. If they dropped the string during the task, participants restarted the task. The dependent variable in the bead threading task was the number of beads threaded in 30 seconds starting from the time they touched the first bead. Children with no cerebellar deficit were expected to thread more beads in 30 seconds than children with a mild cerebellar deficit (Fawcett & Nicolson, 1995c). The published test-re-test reliability reported for the bead threading task is .76. No test-retest or other form of reliability estimate could be calculated for this test for the present sample since the test was administered only once and in a single trial.

Cerebellar Tasks

The cerebellar tasks, which included *the postural stability, the muscle tone, and the toe tapping tasks*, were similar to those used by Fawcett and Nicolson (1996; 1999). Fawcett and Nicolson replicated these tasks from Dow and Moruzzi's (1958) clinical cerebellar test battery. As Fawcett and Nicolson (1999) suggested, these tasks were representative of the three categories of the usual cerebellar test battery (i.e., maintenance of posture, muscle tone, and complex movements). As will be explained, an adaptation of the measurement of postural stability and the muscle tone tasks was attempted in order to increase their sensitivity and if possible their reliability. The three cerebellar tasks are described below. Regarding the postural stability and muscle tone tasks, first the original procedure used by Fawcett and Nicolson (1996; 1999) is explained and then the procedure which was adapted for this study is described.

Postural stability. The materials and administration procedure for the *postural stability* task used in this study followed Fawcett and Nicolson's (1996) Dyslexia Screening Test as well as Fawcett and Nicolson's study (1999) in all regards. The only difference was in the measurement of the task, (i.e., degree of sway,) as will be described starting on page 131. In the original procedure, as explained by Fawcett and Nicolson (1996; 1999) an examinee is asked to stand upright, looking straight ahead, arms at their sides. The shoes are kept on unless they are likely to cause difficulties with balance. Next, the examiner stands behind the examinee and explains that he/she will be pushed gently in the back, while blindfolded, and that the examinee should try to stand still.

The device used to push the examinee is referred to as a balance tester (Fawcett & Nicolson, 1996). As pictured in Figure 6 and explained by Fawcett and Nicolson (1996), the balance tester is a plastic device with a collar that slides on the internal shaft. The collar has a felt washer to control the friction.

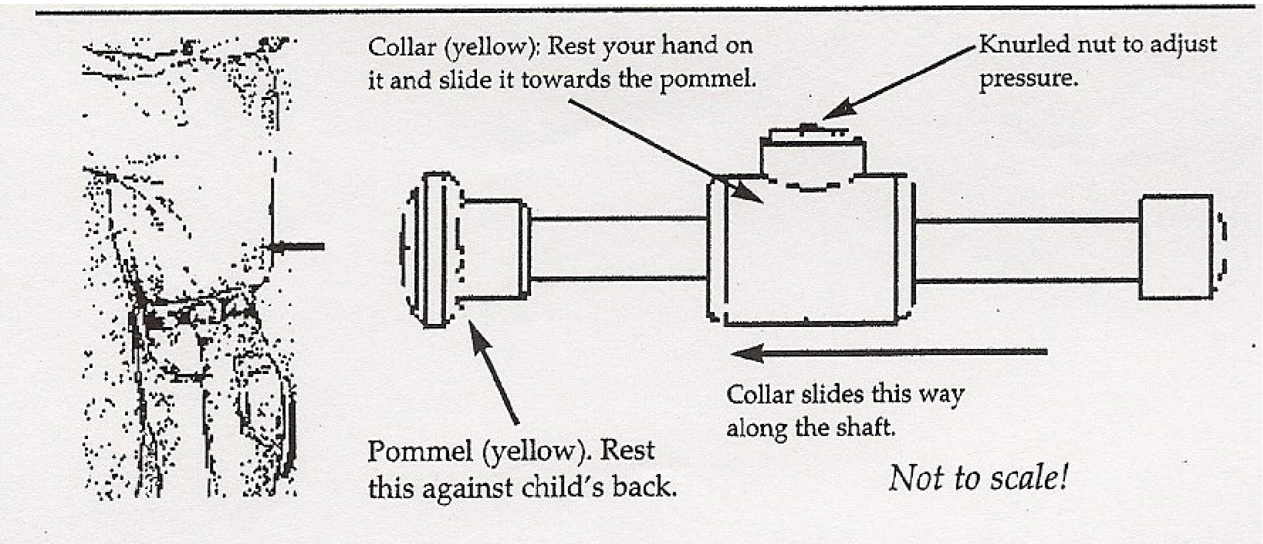


Figure 6. Balance tester in the Dyslexia Screening Test's (DST) postural stability task from Fawcett and Nicolson's (1996) DST manual.

According to Fawcett and Nicolson (1996), by using the washer, the balance tester can be calibrated (e.g., on a kitchen scale) to provide a controlled amount of force. Fawcett and Nicolson (1996) have recommended using a 2.5 kg force to gently push examinees younger than 11.6 years and a 3 kg force to gently push examinees older than 11.6 years.

To gently push the examinee, the examiner holds the collar of the balance tester in their preferred hand and rests the pommel "two vertebrae above the small of the child's back" (Fawcett & Nicolson, 1996, p. 46). The collar slides when sufficient force is applied, while controlling the force. As described by Fawcett and Nicolson (1996), the examiner slides the collar toward the pommel to apply the pressure and push the examinee. Pushing is stopped just before the collar meets the pommel (Fawcett & Nicolson, 1996).

Once the examinee is back into position, he/she is pushed again after 5 seconds using the above procedure. In all, while the examinee is in the upright position with arms at his/her side, he/she is pushed three times. The examinee is then asked to put his/her arms straight out in front (like a sleep walker). Using the same procedure, the examinee is then pushed another three times (Fawcett & Nicolson, 1996).

Each time after the examinee is pushed, the degree of sway is assessed and recorded on a 3-point Likert Scale. A score of zero indicates “good performance,” while a score of one is assigned to “small movement.” Finally, a score of two is assigned to examinees, who after being pushed, step forward or overbalance (Fawcett & Nicolson, 1999). In their study, Fawcett and Nicolson (1999) recorded the degree of sway after each pushing trial. They also checked by videoing a sample of the participants and obtained independent ratings by trained observers who were unaware of the participants’ group. According to Fawcett and Nicolson (Fawcett & Nicolson, 1999), the inter-rater reliability reported between raters varied from .94 to .98 in their study.

As indicated earlier, the material and administration procedure for the postural stability task in the present study followed Fawcett and Nicolson’s (1996) Dyslexia Screening Test and Fawcett and Nicolson’s study (1999). In the same way as described by the authors, a gentle balance challenge was administered to each participant while they were blindfolded, using the aforementioned balance tester. It should be noted that three children in the present study did not like wearing a blindfold and thus were asked to keep their eyes closed during the postural stability tasks. Consequently, the experimenter could not observe whether the child’s eyes were kept shut during the administration of the balance challenge because she was standing behind the child. There were also no other observers in the room. As in Fawcett and Nicolson's study (1999), the balance challenge was administered to each participant across six trials. In three of these trials, participants stood upright, looking straight ahead, and had their arms at their sides. In the other three trials, they stood upright looking straight ahead with their arms straight out in front. To gently push each participant, the pressure was applied to the small of the back for 1.5 seconds and then released (Fawcett & Nicolson, 1999). As recommended by Fawcett and Nicolson (1999), the balance tester was calibrated on a kitchen scale to apply a steady 2.5 kg pressure if the participant was younger than 11.6 years, and 3 kg pressure if the participant was older than 11.6 years. The balance tester was calibrated prior to each session.

The way the degree of sway was assessed in each participant was what was different in the postural stability procedure in the present study. As indicated previously, the degree of sway was assessed and recorded using a Likert Scale in Fawcett and Nicolson's study (1999) as well as in Fawcett and Nicolson's Dyslexia Screening Test (1996). In this study, I attempted to obtain a more sensitive measure by assessing and recording the degree of sway by means of a PASCO 3-axis acceleration sensor.

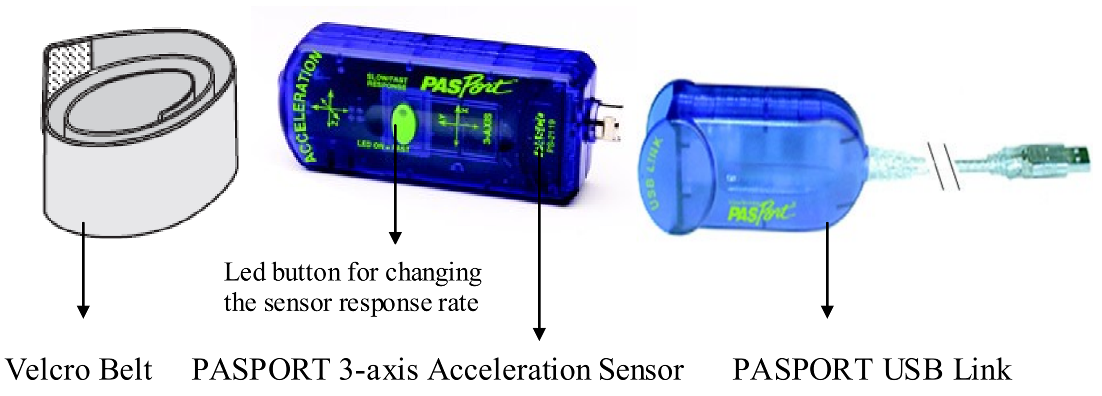


Figure 7. PASPORT 3-axis acceleration sensor, Pasport USB link, and Velcro belt used for recording and measurement of rate of change in position in postural stability task.

The PASCO 3-axis acceleration sensor, as shown in Figure 7, is a device for measuring and recording the acceleration or rate at which an object's velocity (i.e., rate of change in position) changes with time. With a 3-axis acceleration sensor, the magnitude and direction of acceleration can be measured 100 times per second in three dimensions or axes (i.e., x, y, and z). The acceleration recorded is the ratio between the change in velocity (i.e., rate of change in position) measured in meters per second, and time for the velocity to change measured in seconds, as shown by the formula below. Thus, the acceleration is measured in meters per second per second.

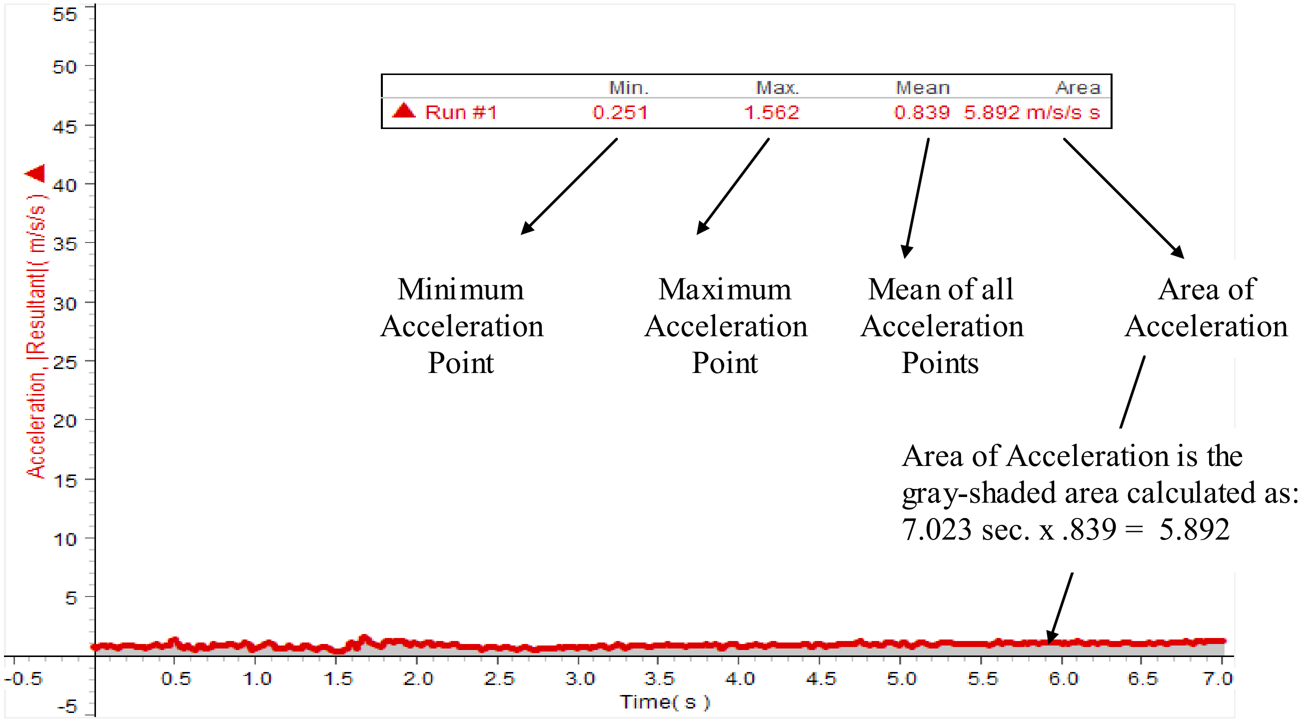
$$\text{Average Acceleration} = \frac{\text{Velocity Change } (\Delta v)}{\text{Elapsed Time } (\Delta t)}$$

According to the manufacturer of the product, the PASCO Foundation, the sensor is sensitive and can measure accelerations ranging up to 5 times the earth's gravitational field with an accuracy of .01 g (g = acceleration of gravity, 9.8 m/s^2). The response rate button on the sensor, as shown in Figure 7, is used in experiments in which the acceleration changes quickly. Pressing the button enables fast response mode. This mode was used to record accelerations in postural stability tasks.

To record accelerations, the sensor is connected to a USB Link, shown in Figure 7. The USB Link is then connected to a USB port on a computer. The data is recorded by the 3-axis acceleration sensor and then analyzed directly through a data collection and analysis interface (i.e., “PASCO Data Studio Data Collection and Analysis Software”) installed on the computer. The data collection and analysis software launches automatically when it detects the PASPORT sensor. The sensor itself is attached to a Velcro belt, which is secured to the hip of the participant. In the present study, prior to securing the accelerometer to the participant’s hip, the experimenter showed the equipment to participants and explained, while demonstrating on herself, how the belt and the accelerometer were attached to the body and how the data was recorded in the computer after they were gently pushed in the back. It was explained to them that they wouldn’t feel anything during the data recording, as the process was completely silent and painless.

Following Fawcett and Nicolson’s approach (1999), the first 3 trials were administered with the participant’s arms at his/her sides, and the second 3 trials with their arms straight out in front. During each trial, accelerations were recorded over a period of approximately 7.023 seconds. In each trial, the balance challenge was administered at some point after 1 second of recording. Recording of accelerations continued until 7 seconds had passed. As indicated earlier, the degree of sway was recorded by the sensor at a rate of 100 times per second. The recordings were the sum of magnitude and direction of accelerations in all three axes. In the postural stability task, since participants were gently pushed in the back, the only axis involved was vertical (i.e., the Y-axis).

The accelerations recorded over the entire 7-second period were graphed and analyzed simultaneously in the PASCO Data Collection and Analysis Software. To illustrate, Figures 8 and 9 show graphic representation and data recorded for 2 participants upon the balance challenge.



	Min.	Max.	Mean	Area
▲ Run #1	0.251	1.562	0.839	5.892 m/s/s s

Minimum Acceleration Point

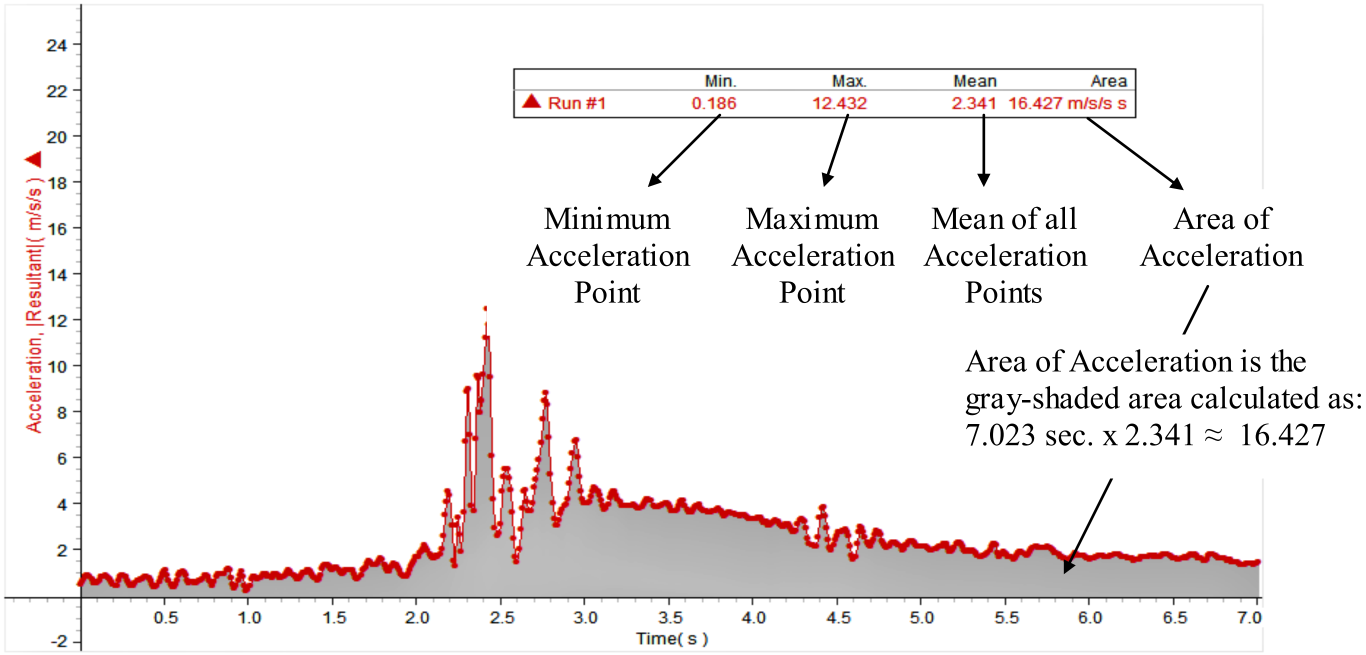
Maximum Acceleration Point

Mean of all Acceleration Points

Area of Acceleration

Area of Acceleration is the gray-shaded area calculated as: 7.023 sec. x .839 = 5.892

Figure 8. Graphic representation and data recorded during postural stability trial for participant with arms at sides.



Minimum Acceleration Point

Maximum Acceleration Point

Mean of all Acceleration Points

Area of Acceleration

Area of Acceleration is the gray-shaded area calculated as: 7.023 sec. x 2.341 ≈ 16.427

Figure 9. Graphic representation and data recorded during postural stability trial for participant with arms straight out in front.

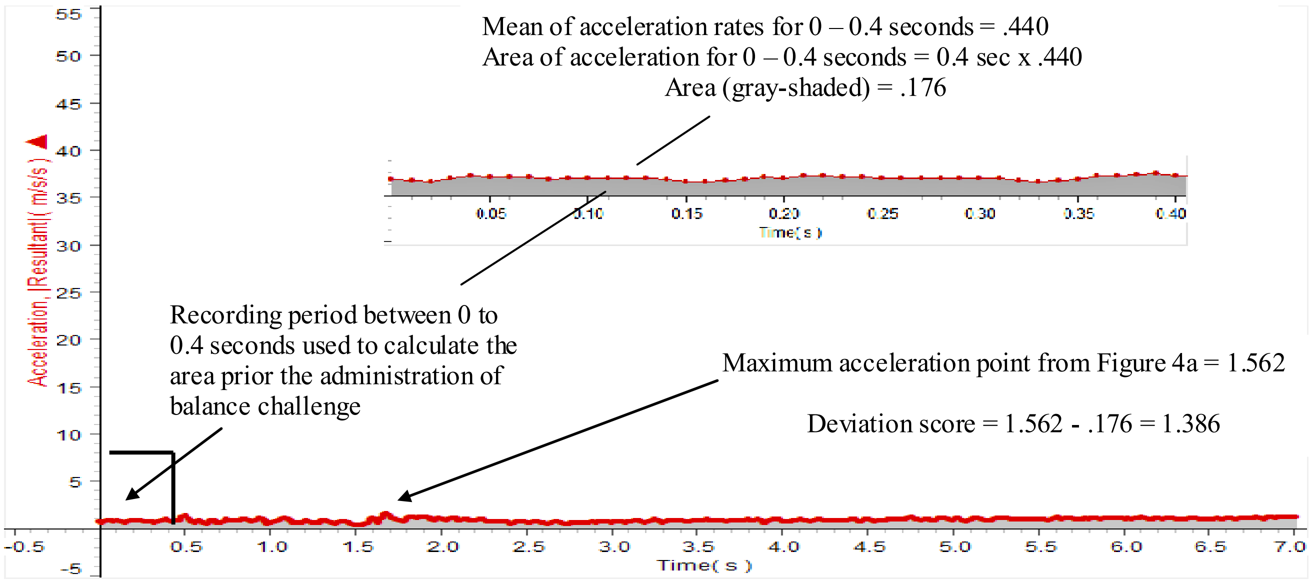
Figure 8 demonstrates a graphic representation of accelerations recorded for a participant with arms held at their sides that remained more or less stable and steady upon the balance challenge. Figure 9 illustrates a graphic representation of accelerations recorded for a participant with arms held straight out in front that lost balance and stepped over in the balance challenge. In both figures, the X-axis represents the time period, in seconds, during which accelerations were recorded for participants and the Y-axis represents the magnitude of accelerations recorded in meter per second per second. At a sampling rate of 100 recordings per second, over 700 acceleration recordings were made for each participant in the sample for each of the postural stability trials.

In each trial, the recording of accelerations began prior to administration of the balance challenge (i.e., from 0 to approximately 1.5 seconds) and continued during and after the administration of the balance challenge. As shown in both figures, part of the summary data reported by the PASCO Data Collection and Analysis Software included the minimum and maximum acceleration points, the mean of all accelerations recorded, as well as the area of acceleration for the entire recording period (i.e., approximately 7.023 seconds). As illustrated in Figures 8 and 9, the area of acceleration is the gray-shaded area. This area is calculated as the area of a rectangle by multiplying the length of the rectangle, which is the total recording time (i.e., 7.023 seconds), by the height of the rectangle, which is the mean of all accelerations recorded over 7.023 seconds.

The data recorded through the PASCO Data Collection and Analysis Software was also used to obtain a measure for degree of sway in participants after the administration of the balance challenge. As recommended by Dr. Paul Stapley (personal communication, October 6, 2008), an Assistant Professor of the Department of Kinesiology and Physical Education with areas of expertise in posture, movement, balance, and motor control, the most appropriate way to measure the degree of sway in this task was to calculate a deviation score for each participant. The deviation score was obtained by subtracting the area of acceleration recorded prior to administering the balance challenge from the maximum acceleration point recorded upon administration of the balance challenge. A deviation score was recommended over other measures calculated by the PASCO Data Analysis Software, such as the mean or area of acceleration for the entire recording period, as these measures were more likely to underestimate the true degree of sway or loss of balance experienced by each participant.

$$\text{Deviation score} = \begin{array}{c}\text{Area of acceleration prior}\\ \text{to balance challenge}\end{array} - \begin{array}{c}\text{Maximum acceleration point after}\\ \text{balance challenge}\end{array}$$

As indicated earlier, for each participant, the time period between 0 to approximately 1.5 seconds included the recordings prior to the administration of the balance challenge. To keep the measure consistent for each participant, the recording area between 0 to 0.4 seconds, shown in Figures 10 and 11, was chosen as the area prior to administration of the balance challenge.



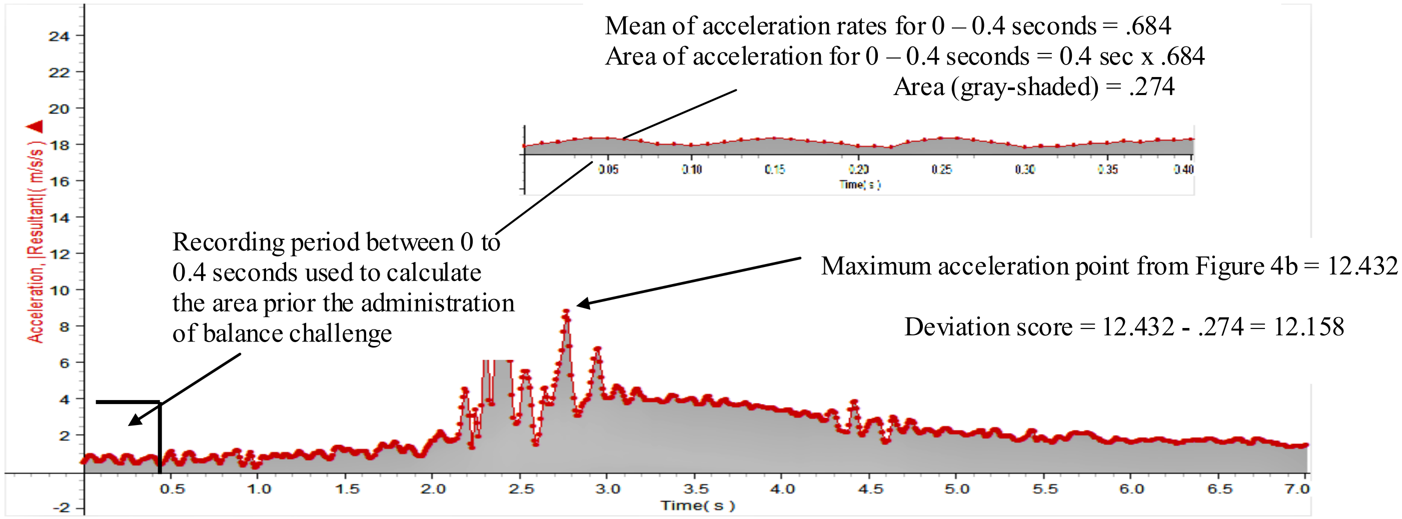
Mean of acceleration rates for 0 – 0.4 seconds = .440
Area of acceleration for 0 – 0.4 seconds = 0.4 sec x .440
Area (gray-shaded) = .176

Recording period between 0 to 0.4 seconds used to calculate the area prior the administration of balance challenge

Maximum acceleration point from Figure 4a = 1.562

Deviation score = 1.562 - .176 = 1.386

Figure 10. Graphic representation of data recording for the participant with arms at sides (from Figure 8) with illustration of recording period prior to administration of challenge (i.e., 0 – 0.4 Sec.).



Mean of acceleration rates for 0 – 0.4 seconds = .684
Area of acceleration for 0 – 0.4 seconds = 0.4 sec x .684
Area (gray-shaded) = .274

Recording period between 0 to 0.4 seconds used to calculate the area prior the administration of balance challenge

Maximum acceleration point from Figure 4b = 12.432

Deviation score = 12.432 - .274 = 12.158

Figure 11. Graphic representation and data recorded for the participant with arms straight out in front (from Figure 9) with illustration of recording period prior to administration of challenge (i.e., 0 – 0.4 Sec.).

This was to control for any forward sway that participants might have engaged in as a result of the anticipation of the balance challenge since they were likely to learn the procedure of the task over the trials and expect to be pushed (P. Stapley, personal communication, October 6, 2008). The area of acceleration for the period of 0 to 0.4 seconds was calculated as the area of a rectangle, which was described earlier. That is, the length of this rectangle, which was the time period of 0.4 seconds, was multiplied by the mean of accelerations recorded over this time period. This time period included 40 acceleration recordings. Consequently, the mean of accelerations recorded for this period was the average of all 40 acceleration rates. For example, the average of 40 acceleration rates recorded between 0 and 0.4 seconds was 0.440 for the participant in Figure 10 and 0.684 for the participant in Figure 11. Thus by multiplying the mean acceleration by the time period of 0.4, the acceleration area obtained was 0.176 (i.e., 0.440 x 0.4) for the participant in Figure 10, and 0.274 (i.e., 0.684 x 0.4) for the participant in Figure 11.

To obtain the deviation score as a measure of degree of sway for each participant, the area of acceleration calculated for the time period of 0 to 0.4 seconds was subtracted from the maximum degree of acceleration rate recorded upon administering the balance challenge. For example, for the participant in Figure 10, the maximum acceleration rate recorded upon administering the balance challenge was 1.562, thus resulting in a deviation score of 1.386 (i.e., 1.562 – 0.176). Similarly, for the participant in Figure 11, the maximum acceleration rate recorded upon administering the balance challenge was 12.432, thus resulting in a deviation score of 12.158 (i.e., 12.158 – 0.274).

For each participant, three deviation scores were obtained for the postural stability task with participants’ arms held at their sides and three deviation scores were obtained for the postural stability task with participants’ arms held straight out in front. In each set of trials, the final score was the average of the three deviation scores calculated over the three trials. Similar to what was predicted in Fawcett and Nicolson's (1999) study, children with no cerebellar deficit were predicted to generate an overall lower score (i.e., less degree of sway and smaller deviation score, greater stability) than children with cerebellar deficit.

For the present sample, the average correlation obtained among the mean deviation scores calculated for the three postural stability trials with arms held at their sides was .61. The internal consistency coefficient alpha obtained for deviation scores for postural stability with arms held at their sides was .79 for all three trials and .70 for the last two trials. Furthermore, the average correlations obtained among the mean deviation scores calculated for the three postural stability trials with arms held straight out in front was .18. The internal consistency coefficient alpha obtained for deviation scores for the postural stability with arms straight out in front was .35 for all three trials and .24 for the last two trials. Because of the fact that correlations and internal reliability for the postural stability task with arms straight out in front was very small, correlations and reliability scores were also obtained for the acceleration areas that were calculated for the postural stability tasks by the PASCO Data Analysis Software for the entire recording period. The average correlation obtained among the mean acceleration areas calculated for the three postural stability trials with arms straight out in front was .43. The internal consistency coefficient alpha obtained for postural stability with arms straight out in front was .69 for all three trials and .57 for the last two trials. The reliability scores obtained for the sample could not be compared to a published reliability score as the latter was not available.

Finally, for the postural stability task, in addition to deviation scores, participants' weight was also obtained in pounds as a control measure to ensure that the degree of sway after the balance challenge was not affected by how much they weighed. Out of 85 participants, one refused to be weighed. A correlation analysis revealed a significant relationship between weight and average deviation score for the postural stability task with arms held at sides and arms held straight out in front (i.e., *r* (84) = -.30, *p* = .006; *r* (84) = -.22, *p* = .03). Consequently, the effect of weight on the deviation score for postural stability will be further explored in the results.

Muscle tone. The second cerebellar task measured *muscle* tone. The procedure used to measure muscle tone was entirely different than the one used in previous studies (Fawcett & Nicolson, 1999). In their study, Fawcett and Nicolson asked participants to sit down with their elbows resting on the chair arm and their hands dangling loosely. The experimenter then held each hand at the wrist and shook it lightly from side to side. Degree of movement was measured on a scale from 1 to 3, with 1 being “little movement” and 3 being “large movement.”

In this study, in order to measure muscle tone, a procedure developed by Wartenberg (1951), referred to as the Wartenberg Pendulum Test was used. The Wartenberg Pendulum test is a diagnostic test of quadriceps muscle tone (Ammann, et al., 2005). The quadriceps (Latin for four-headed) is a large muscle group that includes the four prevailing muscles on the front of the thigh. It is the great extensor muscle of the knee, forming a large fleshy mass which covers the front and sides of the femur (i.e., thigh bone) (Ammann et al., 2005; Biel, 2005).

The Wartenberg Pendulum test measures muscle tone by using gravity to provoke muscle stretch reflexes during passive swinging of the lower leg (Ammann et al., 2005). Originally, Wartenberg developed the pendulum tests as a simple and reliable test of lower-limb muscle tone in patients with Parkinson’s disease (Wartenberg, 1951). In the pendulum test, patients are placed on a bench or edge of a table with their leg hanging freely from the knee. The leg is lifted by the examiner to horizontal position and then released to swing freely like a pendulum under the force of gravity. The leg is expected to swing smoothly with regular and gradually decreasing movements (Brown, Lawson, Leslie, & Part, 1988; Nordmark & Anderson, 2002).

To measure the angles at the knee, an electrogoniometer is used. Various components have been measured to evaluate the test such as maximum velocity of the first swing, swing time, number of swings and amplitude of the first backward swing (Brown, Lawson, Leslie, & Part, 1988; Fowler, Nwigwe, & Wong Ho, 2000; Nordmark & Anderson, 2002). One of the most commonly measured parameters has also been the ratio between the amplitude of the first swing divided by the final angle of the knee (Nordmark & Anderson, 2002). The pendulum test has been tested in different studies with adults, such as healthy elderly or patients with multiple sclerosis (e.g., Brown, Lawson, Leslie, & Part, 1988; Leslie, Muir, Part, & Roberts, 1992) and recently with young children with spastic diplegia ages ranging from 2.5 to 8.8 (e.g., Nordmark & Anderson, 2002) and found to be an objective measure of muscle tone in comparison to methods using a grading scale.

The pendulum test was thus used to measure muscle tone in participants of this study. An electronic goniometer and an angle sensor, as shown in Figure 12, were used to measure and record the angle at the knee joint.

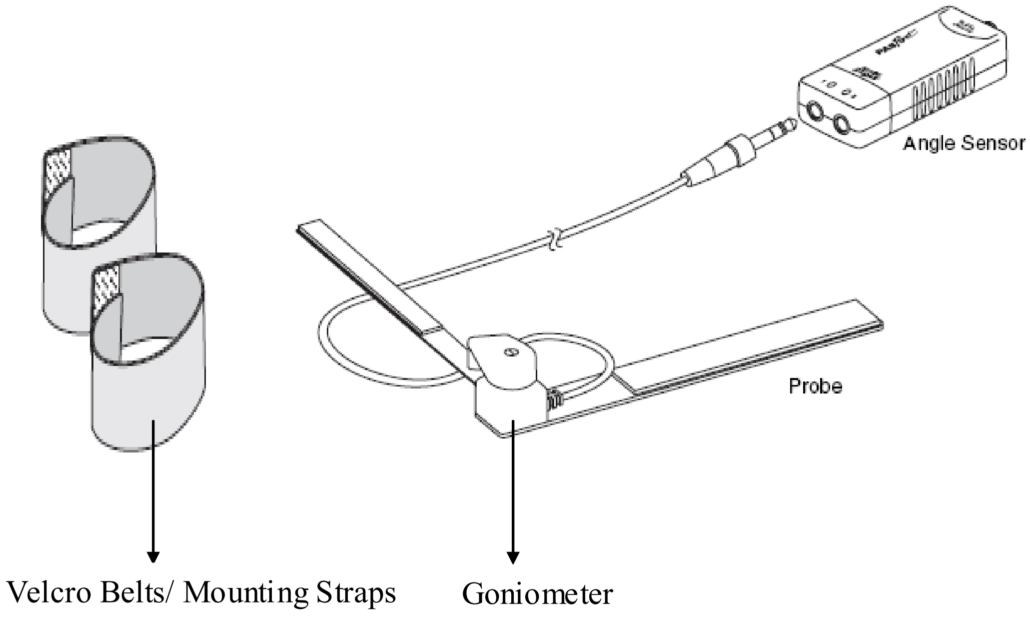


Figure 12. Goniometer, angle sensor, and mounting straps used for the muscle tone task.

As seen in Figure 12, the Goniometer probe consists of two arms and a potentiometer. To place the Goniometer on the lower limbs, the mounting straps, shown in the figure, are used. A larger strap is placed around the thigh just above the knee and a smaller strap is placed around the upper part of the calf, as shown in Figure 13.

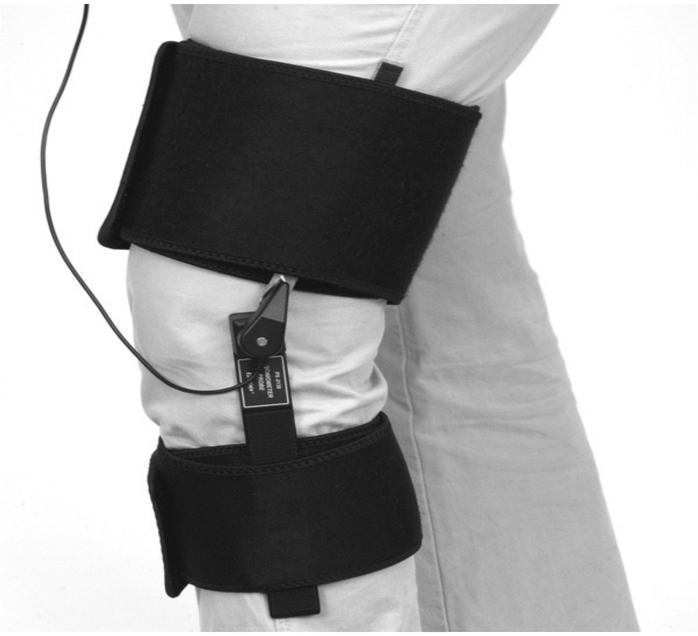


Figure 13. Placement of goniometer and mounting straps on lower limp prior to angle recordings.

The Goniometer probe's hinge is then aligned with the knee. One arm of the probe is attached to the thigh parallel to the thigh bone (i.e., femur) and the other probe arm is attached to the shinbone (i.e., tibia). The probe measures zero degrees when it is fully open. A clockwise rotation of the narrow arm relative to the wide arm is measured as an increasing angle, as shown in Figure 14.

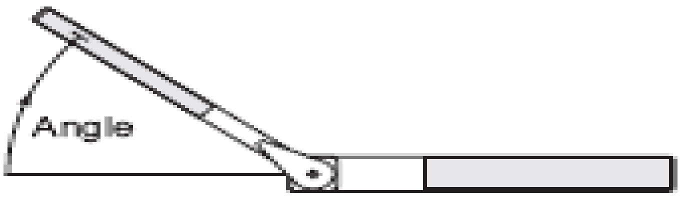


Figure 14. Clockwise rotation of probe narrow arm to measure positive displacement.

In order to measure flexion of knee joint as a positive displacement, the wide arm of the probe should be attached below the knee on the right leg and to the thigh on the left leg, as was done for this study. As the angle between the arms changes, the resistance of the potentiometer changes as well. The resistance of the potentiometer is measured by an Angle sensor, as was shown in Figure 12, which is connected to the Goniometer probe, and converts the resistance to an angle measurement. The Angle Sensor is connected to the computer through a USB Link, analogous to the one used in the postural stability task that was shown in Figure 7. Similar to the postural stability task, data is sent digitally to PASCO interface data collection and analysis software installed in the computer at up to 100 samples per second.

To complete the muscle tone task, participants wore an electronic goniometer at their right knee. Prior to securing the goniometer on participants, the experimenter showed the equipment to them and demonstrated on herself how the leg was lifted and then dropped to swing and how the data was recorded. It was also explained to participants that they would not feel anything during data recording as the process was completely painless and silent.

Upon this demonstration, the goniometer was secured to the knee with the Velcro mounting straps as previously shown in Figure 13. The goniometer was connected to the angle sensor and the angle sensor was connected to the laptop computer via the PASPORT USB Link. The PASPORT Data collection and analysis interface automatically launched when the angel sensor was connected to the computer.

Participants were asked to sit on the edge of a table high enough that their feet did not touch the floor. This position was used by Nordmark and Anderson (2002) with the young children (ages ranging from 2.3 to 8.8) who participated in their study. As they have suggested, this position, while not standardized, had its advantages because children were comfortable and able to tolerate the test well. Additionally, the position itself has not been found to have an important influence on the different measurements recorded in the Pendulum test (Brown, Lawson, Leslie, & Part, 1988). However, the Wartenberg Pendulum test relies on participants being relaxed and not resisting or assisting the pendular movement (Nordmark & Anderson, 2002). Such voluntary movements have been observed and reported in previous studies that have used the Wartenberg Pendulum test (Brown, Lawson, Leslie, & Part, 1988).

In order to decrease measurement error related to voluntary muscle activities, participants needed to completely relax their leg. This was accomplished by devoting time for practice prior to actual recordings until participants felt comfortable with the procedure. In a few cases, and especially with one participant who resisted the pendular movement, relaxation was achieved by closing the eyes during the procedure. The actual task began only when participants understood the process completely.

Once the actual task began, the experimenter lifted the participant’s right foot so that the leg was extended at 180 degrees. The foot was then dropped to swing like a pendulum. Limb oscillation during the pendulum test is shown in Figure 15, with the solid line representing leg at the starting position (i.e., extended leg) and at the final resting position (i.e., flexed leg). This procedure was repeated three times for each participant and each time the angle at the knee joint was recorded by the angle sensor.

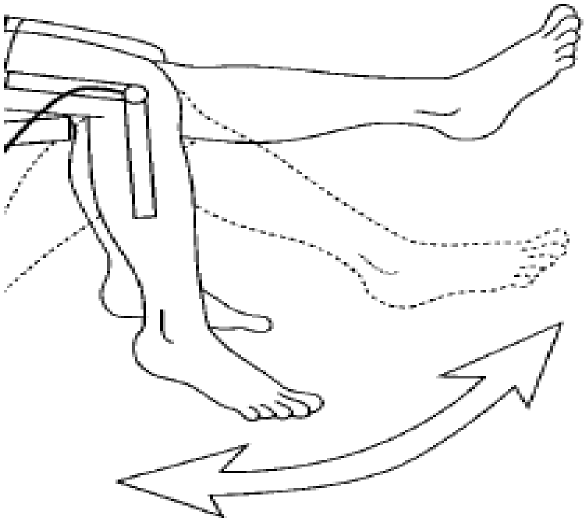
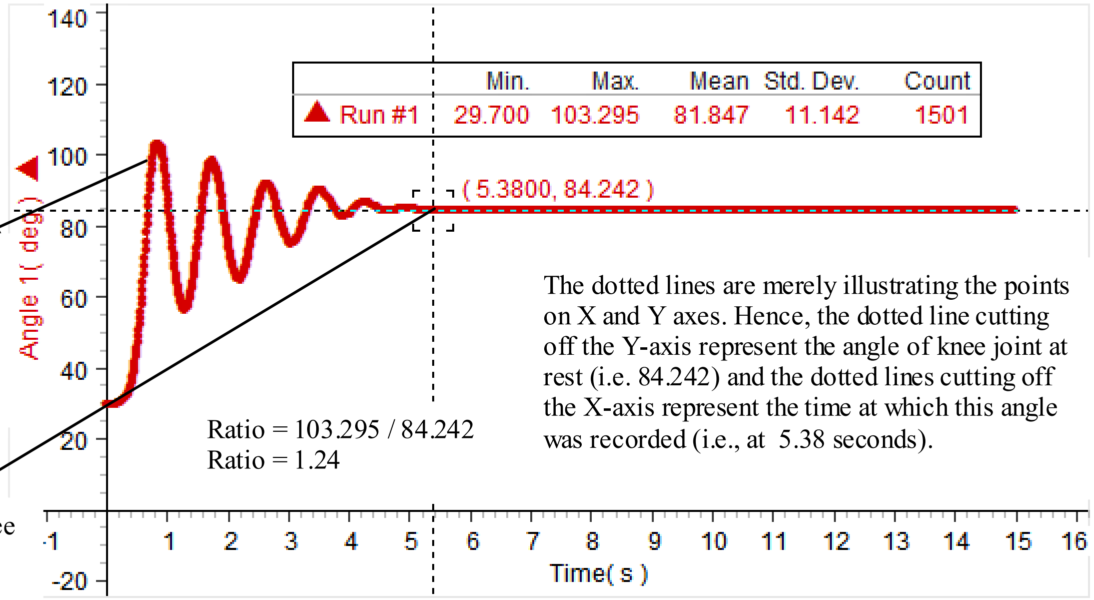


Figure 15. Leg (limb) oscillation during the pendulum test.

The recording time was set for 5 seconds for some of the initial participants. However, because swinging lasted longer than 5 seconds for a few participants, the recording time was increased to 15 seconds in order to ensure that data collection was not terminated prematurely, as this would have led to inaccurate measurements (Stillman, Phty, McMeeken, & Phty, 1995).

The recordings produced by the angle sensor for the muscle tone test are quiet different from the recordings that were produced from the accelerometer sensor for the postural stability task. Consequently, the figures produced from the angle sensor look very different from those that were produced from the accelerometer sensor and thus should not be confused with each other. In both cases (the accelerometer sensor and angle sensor), the X-axis represents the time during which data is recorded. As discussed earlier, accelerometer sensor data recorded were acceleration rates, represented on the Y-axis. In the case of the angle sensor, however, the data recorded were leg swings. Consequently, the Y-axis in figures produced by the angle sensor represents the degree of the recorded angles thus producing different figures. Figure 16 illustrates the graphic representations of angle recordings traces during the leg-swing of a participant in the sample.



The angle of knee joint at maximum bent (i.e., 103.295°)

The angle of knee joint at rest (i.e., 84.242°)

Ratio = 103.295 / 84.242
Ratio = 1.24

The dotted lines are merely illustrating the points on X and Y axes. Hence, the dotted line cutting off the Y-axis represent the angle of knee joint at rest (i.e. 84.242) and the dotted lines cutting off the X-axis represent the time at which this angle was recorded (i.e., at 5.38 seconds).

Figure 16. Angle recording traces using electronic goniometer and angle sensor for a participant in sample.

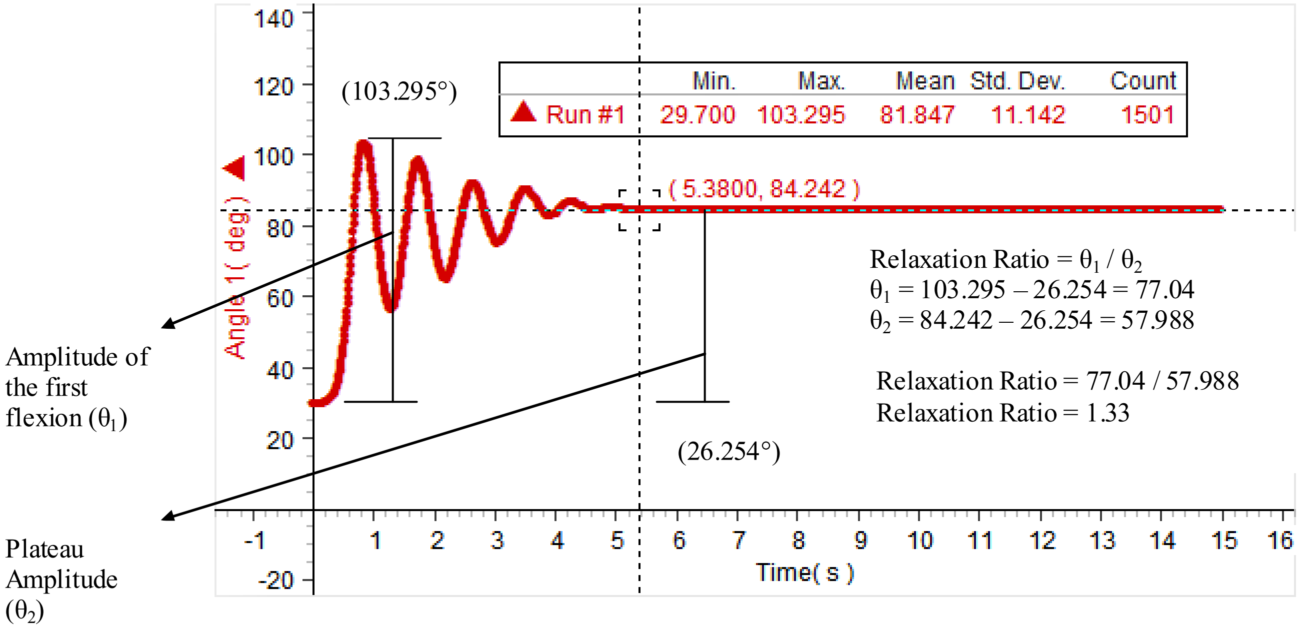
As indicated earlier, the X-axis in Figure 16 represents the time during which angles and leg swings were recorded and the Y-axis represents the degree of the recorded angles. Part of the data analyzed through the PASCO Data Analysis interface was the minimum and maximum angles and the mean of all angles recorded, which are illustrated in both figures.

As discussed earlier, the most commonly measured parameter, which has been considered a reliable measure of muscle tone, is the ratio between the amplitude of the first swing divided by the final angle of the knee (e.g., Brown, Lawson, Leslie, MacArthur et al., 1988; Brown, Lawson, Leslie, & Part, 1988; Nordmark & Anderson, 2002). To calculate this ratio, the measurements taken from each recording were the angle of the knee joint at its maximum bend and at its resting position. For example, for the participant in Figure 16, the angle of the knee joint at its maximum bend was 103.295° and the angle of the knee joint at its resting position was 84.242°, as shown in the figure. Thus the ratio for the participant in Figure 16 would be 1.24 (i.e., 103.295/84.242). This ratio was calculated for each participant in the sample in all three trials and the final score was the average of the scores for the three trials.

The ratio described has also been measured differently in some studies (e.g., Brown, Lawson, Leslie, & Part, 1988; Valle, et al., 2006), namely as the ratio between the amplitude of the first or initial flexion upon dropping the leg (i.e., θ_1) and the plateau amplitude (i.e., θ_2) referred to as the relaxation ratio.

$$\text{Relaxation Ratio} = \frac{\theta_1}{\theta_2}$$

To illustrate, Figure 17 represents how the ratio is calculated using this ratio for the same participant shown in Figure 16.



	Min.	Max.	Mean	Std. Dev.	Count
▲ Run #1	29.700	103.295	81.847	11.142	1501

Relaxation Ratio = θ_1 / θ_2
$\theta_1 = 103.295 - 26.254 = 77.04$
$\theta_2 = 84.242 - 26.254 = 57.988$

Relaxation Ratio = 77.04 / 57.988
Relaxation Ratio = 1.33

Figure 17. Angle recording traces using electronic goniometer and angle sensor for participant shown in Figure 16.

As shown in Figure 17, the amplitude of the first flexion or θ_1 is calculated by subtracting the angle recorded at the first flexion (i.e., angle at maximum bend which was 103.295°) from the initial extension angle (i.e., the angle recorded when leg was extended just prior to the drop) which is 26.254°, illustrated in Figure 17. Thus θ_1 would be 77.04 (i.e., 103.295 – 26.254). The plateau amplitude or θ_2 is calculated by subtracting the angle recorded at a resting position (i.e., final angle which was 84.242°) from the initial extension angle, which as indicated is 26.254°. As a result θ_2 would be 57.988 (i.e., 84.242 – 26.254). Hence, the relaxation ratio for the same participant who was previously illustrated in Figure 16 would be 1.33 (i.e., 77.04 / 57.988). A relaxation ratio obtained in this manner was also calculated for each participant in the sample for all three trials. A correlation analysis was conducted to explore the relationship between the two ratios obtained in the two different manners explained. Because of the fact that a significant relationship was revealed between the two ratios, r (85) = .81, $p < .001$), only the first ratio (i.e., ratio between angle at maximum bend and at resting position) was used in the analyses of results. Children with more optimal cerebellar functioning were expected to have lower scores (lower range of movement, greater muscle tone) than children with poorer cerebellar functioning (Fawcett & Nicolson, 1999; Fawcett et al., 1996).

It should be noted that despite the time devoted to practice, a few participants still continued to assist the pendular movement during the actual recording trials after two or three unassisted swings. However, this co-operation was recognizable in the recordings and thus could be controlled. That is, the measurement for the denominator of the ratio was obtained from the place where the actual unassisted swing ended and before the cooperation started.

Furthermore, to control for assisted swings during the Pendulum Test, the amplitude of the first flexion or backward swing (i.e., θ_1 or the difference between the angle recorded at the first flexion from the initial extension angle as calculated earlier in Figure 17) has been recommended (Brown, Lawson, Leslie, & Part, 1988) as an additional measure because this measure is not affected by participants' voluntary swings. Since this measure was available for all participants in the sample, a correlation analysis was conducted to explore the relationship of this measure with the ratio calculated for all participants as well as with reading measures used in the study. The analysis revealed a significant relationship between this amplitude of the first flexion and the ratio calculated for participants, r (85) = .51, $p < .001$). Additional correlations conducted to explore the relationship between the amplitude of the first flexion and reading measures used in the study did not reveal any significant relationship. Consequently, this measure was not added in the analyses in order to maintain the ratio between the number of variables already included in the study to the number of participants.

The reliability of the Pendulum test has been measured in some studies for spasticity. For example, the published test-retest reliability for spasticity reported by Bohannon (1987) was .96 for 30 participants who were tested four times consecutively. Stillman et al. (1995) have also reported a test-retest reliability of .84 for 14 participants who were retested in their study 26.8 days apart. For the present sample, the average correlation between the muscle tone ratios obtained for all three trials was .76. The internal consistency coefficient alpha computed was .85 for muscle tone ratios obtained for the last two trials and .90 for all three trials. These reliabilities were similarly high as the published reliabilities indicated above that were obtained for spasticity. Unfortunately, the reliabilities obtained for the sample in this study could not be directly compared to a published reliability of the Pendulum test for measuring hypotonia, as the latter was not available.

Toe tap speed. The last task measured *toe tap speed*. Following the procedure used in Fawcett and Nicolson's study (1999), participants were asked to tap their foot as fast as they could on a tin lid that was secured onto the floor. A microphone was placed under the tin lid and connected to the laptop computer to record the tapping sound.

Prior to asking participants to tap their foot, tapping was demonstrated to the child by the experimenter. Children were given some time to practice tapping on the lid with both their right and left foot until they felt comfortable with the procedure. Practice time for all children took under a minute but was not recorded. Once children felt comfortable with the task the actual task began. Sounds were then recorded on the computer and the speed of tapping was assessed accurately using a digital audio editor, namely the Audacity sound wave analysis software. The score was the speed at which the 10 taps were completed, which was recorded as the time it took to complete the task. Cerebellar dysfunction was expected to lead to slower tapping times (Fawcett & Nicolson, 1999; Fawcett et al., 1996), although as Fawcett and Nicolson (1999) have suggested, many other factors could also affect speed of tapping. No reliability estimate could be calculated for the toe tapping task as the test was administered once in a single trial for each foot. However, for the sample in this study correlation between "toe tapping time" for the left and "toe tapping time" for the right foot was .80. The Spearman-Brown split-half reliability as well as the internal consistency coefficient alpha obtained for the tapping time for the left and right foot was .89 in both cases.

CHAPTER 4

Results

Part 1: Preliminary Data Analyses

Prior to main multivariate analyses, all variables were examined, using the standard approach recommended by Tabachnik and Fidell (2007), for accuracy of data entry, univariate and multivariate outliers, skewness, kurtosis, normality, linearity, and homoscedasticity.

Accuracy

Following Tabachnik and Fidell's (2007) recommendations, frequencies of continuous variables, including their minimum, maximum, means and standard deviations were inspected for accuracy. Part of the original data was also proofread manually against the computerized data file.

Univariate outliers

According to Tabachnik and Fidell (2007), cases with standardized scores larger than 3.29 ($p < .001$, two tailed test) are considered to be potential univariate outliers. All measures were thus additionally saved as z-standardized scores. Measures included ADHD Index Raw scores, IQ Standard Score, Word Identification, Word Attack, TOWRE Word and Non-Word Reading, Elision, Rapid Digit, Letter, and Object Naming raw scores, as well as two postural stability measures (i.e., with hands at sides and hands stretched out in front), average muscle tone ratio, two toe tapping measures (i.e., with right and left foot), bead threading, and peg moving. For all the enumerated measures, frequencies of their z-standardized scores were examined. To inspect measures graphically, box plots and steam-and-leaf graphs, normal probability plots and detrended normal probability plots were used to reveal extreme cases as has been suggested by

Newton and Rudestam (1999) and Tabachnik and Fidell (2007). Using these methods, eight different cases with *z*-scores above 3.29 were found to be univariate outliers on the three Rapid Naming tasks (i.e., $z = 3.50$ on Rapid Digit Naming, $z = 4.64$ on Rapid Letter Naming, and $z = 4.56$ on Rapid Object Naming), as well as on the two postural stability (i.e., *z*-scores 3.99 and 4.02) and the two toe tapping measures (i.e., *z*-scores 4.28, 4.15, and 3.86). One additional case was found to be a marginal outlier on Word Identification (*z*-score = 3.10). As suggested by Tabachnik and Fidell (2007), it was investigated whether these were random cases or if there was an underlying reason for the cases to be outliers. This inspection revealed that the outliers seemed to be random occurrences (i.e., the same cases were not outliers on all measures). The procedure for correcting the outliers is discussed later in the "Normality" subsection.

Multivariate outliers

Following standard procedures recommended for examining data for multivariate outliers and influential cases among the set of variables, statistics assessing Mahalanobis and Cook's distance were computed through the SPSS Regression (Newton & Rudestam, 1999: Stevens, 1996: Tabachnik & Fidell, 2007). Usually the 10 most extreme cases are reported in the regression. Cases are considered multivariate outliers at a Mahalanobis Distance with $p < .001$ for the χ^2 value from the centroid (or point created at the intersection of the mean of all variables) (Tabachnik & Fidell, 2007). If a case is considered an outlier, then the Cook's Distance for that case is examined. Cook's Distance, which is considered to be useful in identifying influential points, measures the combined influence of an outlier on both the dependent and independent variables in a regression (Stevens, 1996; Tabachnik & Fidell, 2007). If the Cook's Distance for the case is also greater than one, then the case is considered influential and deleted from the data.

For the present data, no multivariate outliers and influential cases were identified and no data points were deleted from the sample.

Normality

All continuous variables listed earlier were screened for normality both statistically (using skewness and kurtosis) and graphically as has been suggested by Tabachnik and Fidell (2007). An alpha level of .001, which is suggested as a conventional but a conservative level for samples of small or moderate size (Tabachnik & Fidell, 2007), was used to evaluate the significance of skewness and kurtosis. The departure from normality was also inspected graphically through normal probability plots as well as detrended normal probability plots in which scores are ranked and sorted (Newton & Rudestam, 1999; Tabachnik & Fidell, 2007). An expected normal value (i.e., the *z*-score a case with that rank holds in a normal distribution) is then computed and compared with the actual values for each case. If the distribution is normal, then most points for cases fall along the diagonal running from lower left to the upper right. However, in cases of deviation from normality, points shift away from the diagonal (Newton & Rudestam, 1999; Tabachnik & Fidell, 2007).

Normality within reading measures. Within reading measures, Word Identification and all three rapid naming variables were considered skewed, which was possibly due to outliers that were found on these measures. Although in such cases transformation of variables is generally recommended, some researchers such as Tabachnick & Fidell (2007) have pointed that transformations can increase difficulty in interpretation depending on the scale in which the variable is measured. For example, transformation may often hinder interpretation in cases where a scale, such as intelligence or reading measures in this study, is meaningful and widely used. In such cases, transformation is not recommended in ungrouped data. Hence, these measures were not transformed. However, the outliers found on these variables were corrected using two methods. The first method to correct the outliers was regression in which variables with large correlations, namely correlations of .50 and larger (Newton & Rudestam, 1999), were used to predict the variable with outliers. For example, variables Word Attack, TOWRE Word and Non-word Reading, and Elision were used to predict Word Identification. In each regression, the unstandardized predicted values were then saved and the predicted values were used to replace the outlier cases. The other method used to correct the outliers was using rows and columns to calculate the expected values for variables with outliers using the following formula:

$$\text{Expected value for outlier case} = \frac{\text{Row Total x Column Total}}{\text{Grand Total}}$$

For example, to calculate the expected value for the outlier case in Word Identification, rows and columns of all reading variables entered as raw scores (i.e., Word Attack, TOWRE Word and Non-Word Reading, Elision, and Rapid Digit, Letter, and Object Naming) were used. The outlier case was then replaced with the expected value that was calculated. To ensure the method of correction did not affect the results, the analyses were conducted with uncorrected variables as well as with the corrected variables using both methods of correction.

Normality within motor/cerebellar measures. Within motor and cerebellar measures, both toe tapping and posture control tasks were also significantly positively skewed. As discussed earlier, outliers were found on both of these tasks. As opposed to reading and intelligence measures, motor and cerebellar measures are not standardized and are not widely used measures. The distribution of these variables has been reported to deviate from normality in other studies as well (Savage et al., 2005b; White et al., 2006). Consequently, transformations were applied to correct these variables. The recommended transformations by Newton and Rudestam (1999) and Tabachnik and Fidell (2007) to correct normality in case of positive skewness include: (a) a square root transformation for moderate deviation from normality, (b) a log transformation for substantial deviation from normality, (c) and the inverse log transformation for severe deviation from normality. As Tabachnik and Fidell (2007) have recommended, the transformations of the variables started with a square root transformation. Because skewness was not corrected in any of the variables, a log transformation (using the formula LG10) was applied to each variable. Inspections revealed that the skewness of the two postural stability measures was corrected to a great extent using the log transformation. However, toe tapping measures were corrected using the inverse log transformation (using formula 1/toe tapping). Again, to ensure that transforming variables or the method of correcting outliers did not affect the results, the main analyses were repeated using untransformed variables as well as the corrected transformed variables using both methods of correction.

Linearity

Following standard procedures recommended by Tabachnik and Fidell (2007), linearity was assessed by inspecting bivariate scatter plots. As the authors have suggested, since numerous variables were involved in this study, all possible combinations of bivariate relationships were inspected between variables that were found to be skewed (i.e., Word Identification, all three Rapid Naming measures, the two postural stability measures, and the two toe tapping measures as well as the transformed versions of these variables). Results did not reveal violations of linearity (e.g., no curvilinear relationships were found).

Homoscedasticity

Following Tabachnik and Fidell's (2007) recommendations, to examine if this assumption was met, a scatter plot of residuals was observed (i.e., a scatter plot of standardized residuals against standardized predicted values) using the SPSS Regression function. When the assumption of homoscedasticity is met, no trend is observed in the scatter plot. This was the case for the present data, as no trend was observed in the scatter plot of residuals.

Missing Values

For all 85 participants, scores on all reading, motor and cerebellar measures were available. However, the parent's behavior rating on the adapted Conner's Parent Rating Scale was missing for 2 participants. The two cases were kept in the data since all other important demographic information and continuous measures were available for them. However, in all subsequent main analyses that included the ADHD index, analyses were repeated both with and without the 2 participants to ensure there were no differences in the results. Additionally, as explained in the method section, while the parents of most participants (n = 57) completed the adapted Conner's Parent Rating Scale, for some of the participants (n = 28), teachers instead of parents completed the form. To investigate whether there were any differences in reading, motor and cerebellar related measures based on whether parents or teachers completed the adapted Conners' Parent Rating Scale, the data was divided into two groups reflecting the parent- vs. teacher-completed Conners questionnaire. Several *t*-tests were conducted to evaluate whether participants in the two groups differed on any of the reading, rapid naming, motor and cerebellar measures. Results of *t*-tests on reading and rapid naming measures did not reveal any significant differences.

Prior to investigating differences in motor and cerebellar measures, three summary scores were created for (a) the two postural stability, (b) the two toe tapping, and (c) the two motor (i.e., bead treading, peg moving) tasks. For example, to create a postural stability summary score, *z*-standardized scores on each postural stability task were combined across participants. To create summary scores for postural stability and toe tapping measures, both transformed and untransformed variables were used. The results reported here are those performed on the summary scores that were created using the transformed measures, since those performed on the untransformed measures yielded the same pattern of findings. Among the motor and cerebellar measures, significant differences were found only for muscle tone, t (83) = 2.673, $p < .01$, and toe tapping summary score, t (83) = 2.90, $p < .01$. That is, in the group for which parents completed the Conners rating scale as compared to the group for which teachers completed the form, participants on average seemed to have higher muscle tone ratios (i.e., weaker muscle tone). They also seemed to take less time to tap both their left and right foot 10 times on the floor.

Because of the fact that the data included a wide age range, the question of whether mean differences between the groups in muscle tone and toe tapping would still exist after adjustments were made for age was explored. This was addressed by conducting two separate one-way between subjects analyses of covariance with group as the between subject and muscle tone and toe tapping summary score as the dependent measures in each analysis. After adjustment for age, the groups no longer differed in toe tapping, F (1, 82) = 2.61, $p = .11$, but differences in muscle tone still persisted, F (1, 82) =

7.49, $p < .01$. Hence, to ensure that the main findings were not affected, muscle tone was treated as an extraneous variable and all main analyses were repeated on the two groups (i.e., the group with weaker and the group with stronger muscle tone).

Part 2: Main Analyses

As was indicated in the previous chapter, prior to addressing the main questions of this study a preliminary step was taken to determine whether motor and cerebellar measures could be reduced into factors for subsequent analyses. An exploratory Principal Component analysis approach, using varimax rotation, was conducted on these measures as suggested by Stevens (1996) and Tabachnik and Fidell (2007). Items entered in the Principal Component analysis included both postural stability measures, both toe tapping measures, muscle tone measure, as well as bead threading and peg moving measures. Following Tabachnik and Fidell's (2007) recommendation, the transformed variables for the postural stability and toe tapping measures were entered in the Principal Component analysis since the original variables were skewed.

Preliminary analyses provided support that the data was suitable for conducting the Principal Component analysis. These analyses included (a) the correlation matrix that included several correlations above .3; (b) the Keizer-Meyer-Olkin value of .64 which exceeded the minimum recommended value of .6; and (c) the statistically significant Bartlett's test of sphericity, $\chi^2 = 165.96$, $p < .001$ (Tabachnik & Fidell, 2007). The Principal Component analysis revealed two components, each with Eigen values exceeding one. The first component explained 37.35%, and the second 20.16% of the variance. Loadings of variables on components as well as their communalities are illustrated in Table 8.

Table 8

Component Loadings and Communalities (h^2) for Principal Component Analysis and Varimax Rotation on Motor and Cerebellar Measures

Measure	Component 1	Component 2	Communality h^2
Log10 Postural Stability (Hands at Sides)	.80	.02	.64
Log10 Postural Stability (Hands Stretched out in Front)	.77	-.03	.59
Average Muscle Tone Ratio	.34	.38	.25
Inverse Log Toe Tapping (Left Foot)(Right Foot)	-.24	.88	.84
Inverse Log Toe Tapping (Right Foot)(Left Foot)	-.10	.88	.79
Bead Threading	-.54	.29	.38
Peg Moving	-.58	.45	.54
% of Variance Explained by Component	37.35	20.16	

According to Tabachnik and Fidell (2007), loadings between .32 and .44 are considered “poor.” As indicated by the underlined loadings, the log10 postural stability measures loaded strongly on the first component with loadings higher than .63, which are considered “very good” (Tabachnik & Fidell, 2007). The inverse log toe tapping measures loaded strongly on the second component. Bead threading and peg moving, which were considered to be pure motor measures (Fawcett & Nicolson, 1995c) did not seem to be as well defined as the transformed postural stability and toe tapping measures since they loaded on both components one and two. Both bead threading and peg moving seemed to load well on the first component with the transformed postural stability measures (i.e., with loadings of -.54 and -.58 which are considered “fair” and “good,” respectively; Tabachik & Fidell, 2007). Peg moving also loaded fairly with the transformed toe tapping measures on the second component.

Finally, muscle tone, which was considered along with postural stability as the most cerebellar-based (Fawcett & Nicolson, 1999), did not seem to load strongly on any of the two components. Further inspection also indicated that muscle tone exhibited poor communality (.27), as indicated in Table 8. Additionally, the results of the correlation matrix in the Principal Component analysis indicated that muscle tone did not correlate significantly with any of the other variables. The correlations were as follows: (a) log10 postural stability (hands at sides), *r* (85) = .10, *p* = .20; (b) log10 postural stability (hands stretched out in front), *r* (85) = .13, *p* = .11; (c) inverse log toe tapping (right foot), *r* (85) = .08, *p* = .24; (d) inverse log toe tapping (left foot), *r* (85) = .09, *p* = .20; (e) bead threading, *r* (85) = -.00, *p* = .49; and (f) peg moving, *r* (85) = .05, *p* = .32. Altogether, poor loading on components and communality as well as poor correlations with other variables were an indication that muscle tone was an outlier and could be removed from this Principal Component analysis, as recommended by Tabachnik and Fidell (2007).

A second Principal Component analysis was conducted on the same measures excluding muscle tone. The Keizer-Meyer-Olkin value remained at .64 as in the previous analysis. The Bartlett's test of sphericity was also strong and statistically significant, ($\chi^2 = 163.40$, $p < .001$), thus supporting factorability of the correlation matrix (Tabachnik & Fidell, 2001). Again, two components were extracted from this analysis, each with Eigen values exceeding 1. The first component explained 43.58%, and the second 22.06% of the variance. Loadings of variables on components as well as their communalities are illustrated in Table 9.

Table 9

Component Loadings and Communalities (h^2) for Principal Component Analysis and Varimax Rotation on Motor and Cerebellar Measures (Excluding Muscle Tone)

Measure	Component 1	Component 2	Communality h^2
Log10 Postural Stability (Hands at Sides)	.82	.03	.68
Log10 Postural Stability (Hands Stretched out in Front)	.76	-.05	.57
Inverse Log Toe Tapping (Left Foot)(Right Foot)	-.20	.92	.88
Inverse Log Toe Tapping (Right Foot)(Left Foot)	-.05	.93	.87
Bead threading	-.57	.26	.40
Peg Moving	-.61	.41	.55
% of Variance Explained by Component	43.58	22.06	

Similar to the previous analysis, the underlined loadings, which are higher than 0.4, indicated that the log10 postural stability measures loaded strongly on the first component and the inverse log toe tapping measures loaded strongly on the second component. Loadings for bead threading and peg moving measures in particular were still not well defined as was found in the previous Principal Component analysis. Both measures loaded well with the transformed postural stability components on the first component, while peg moving also loaded to some extent with the transformed toe tapping measures on the second component.

Exploring the correlation matrix indicated that bead threading was significantly correlated with log10 postural control (hands stretched out in front), r (85) = -.22, p <.05, log10 postural control (hands at sides), r (85) = -.27, p < .01, and inverse log toe tapping (right foot), r (85) = .32, p < .01. Peg moving was significantly correlated with all other variables as follows: (a) log10 postural stability (hands at sides), r (85) = -.35, p < .001; (b) log10 postural stability (hands stretched out in front), r (85) = -.30, p < .01; (c) inverse log toe tapping (right foot), r (85) = .40, p < .001; and (d) inverse log toe tapping (left foot), r (85) = .32, p < .01. However, the highest correlation seemed to be between the two measures of peg moving and bead threading, r (85) = .46, p <.001.

Although components extracted with Eigen-values greater than one is the general rule for retaining the most important components (Stevens,1996; Tabachnik & Fidell, 2007), exploring variances and graphical representations of components have also been recommended in this decision making process (Stevens, 1996). The graphical method proposed is the scree test in which the magnitude of Eigen-values, which are represented on a vertical axis, are plotted against their ordinal numbers (i.e., whether it was the first, second, or third Eigen-values and so on) (Stevens, 1996; Tabachnik & Fidell, 2007). As

Stevens (1996) explains, usually the magnitude of successive Eigen-values drop off sharply and then level off. The most common method suggested is “to retain components before the first one on the line where they start to level off” (Stevens, 1996, p. 366).

Inspecting the scree test, shown in Figure 18, in the Principal Component analysis indicated that after the second Eigen-value, there was still a clear descent. The descent continues and the line appears to level off by the fourth Eigen-value. Thus, considering Stevens’s (1996) recommendation, the third Eigen-value, which had a value of .89 which is close to one, could also be retained. This Eigen-value represented a component that explained an additional 14.89% of the variance.

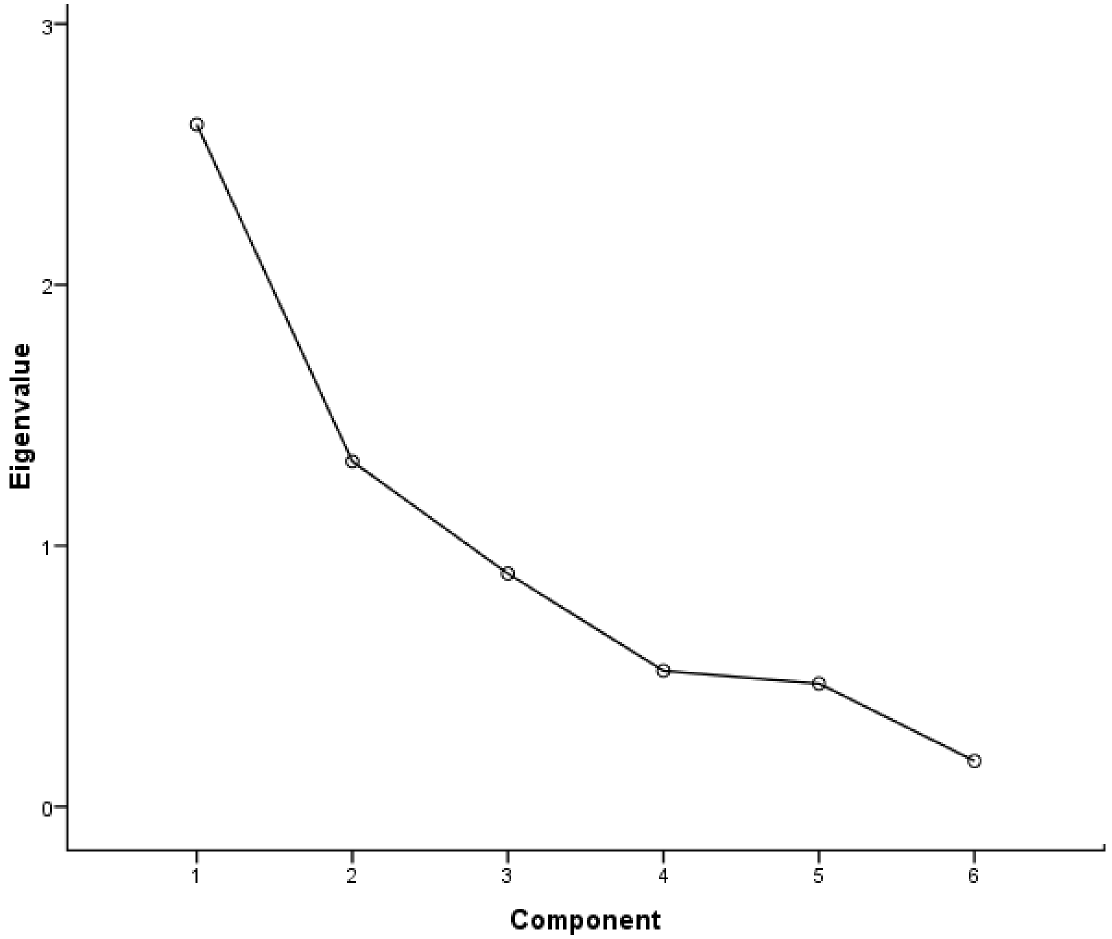


Figure 18. Scree plot for exploratory principal component analysis.

Hence, a third Principal Component analysis was conducted on the same measures as reported in Table 9. However, this time the extract function of the SPSS factor analysis was changed so that components with Eigen-values over .88 were also extracted. Three components were obtained from this analysis. As indicated earlier in the previous Principal Component analysis, the first two components explained 65.63% of the variance (i.e., 43.56% and 22.06% respectively). With the addition of the third component in this analysis, the variance explained was improved by 14.89%. Thus in contrast to the previous analysis, the total variance explained in this Principal Component analysis increased to 80.52%. Loadings of variables on components as well as their communalities are illustrated in Table 10.

Table 10

Component Loadings and Communalities (h^2) for Principal Component Analysis and Varimax Rotation on Motor and Cerebellar Measures with Eigen Values Above .90

Measure	Component 1	Component 2	Component 3	Communality h^2
Log10 Postural Stability (Hands at Sides)	-.01	.83	-.24	.75
Log10 Postural Stability (Hands Stretched out in Front)	-.14	.87	-.07	.77
Inverse Log Toe Tapping (Left Foot)(Right Foot)	.91	-.12	.25	.90
Inverse Log Toe Tapping (Right Foot)(Left Foot)	.95	-.05	.07	.92
Bead threading	.06	-.09	.90	.83
Peg Moving	.28	-.27	.72	.67
% of Variance Explained by Component	43.56	22.06	14.89	

As indicated by the underlined loadings, components extracted in this analysis were well defined with the inverse log toe tapping measures loading strongly on the first component, the log10 postural stability measures loading strongly on the second component, and finally the bead threading and peg moving measures loading strongly on the third component. Considering the improvement in loadings as well as communalities, these components were saved. The three components that were saved and used in subsequent analyses will be referred to as (a) toe tapping component which represented the two transformed toe tapping measures (i.e., inverse log toe tapping - right foot and inverse log toe tapping - left foot), (b) postural stability component which represented the two transformed postural stability measures (i.e., log10 postural stability - hands stretched out in front and log10 postural stability - hands at sides), and (c) motor component which represented the bead threading and peg moving measures. As discussed in the results for the first Principal Component analysis, muscle tone was removed from the analysis due to the fact that it did not load with any of the extracted components. Thus, this measure was used as a separate cerebellar variable in the subsequent analyses.

The Principal Component analyses explained in this section were also repeated using the untransformed variables, but the pattern of results remained the same. Furthermore, the last analysis for which components were saved was repeated using Principal Axis Factoring to investigate if similar results would be obtained. As Tabachnik and Fidell (2007) have suggested, one test of the stability of a factor analysis solution is that regardless of the extraction technique used, solutions remain the same. Similar to the Principal Component analysis, the Principal Axis Factoring revealed three factors (i.e., a top tapping, a postural stability and a motor factor). Although the loadings were somewhat lower than those in the Principal Component analysis, they were still strong and above the cutoff point of .45.

Additionally, the last Principal Component analysis for which three components were saved was also repeated including the muscle tone measure to investigate if the results would change. The analysis revealed the same three components. While the toe tapping and motor components remained well defined, the pattern of loadings for the postural stability component was somewhat different. Both postural stability measures loaded well on the first component (loadings of -.70 and -.57). However, the muscle tone measure also loaded strongly on the third component (loading of .83) along with a fair loading (.53) on postural stability (hands stretched out in front). To test if the solutions extracted from this Principal Component analysis would remain the same, the analysis was repeated using Principal Axis Factoring. Although all loadings remained similar in this analysis, the loading for muscle tone was decreased significantly and it no longer loaded strongly on the postural stability factor (loading of .21). Consequently, the previous decision to use this measure as a separate cerebellar variable remained unchanged. Finally, the stability of solutions extracted from the Principal Component analysis was also tested using oblique instead of varimax rotation and results remained unchanged.

Question 1:

Is there a relationship between word reading as measured by word identification and (a) phonological awareness, (b) reading fluency and rapid automatized naming, and (c) purported cerebellar processing tasks?

A correlation analysis was performed to examine the relationship between Word Identification (*z*-standardized scores) with the following measures used in this study: (a) non-word reading or phonological recoding as measured by Word Attack (*z*-standardized scores); (b) word reading fluency as measured by the TOWRE Word and Non-Word Reading (*z*-standardized scores); (c) phonological awareness as measured by the Elision (*z*-standardized scores); (d) rapid naming skills as measured by the Rapid Digit, Letter, and Object Naming tasks (z-standardized scores); (e) postural stability component; (f) muscle tone ratio; (g) toe tapping component; and (h) motor component. For Word Identification and the three rapid naming measures, the corrected variables (using rows and columns technique) were entered in the correlation analysis as standardized *z*-scores. Additionally, age, IQ standard scores, and ADHD index raw score (entered as *z*-standardized scores) were also included in the correlation analysis as control measures. The correlation analysis was repeated again with the same measures, but the effect of age was controlled by using the SPSS partial correlation program. The inter-correlations of the variables are presented in Tables 11. The bottom half of the table represents the results of the partial correlation controlling for chronological age.

Table 11

Correlations Between Age, IQ, Reading Measures, Cerebellar and Motor Components

Subscale	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
								n = 85							
1. Age	_	-.31**	.26*	.23*	.10	.21*	.24*	.05	-.13	-.15	-.22*	-.04	-.02	.30**	.50***
2. IQ		_	-.23*	.36**	.42***	.26*	.36**	.36**	.06	.02	-.15	-.18	.15	-.17	.04
3. ADHD Index		-.16	_	-.26*	-.20	-.25*	-.15	-.33**	.32**	.25*	.25*	.18	.13	.12	.10
4. Word Identification		.47***	-.34**	_	.86***	.85***	.86***	.62***	-.44***	-.36**	-.48***	-.11	-.02	.07	.29**
5. Word Attack		.48***	-.24*	.87***	_	.80***	.88***	.73***	-.38***	-.36**	-.43***	-.17	.05	.05	.21
6. TOWRE Word		.36**	-.33**	.85***	.81***	_	.87***	.58***	-.59***	-.48***	-.54***	-.11	-.08	.10	.26*
7. TOWRE Non-word		.47***	-.23*	.86***	.88***	.87***	_	.63***	-.47***	-.40***	-.49***	-.12	.05	.16	.30**
8. Elision		.41***	-.36**	.62***	.74***	.57***	.63***	_	-.39**	-.30**	-.27*	-.09	-.07	.03	.08
9. Rapid Digit Naming		.01	.37**	-.43***	-.38**	-.57***	-.46***	-.38**	_	.60***	.44***	.05	.03	-.04	-.18
10. Rapid Letter Naming		-.03	.30**	-.33**	-.35**	-.45***	-.37**	-.28**	.58***	_	.55***	.11	-.02	.02	-.17
11. Rapid Object Naming		-.25*	.33*	-.45***	-.42***	-.51***	-.46***	-.24*	.41***	.53***	_	.12	-.05	-.03	-.17

Table 11 (cont'd)

Subscale	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
							n = 85								
12. Postural Stability Comp.		-.20	.20	-.10	-.16	-.08	-.09	-.07	.03	.09	.10	_	.15	.00	.00
13. Muscle Tone		.17	.14	-.01	.06	-.06	.07	-.04	.02	-.04	-.07	.12	_	.10	.04
14. Toe Tapping Comp.		-.09	.04	.00	.02	.04	.10	.02	-.00	.07	.04	.02	.11	_	.00
15. Motor Factor Comp.		.23*	-.03	.21	.19	.20	.22*	.08	-.15	-.12	-.09	.02	.05	-.19	_

Note. *$p < .05$. **$p < .01$. ***$p < .001$. Bottom half of table illustrates correlations between variables after controlling for effect of age.

Correlations with control measures IQ, age, and ADHD Index. As results indicated in the top half of the table, IQ measure was significantly correlated with measures of word identification, word attack, TOWRE word and non-word reading fluency, and phonological awareness (Elision) in the present sample. Furthermore, significant positive correlations were found between age and measures that assessed word identification, and word reading fluency (including both TOWRE word and non-word reading). This indicated that sample performances on word identification and word reading fluency seemed to increase with increasing age. The size of these correlations were below .30, which is considered small (Newton & Rudestam, 1999). A negative correlation was also found between age and IQ, indicating that sample performance on the measure assessing IQ seemed to be lower with increasing age. This negative correlation seemed to be due to the nature of the sample, which included older participants with dyslexia who on average performed considerably poorer than some of the younger participants on the above measure. This is not surprising given that intelligence tests include some abilities (e.g., language-related abilities) that can be deficient in children with dyslexia (Reid, et al., 2001; Siegel & Ryan, 1989).

Finally, results also revealed significant negative correlations between the ADHD Index and measures of IQ, word identification, word attack, TOWRE word reading fluency, and phonological awareness. This indicated that higher ADHD index raw scores were associated with lower performance on IQ, reading, and phonological awareness measures, with the strongest association being between ADHD index score and phonological awareness, r (85) = .33. Significant positive correlations were also revealed between the ADHD index and all three rapid naming measures, indicating that higher ADHD index raw scores were associated with longer times to complete the rapid naming tasks. No significant correlations were found between the ADHD index and any of the cerebellar-related motor measures.

The pattern of correlations was somewhat different after chronological age was partialled out. One difference was that the correlation between ADHD and IQ was no longer significant after controlling for the effect of chronological age. Furthermore, two more significant negative correlations were revealed between ADHD and two reading related measures, namely word attack and TOWRE non-word reading fluency. This indicated that higher ADHD index raw scores were associated with lower performance on these two measures. While all other correlations remained significant after controlling for the effect of chronological age, those between ADHD and word identification, TOWRE word reading fluency and rapid naming measures increased modestly.

Correlations among reading measures. As expected, results revealed significant correlations between word identification and the measures word attack, word and non-word reading fluency, and phonological awareness. The size of these correlations was above .50, which is considered large (Newton & Rudestam, 1999). These correlations indicated that better performances on word identification seemed to be significantly related to better performance on word attack, better word and non-word reading fluency, and better phonological awareness skills and vice versa. Results also revealed significant negative correlations among word reading and all three rapid naming measures. This meant that better performance on word reading was related to better rapid naming skills since significantly less time was taken to name these stimuli. In contrast to the large size correlations obtained between word reading and measures of word attack, r (85) = .86, TOWRE word reading fluency r (85) = .85, TOWRE non-word reading fluency r (85) = .86, and phonological awareness, r (85) = .62, correlations obtained between word reading and measures of rapid naming were considered moderate (Newton & Rudestam, 1999), rapid digit naming *r* (85) = -.44, rapid letter naming *r* (85) = -.36, rapid object naming *r* (85) = -.48.

Further inspection of correlations also indicated that word attack, word and non-word reading fluency, and phonological awareness were all significantly related to each other in a positive direction. Word attack, word and non-word reading fluency, and phonological awareness were also significantly correlated with all three rapid naming measures, indicating that better skills on these measures were related to better rapid naming skills and vise versa. Finally, results also revealed that the three rapid naming measures were significantly correlated with each other, with large correlations among rapid letter naming and rapid digit and object naming and a medium size correlation between rapid digit and rapid object naming.

Correlations among reading and motor/cerebellar measures. Contrary to expectations and what the cerebellar deficit theory suggests, no significant relationships were revealed in the present sample among reading related measures including word attack, word and non-word reading fluency, phonological awareness, and rapid naming measures and measures related to postural stability component, muscle tone, and toe tapping components. Nonetheless, significant correlations were revealed among word identification and word reading fluency (including both TOWRE word and non-word reading) and the motor component. This indicated that better performance on these reading measures seemed to be related to better motor skills (i.e., bead threading and peg moving). Additionally, a large size correlation was also revealed between the motor component and age, *r* (85) = .50, as well as medium-sized correlation between the toe tapping component and age, *r* (85) = .30. These correlations indicated that an increase in age seemed to be related to better performances on toe tapping and motor skills (i.e., bead threading and peg moving).

The correlation analysis was repeated again while the effect of age was controlled using the SPSS partial correlation program. The results of this correlation analysis are presented in the bottom half of Table 11. After controlling the effect of chronological age, the small sized correlations among the measures of word identification and TOWRE word reading fluency and the measure of motor component were no longer significant. Additionally, the moderate size correlation between TOWRE non-word reading fluency and motor component (r (85) = .30) decreased to a small size correlation (r (85) = .22). A small correlation was also found between IQ and motor component, r (85) = .23. Generally, given the size of these two correlations and the number of correlations conducted in the analysis, it is possible that these correlations and other small sized correlations found were due to chance. The correlation analyses were repeated using uncorrected word identification and uncorrected rapid naming measures and the pattern of results remained the same.

Question 2:

Does a subgroup of children with dyslexia selected from the sample differ in their performance on any of the motor, cerebellar, reading, phonological, and rapid naming related measures when compared to two control subgroups that were selected from the same sample and matched to the dyslexia subgroup based on (a) their reading level, and (b) chronological age?

Prior to addressing the second research question, several preliminary steps were required which included: (a) selection of children with dyslexia from sample, (b) selection of a reading-age match control subgroup from the sample, (c) selection of a chronologically-age match subgroup from the sample, (d) inspection for success of the match on reading level and chronological age, (e) inspection for match on first language spoken by participants, and (f) inspection for univariate and multivariate outliers in each subgroup. These steps are described below.

(a) Selection of Children with Dyslexia from Sample

As a first criterion, in order for participants from the sample to be included in the dyslexia subgroup, they had to be in grade 3, 4, or 5. As a second criterion for inclusion of participants in the dyslexia subgroup, performance on the Word Identification subtest had to be below average. This included children in grades 3, 4, or 5 who obtained standard scores below 90 (i.e., below 25th percentile), which is a performance of more than .67 standard deviation points below the mean for the Word Identification subtest. This cut-off point has been used in other studies as well (Fletcher et al., 1994; Juel, 1988; Shaywitz et al., 2004; Siegel & Ryan, 1988).

Higher or lower cut-off points have also been used in intervention studies that screen for young at-risk readers. For example, the cut-off point used in Vellutino, Scanlon, Zhang, and Schatschneider (2008) to identify at-risk readers in kindergarten was performance at or below the 30th percentile on a letter-word identification test (i.e., standard scores at or below 92, which falls more than .50 standard deviation below the mean). This cut-off point has also been used in other intervention studies (e.g., Simmons, et al., 2008; Vellutino, Scanlon, & Lyon, 2000; Vellutino et al., 1996; Vellutino, Scanlon, & Tanzman, 1998). Lower cut-off points, such as below the 21st percentile (i.e., more than .75 SD below the mean) (O'Connor, Fulmer, Harty, & Bell, 2005), or below the 15th percentile (i.e., more than 1 SD below the mean) (Savage et al., 2005a; Scanlon, Vellutino, Small, Fanuele, & Sweeny, 2005; Torgesen, et al., 1999) have also been used to identify poor readers.

Choosing a cut-off point becomes more important especially in intervention studies that identify at-risk readers since it is related to the degree of accuracy with which struggling readers are identified. Given that this was not an intervention study, the standardized cut-off point for the Word Identification subtest, which was in between the lower and higher cut-off points used in other studies, was chosen to identify individuals with dyslexia in this sample. Using the 25th percentile cut-off, 17 participants were identified with dyslexia in the sample (13 in grade 3, one in grade 4, and three in grade 5). No other criteria were applied to the selection of participants in this subgroup.

(b) Selection of Reading-Age (RA) Match Control subgroup from Sample

The first criterion for the inclusion of participants in this subgroup was that participants had to be in grades 1 or 2. Next, following procedures used in other studies (e.g., Baddeley, Logie, & Ellis, 1988; Grainger, Bouttevin, Truc, Bastien, & Ziegler, 2003; Olson, Wise, Conners, Rack, & Fulker, 1989; Savage et al., 2005a; Snowling, 1981) participants were matched to the dyslexia subgroup based on their reading age. Similar to the procedure used by Savage et al. (2005a), children who were included in the RA-match subgroup were those whose reading age was close to their chronological age. Hence, within grades 1 and 2, the initial selection was to choose among those whose reading age was exactly equal to their chronological age, then move outward to select children whose reading age was as close as possible to their chronological age until 17 participants were included in the RA-match subgroup (three in grade 1, 14 in grade 2). No other criteria were applied to the selection of participants in this subgroup.

(c) Selection of Chronologically- Age (CA) Match Control subgroup from Sample

Participants in the CA-match subgroup had to be in grades 3, 4, and 5 so that their chronological age matched the children in the dyslexia subgroup. Within these higher grade levels, the selection criteria was similar to the one used for the RA-match subgroup. Following Savage et al.'s (2005a) procedure, first those whose reading age was exactly equal to their chronological age were selected, then selection moved outward to include children with the smallest difference between their reading and chronological age until 17 children were selected in the CA-match subgroup (16 in grade 3, one in grade 4). No other criteria were applied to the selection of participants in this subgroup.

(d) Inspection for Successful Matching of Subgroups

To confirm if the subgroups were matched appropriately, six one-way analyses of variance were used using the SPSS GLM univariate program. There was one between-subjects factor with three levels (Group: dyslexia vs. reading-age match comparison vs. chronologically-age match comparison subgroups) in each of these analyses. For two of the analyses, data on reading age and chronological age were submitted to confirm if a good match was achieved for reading level and chronological age. For another three analyses, data on IQ were submitted to investigate if the three subgroups matched on their level of intellectual functioning (including their verbal and non-verbal skill). Finally, the last one-way analysis was performed to determine whether the three subgroups differed in their ADHD index raw scores. Means, standard deviations and results of the analyses of variances are displayed in Table 12.

Table 12

Mean and Standard Deviations on Reading Age, IQ and IQ subscales, Chronological Age, and ADHD Index by Group

Source	Dyslexia		RA-Matched		CA-Matched		df	F	p	Source of Effect
	M	SD	M	SD	M	SD	Between Subjects			
Reading Age (in months)	83.53	8.97	86.41	5.81	103.59	8.83	2	31.22	.00	(1=2)<3
IQ (standard scores)	92.59	13.68	101.24	12.26	101.18	14.24	2	2.34	.11	1=2=3
Vocabulary (scaled scores)	8.59	2.69	9.47	2.07	9.76	3.01	2	.93	.40	1=2=3
Matrix Reasoning (scaled scores)	8.53	3.28	10.82	3.38	10.65	3.32	2	2.51	.09	1=2=3
Chronological Age (in months)	113.00	17.70	87.47	4.84	103.76	8.81	2	20.57	.00	(1=3)>2
ADHD Index (raw scores)	14.65	9.35	6.76	7.59	7.88	9.45	2	3.95	.03	1>(2=3)

Reading level. In the first analysis, data on reading age on the word identification subtest were entered as the dependent variable. As indicated in Table 8, the analysis of variance revealed a significant main effect of group, F (2, 48) = 31.22, $p < .001$. To check for homogeneity of variance across the subgroups, Levene's test of error variance was selected in the GLM univariate program, as recommended by Howitt and Cramer (2008). Results of Levene's test of error variance for word reading was not significant, F (2, 48) = .73, $p = .489$, indicating that the assumption of homogeneity of variance was met across the subgroups. Simple main effect comparisons confirmed that a good match was achieved for reading level as the significant effect was due to the fact that the reading age of participants in both dyslexia and RA-match subgroups was lower than that of the CA-match subgroup. This means that, as would be expected, participants in the dyslexia and RA-match subgroups read significantly fewer words than the CA-match subgroup.

Chronological age. In this analysis, data on chronological age were entered as the dependent variable. The assumption of homogeneity of variance was not met for this analysis, F (2, 48) = 7.22, $p = .002$. Results revealed a significant main effect of group, F (2, 48) = 20.57, $p < .001$. In the case of violation of the assumption of homogeneity of variance, it is recommended to use post hoc comparison tests such as Dunnett's C which allow for unequal variances (Howitt & Cramer, 2008). The Dunett's C comparisons indicated that a good match for chronological age was achieved, as the significant effect was because of the fact that participants in dyslexia and CA-match subgroups were significantly older than those in the RA-match subgroup.

Intelligence. To investigate whether the three subgroups matched on the level of intellectual functioning, data on IQ standard scores were entered as the dependent variable in the next analysis of variance. Whether the subgroups differed on the two IQ subscales was also explored, namely, verbal ability as measured by the Vocabulary and non-verbal ability as measured by the Matrix Reasoning tasks. Two additional analyses of variance addressed this with Vocabulary and Matrix Reasoning scaled scores as the respective dependent variables. For these three analyses of variance, type I error was controlled with the Bonferroni method to adjust the α level to .016 (α of .05 divided by 3).

As shown in Table 12, for the analysis of variance with the dependent variable IQ, the main effect of group was not significant, $F(2, 48) = 2.34$, $p = .11$, indicating that the 3 subgroups did not differ in their level of intelligence. Similarly, results for the next two analyses of variance with the Vocabulary and Matrix Reasoning as the dependent variables did not reveal any significant main effects, $F(2, 48) = .93$, $p = .40$, and $F(2, 48) = 2.51$, $p = .09$), indicating that the 3 subgroups also did not differ in their verbal and non-verbal skills. Levene’s test of equality of error variances was not significant for any of the above three analyses, (a) IQ, $F(2, 48) = .38$, $p = .686$; (b) Vocabulary, $F(2, 48) = 1.69$, $p = .196$; and (c) Matrix Reasoning, $F(2, 48) = .07$, $p = .930$. This confirmed that variances on the dependent measures were homogeneous across subgroups for all three analyses of variance.

ADHD index. The final one-way analysis of variance was conducted to determine whether the three subgroups differed in their ADHD index scores. Data on ADHD index raw scores were entered as the dependent variable. As discussed in the preliminary analysis section, the ADHD scores were missing for 2 participants in the sample. One of these 2 participants was still included in the matched data, in the CA-match subgroup. As a result, for this analysis of variance, the CA-match subgroup included only 16 instead of 17 participants. The assumption of homogeneity of variance was met, $F(2, 48) = .42$, $p = .662$. As shown in Table 12, the analysis of variance revealed a significant main effect of the group, F (2, 48) = 3.95, $p < .05$. Simple comparison tests indicated that on average participants in the dyslexia subgroup had significantly higher ADHD index raw scores than participants in both the RA and CA-match subgroups.

(e) Inspection for Match on Demographic Variables

In addition to reading level, chronological age, and IQ, whether the 3 subgroups matched on the proportion of the first language (L1) spoken by participants was investigated. To address this, a χ^2 (chi-square) test was performed using the SPSS Crosstabs analysis in the Descriptives function. Since information on the first language spoken was missing for a few cases (i.e., parents did not provide the information), the χ^2 reported here was conducted, excluding the cases for which this information was missing. The results indicated that the 3 subgroups matched on the first language spoken by participants χ^2 (2, N = 47) = 2.94, p = .569. In other words, the proportion of first languages spoken (as reported by parents) was similar within all 3 subgroups. For the majority of participants within all three subgroups, English was reported to be L1. For a smaller proportion of the participants within each subgroup, French and a language other than English or French were reported as L1. Nonetheless, as previously explained in the Method chapter on pp. 95-96, the conditions required for children in Quebec to be allowed to attend English schools guaranteed English exposure at home. This means that for the participants of this study (or at least for most of them), English appears to be first language or among one of the first languages at home. Therefore, the sample generally represented English-dominant bilinguals who were exposed concurrently to English and another language. Subsequently, language as such would not explain group differences observed in reading processes.

(f) Inspection for Univariate and Multivariate Normality in Each Subgroup

Once the 3 subgroups were selected, Tabachnik and Fidell's (2007) recommendation for grouped data was followed and each subgroup was separately inspected for univariate and multivariate normality and the presence of outliers. The procedure used to inspect for univariate and multivariate normality and the presence of outliers was similar to that explained in the preliminary analysis section. The following variables were inspected: Word Attack, Elision, TOWRE Word and Non-word Reading Efficiency, and Rapid Digit, Letter, and Object Naming raw scores, postural stability component, toe tapping component, motor component, and muscle tone ratio. In addition to the Word Attack variable, which included all 45 items, another variable for Word Attack was also created which represented performance of the three subgroups on the eleven polysyllable items only. This new variable was also inspected for normality and the presence of outliers.

The assumption of normality was not met for several variables within each subgroup. Elision variable was skewed in the dyslexia and RA-match subgroups. Rapid Letter Naming variable was skewed in the RA-match and CA-match subgroups. Additionally, Rapid Digit Naming variable was skewed in the dyslexia subgroup and Word Attack variable (representing all items) was skewed in the CA-match subgroup. Finally, the Word Attack variable representing polysyllable items for the dyslexia subgroup was skewed and markedly different than the other two subgroups. This was because all except 3 participants in the dyslexia subgroup had a score of zero on polysyllable items of the Word Attack. As shown in Table 13, altogether 26 outliers were identified within the 3 subgroups (i.e., 14 in the dyslexia subgroup, 10 in the RA-match, and 2 in the CA-match subgroups) on Word Attack, TOWRE Word and Non-word

Reading, Elision, as well as Rapid Digit, Letter, and Object Naming variables. No multivariate outliers were found in any of the subgroups.

Table 13

Number of Outliers on Variables within Each Subgroup

Variables	Number of Outliers		
	Dyslexia	RA-match	CA-match
Word Attack (all items)	2		1
Word Attack (polysyllables)	3		
TOWRE Word Reading	2		
TOWRE Non-word Reading	2		
Elision	3	6	
Rapid Digit Naming	1	2	
Rapid Letter Naming		2	
Rapid Object Naming	1		1
Total *n* for outliers	14	10	2

The inspection for normality was repeated for all skewed variables except Word Attack (polysyllables) with the outlier participants excluded. The assumption of normality was met this time. Therefore, the outliers on these variables were kept in the matched data. Several methods were used to correct the outlying cases. Two of these methods were explained earlier on page 161 of this chapter. One included regression in which variables with large correlations were used to predict the variable with outliers. The other method was using rows and columns to calculate the expected values for variables with outliers. An additional method used here to correct outliers was “score alteration”, which has also been recommended by Tabachnik and Fidell (2007). In this method, the score(s) on the outlying cases were changed so that they were deviant, but not as deviant as they were previously (Tabachnik & Fidell, 2007). For example, for Word Attack variable (all items), one of the outlying cases in the dyslexia subgroup was assigned a raw score that was one unit larger than the next most extreme score in the distribution and the other was assigned a raw score that was one unit smaller than the next most extreme score in the distribution. Upon correcting all outliers, the assumption of normality was re-checked for variables after each method of correction was used. For Word Attack (all items) and Elision variables, the row and column method was not successful in correcting the outlying cases. The score alteration and regression methods were also repeated several times in order to correct the outlying cases on these two variables.

To ensure that the pattern of results was not affected and remained the same, the main analyses, which are discussed next, were performed several times, namely (a) with uncorrected variables, (b) with outliers excluded, and (c) with corrected variables. The pattern of results obtained from these analyses indicated that for Word Attack (all items), TOWRE Word and Non-word Reading, and the three Rapid Naming variables, all analyses yielded similar findings. Therefore, the results presented for these variables are those performed on the uncorrected variables. There were two exceptions. One was the Elision variable which yielded different pattern of results. For this variable, the pattern of results for the analyses using the corrected variables was the same as the analyses performed excluding the outlying cases. Nonetheless, as was indicated earlier, a large number of cases were outliers on this variable (i.e., 3 in the dyslexia subgroup and 6 in the RA-match subgroup). Consequently, a large proportion of cases had to be transformed to normalize Elision. Given the fact that this large portion of transformed data was no longer the actual data obtained from participants, non-parametric tests were performed on the uncorrected Elision. The results of the non-parametric test were similar to the analyses performed without the outliers and those performed on the corrected variable. Consequently, the results presented for the Elision variable are those performed using non-parametric tests since the actual unchanged data was used here. The second exception was Word Attack variable that represented the polysyllables. As indicated earlier, the distribution of this variable for the dyslexia subgroup was markedly different than the other two subgroups. Correction of the variable or transformation of scores was also not possible for this variable given that 14 of 17 participants in the dyslexia subgroup had a score of zero. As a result, non-parametric tests were performed on the uncorrected Word Attack (polysyllables) variable.

Addressing Question 2: Group Differences in Reading and Cerebellar Measures

As discussed in the beginning of this section, the second question investigated whether a subgroup of children with dyslexia selected from the sample differed in their performance on any of the motor, cerebellar, reading, phonological, and rapid naming related measures, when compared to two control subgroups that were selected from the same sample and matched to the dyslexia subgroup based on (a) their reading level, and (b) chronological age.

This question was addressed using 10 one-way analyses of variance with Word Attack (all items), TOWRE Word and Non-word reading, Rapid Digit, Letter, and Object

Naming, Postural Stability Component, Muscle Tone ratio, Toe Tapping and Motor Component as the dependent variables. For these analyses, there was one between-subjects factor with three levels (Group: dyslexia vs. reading-age match vs. chronologically-age match) in each of the analyses. The results of these analyses are displayed in Table 14. As indicated earlier, the results reported are based on the analyses performed on the uncorrected variables. For Elision and Word Attack (polysyllables) variables, non-parametric tests were used that are described in the text.

Table 14

Mean and Standard Deviations on Word Attack, Elision, TOWRE Word and Non-word Reading, and Rapid Naming Raw scores and Motor/Cerebellar Components

Source	Dyslexia		RA-matched		CA-matched		df	*F*	*p*	Source of Effect
	M	SD	M	SD	M	SD	Between Subjects			
Word Attack (all items)	10.65	5.42	14.76	8.50	25.88	5.89	2	23.24	.00	(1 = 2) < 3
Word Attack (polysyllables)	.47	1.13	1.53	1.77	4.71	2.29	2	27.29[a]	.00	1 < 2 < 3
Elision	6.82	2.92	8.94	4.52	12.47	4.38	2	16.73[b]	.00	1 < 2 < 3
TOWRE Word Reading	35.65	15.31	33.35	14.20	57.82	11.39	2	16.46	.00	(1 = 2) < 3
TOWRE Non-word Reading	13.00	9.23	12.76	7.85	27.65	10.42	2	14.53	.00	(1 = 2) < 3
Rapid Digit Naming	46.71	12.64	46.59	12.53	40.00	8.44	2	1.94	.16	1 = 2 = 3
Rapid Letter Naming	49.35	12.77	46.76	11.62	45.88	16.34	2	.29	.75	1 = 2 = 3
Rapid Object Naming	70.71	14.45	68.00	9.51	66.82	13.15	2	.43	.65	1 = 2 = 3

Table 14 (cont’d)

Source	Dyslexia		RA-matched		CA-matched		df	*F*	*p*	Source of Effect
	M	SD	M	SD	M	SD	Between Subjects			
Postural Stability	.04	.88	-.29	1.05	-.36	.85	2	.88	.42	1 = 2 = 3
Toe Tapping	.16	.70	-.38	1.02	.007	1.03	2	1.49	.24	1 = 2 = 3
Motor	.50	.91	-.42	1.00	.15	.86	2	4.30	.02	1 > 2 = 3
Muscle Tone Ratio	.02	1.04	.06	.81	-.27	1.19	2	.52	.60	1 = 2 = 3

Note. For Word Attack (polysyllables) and Elision variables, non-parametric tests were used.

$^{a}\chi^{2}$ value for Kruskal-Wallis H omnibus test for Word Attck (polysyllables). $^{b}\chi^{2}$ value for Kruskal-Wallis H omnibus test for Elision.

Non-Word Reading

For the first analyses of variance, which included Word Attack (all items) raw scores as the dependent measures, the assumption of homogeneity of variance was met, as Levene's test of equality of error variance was not significant, F (2, 48) = 2.91, $p = .07$. As shown in Table 14, results of the analysis revealed a significant main effect of group, F (2, 48) = 23.24, $p < .001$. Comparisons indicated that the significant effect was due to the fact that non-word reading performance of participants in the dyslexia and RA-match subgroups were significantly poorer as compared to those in the CA-match subgroup, while the former two subgroups did not differ in their performance.

As indicated earlier, non-parametric tests were used for the Word Attack (polysyllables) variable. The Kruskal-Wallis H test, recommended by Howell (2010) was used. The results of this omnibus test indicated an overall significant difference between the three subgroups, $\chi^2(2, N = 51) = 27.29$, $p < .001$. According to Howell (2010), pairwise comparisons following the significant omnibus test can be performed using non-parametric Mann-Whitney U tests. Family wise error rate is controlled for up to three pairwise comparisons (Howell, 2010). The comparisons made using Mann-Whitney U tests indicated that as compared to both the RA-match ($U = 92.50$, $p = .03$) and CA-matched subgroups ($U = 10.50$, $p = .000$), participants in the dyslexia subgroup demonstrated significantly weaker phonological awareness skills. The difference in performance of RA- and CA-matched subgroups was also significant, $U = 41.00$, $p = .000$.

Phonological Awareness

As noted previously, a large portion of data had to be transformed to normalize the Elision variable. Consequently, non-parametric tests were performed on the uncorrected Elision since the actual unchanged data were entered in the analysis. The results for the Kruskal-Wallis H test indicated an overall significant difference between the three subgroups, $\chi^2(2, N = 51) = 16.733, p < .001$. The comparisons made using Mann-Whitney U tests indicated that as compared to both the RA-match ($U = 79.50, p = .02$) and CA-matched subgroups ($U = 30.50, p = .000$), participants in the dyslexia subgroup demonstrated significantly weaker phonological awareness skills. The difference in performance of RA- and CA-matched subgroups seemed to be marginally significant, $U = 87.50, p = .05$.

Rapid Naming Skills

The next three analyses of variance included corrected Rapid Letter, Digit, and Object Naming scores as their respective dependent variables. For these three analyses, type I error was controlled with the Bonferroni method to adjust the α level to .016 (α of .05 divided by 3). For all three analyses, Levene's test of equality of error variance indicated that variances were homogenous across subgroups (Rapid Digit Naming $F\ (2, 48) = .58, p = .57$; Rapid Letter Naming $F\ (2, 48) = 1.30, p = .282$; Rapid Object Naming $F\ (2, 48) =.94, p = .40$). The main effect of group was not significant for any of the three analyses, which indicated that the three subgroups did not differ significantly in the time they took to rapidly name digits ($F\ (2, 48) = 1.94, p = .16$), letters ($F\ (2, 48) =.29, p = .75$), and objects ($F\ (2, 48) =.43, p = .65$).

Word Reading Efficiency

Two analyses of variance investigated whether the subgroups differed in their fluency of word and non-word reading. Type I error was controlled with the Bonferroni method to adjust the α level to .025 (α of .05 divided by 2). The assumption of homogeneity of variance was met for both analyses as Levene's test of equality of error variance was not significant, (TOWRE Word Reading F (2, 48) = .48, p = .62; TOWRE Non-Word Reading, F (2, 48) = .56, p = .58). As shown in Table 14, results of the analyses revealed a significant main effect of group for both fluency in word reading, F (2, 48) = 16.46, $p < .001$, as well as fluency in non-word reading, F (2, 48) = 14.53, $p < .001$. Comparisons indicated that the dyslexia and RA-match subgroups did not differ in their fluency in word and non-word reading. However, both subgroups demonstrated significantly poorer word and non-word reading efficiency as compared to the CA-match subgroup.

Motor and Cerebellar Measures

The final four analyses of variance were conducted to investigate differences between the three subgroups on the postural stability component, toe tapping component, motor component, and muscle tone ratio. Type I error was controlled with the Bonferroni method to adjust the α level to .0125 (α of .05 divided by 4). For all four analyses, Levene's test of equality of error variance indicated that variances were homogeneous across subgroups (postural stability component F (2, 48) = .76, p = .48; toe tapping component F (2, 48) = 1.50, p = .23; motor component F (2, 48) = .05, p = .95; muscle tone ratio F (2, 48) = 1.07, p = .35).

As indicated in Table 14, the results of the analyses did not reveal a significant main effect of group for the postural stability component, F (2, 48) = .88, p = .42, toe tapping component, *F* (2, 48) = 1.49, *p* =. 24, and muscle tone ratio, *F* (2, 48) = .52, *p* = .60. In other words, the three subgroups did not differ significantly in their performance on postural stability tasks. They also did not significantly differ in the average time taken to tap their foot 10 times on the floor. Finally, the three subgroups did not differ significantly in their average muscle tone ratios. For postural stability component, recall from pages 142-143 of the Method chapter, that participant's weight was also obtained as a control measure. Because significant correlations were revealed between weights of participants and the deviation scores for postural stability (see p. 143), the main analysis for postural stability was repeated while the effect of weight was controlled statistically. However, results remained unchanged.

For the motor component, while the main effect of group was not significant at the corrected α level, it was significant at α. level of .05, *F* (2, 48) = 4.30, *p* < .05. Main effect comparisons indicated that on average participants in the dyslexia subgroup performed better on motor tasks as compared to those in the RA-matched control subgroup. However, no significant differences on motor tasks were found between the dyslexia and CA-matched subgroups. Surprisingly, no significant differences on motor tasks were found between the RA- and CA-match subgroups either. Nonetheless, the mean for motor performance was higher for the CA-match as compared to the RA-match subgroup.

Magnitude of Effects in Cohen's d

For all 12 analyses, the magnitude of effects was also explored in terms of effect size ratios which have been used in some of the previous studies (Fawcett & Nicolson, 1999; Fawcett et al., 1996; Savage et al., 2005a). There is a tendency for positive bias in effect size estimation as a result of design features such as smaller sample size or increased number of measures involved in a study (Thompson, 2002). Although it is recommended to correct for this bias, comparisons to other research findings have to be made using the uncorrected effect sizes (Durlak, 2009) because the correction for bias is usually not performed in other studies, as was the case here. Therefore, the effect sizes reported here were not corrected for bias so they could be compared to other research findings available.

The type of effect size calculated was Cohen's *d* which is a difference between standardized means (Cohen, 1988; Tabachnick & Fidell, 2007). To calculate effect sizes to contrast performance of the dyslexia subgroup with the RA-matched control subgroup, the mean obtained for the dyslexia subgroup was subtracted from the mean obtained for the RA-matched subgroup for each effect and then divided by the pooled standard deviation (Cohen, 1988, 1992). For example, as shown below, Cohen's *d* for non-word decoding as measured by Word Attack was calculated by subtracting the Word Attack mean for the dyslexia subgroup (i.e., 10.65) from the Word Attack mean for the RA-match control subgroup (i.e., 14.76), and dividing the quotient by the pooled standard deviation (6.72). The pooled standard deviation is calculated using the formula below (Thalheimer & Cook, 2002) with subscript *s* referring to standard deviation, *n* to number of subjects, *t* referring to the treatment group (i.e., dyslexia subgroup), and *c* referring to control condition (i.e., RA-match).

$$s_{pooled} = \sqrt{\frac{(n_t - 1)s_t^{\ 2} + (n_c - 1)s_c^{\ 2}}{n_t + n_c}} \qquad \text{Cohen's } d = \frac{14.76 - 10.65}{6.72} = .61$$

In a similar manner, effect sizes were calculated to contrast the dyslexia with the CA-match subgroup and to contrast the RA-match with the CA-match subgroup. According to

Cohen's *d* guidelines (1992), *d* values of .20 are considered small effects, *d* values of .50 are considered moderate effects, and *d* values of .80 are considered large effects.

The formula described above, which includes the pooled standard deviation across the two subgroups in the denominator, is used in cases in which the assumption of homogeneity of variance has been met (Olejink & Algina, 2000). This condition applied to all contrasts except for the contrast between the dyslexia and the CA-match subgroup on Elision variable. In cases where the assumption of homogeneity is not met, the standard deviation for the control group is typically used in the denominator of the formula (Olejink & Algina, 2000). The effect sizes calculated for the sample are shown in Table 15. Throughout the table, a negative score indicates poorer performance.

Table 15

Mean Effect Sizes Measured in Cohen's d for the Dyslexia Subgroup versus the Reading- and Chronologically-Age Match Controls and for the Reading-Age Match versus the Chronologically-Age Match Control

Variable	Dyslexia vs.		RA- vs.
	RA-matched	CA-matched	CA-matched
Word Attack (all items)	.59	2.77	1.57
Word Attack (polysyllables)	.74	2.49	1.39
Elision	.57	1.29[a]	.82
Word Reading Efficiency	-.16	1.69	1.96
Non-Word Reading Efficiency	-.03	1.53	1.66
Rapid Digit Naming	.01	.64	.64
Rapid Letter Naming	.22	.24	.06
Rapid Object Naming	.23	.29	.11
Postural Stability	.35	.48	.08
Toe Tapping	-.64	-.18	.39
Motor	-.99	-.41	.63
Muscle Tone Ratio	-.04	.27	.33

[a] Assumption of homogeneity not met, SD for control used for contrast.

As indicated in the table, moderate positive effect sizes were found for non-word reading and phonological awareness for contrasts that were made between the dyslexia subgroup and the RA-match control. This was combined with large positive effect sizes for contrasts of the dyslexia subgroup and the CA-match control as well as the RA- and

CA-match control subgroups for the same measures. In comparison to the moderately large effect size for non-word reading that represented all test items, the effect size for polysyllable non-words for the contrast between the dyslexia subgroup and the RA-match control was large. The effect sizes for word and non-word reading efficiency were small and negative for contrasts made between the dyslexia and the RA-match subgroups and large positive for contrasts between the dyslexia subgroup and the CA-match control as well as the RA- and CA-match control subgroups. Among the rapid naming measures, a moderate positive effect size was found for rapid digit naming when the dyslexia subgroup's performance was contrasted against the CA-match control and when the performance of the RA-match subgroup was contrasted against the CA-match subgroup.

Among motor and cerebellar measures, small positive effect sizes were found when the dyslexia subgroup's performance on postural stability was contrasted against performance of the RA- and CA-matched controls. Conversely, for contrast of the dyslexia subgroup with RA-match control, a moderately negative effect size was found for the toe tapping component. Additionally, a large negative effect size was also found for the motor component for contrast of the dyslexia subgroup with the RA-match control combined with a small negative effect size for contrast of the dyslexia subgroup and the CA-match control and a moderate positive effect size for contrast of the RA- and CA-match control subgroups.

Question 3:

Do any group differences in performance on the above reading and cerebellar related measures emerge when the effect of attention is controlled statistically?

To address this question, 12 one-way analyses of covariance were performed. Each analysis included one between-subjects factor with three levels (Group: dyslexia vs.

reading-age match vs. chronologically-age match). The effect of attention was adjusted by entering the ADHD index raw score as a covariate in each analysis. As discussed previously in the preliminary analysis section on page 189, because the ADHD scores were missing for one participant in the CA-match subgroup, the total number of participants was reduced from 51 to 50 for all analyses of covariance since the SPSS excluded the missing case from the analyses. Prior to conducting the analyses of covariance to address the research question, preliminary analyses had to be performed to ensure that assumptions related to the covariate were met for all analyses. These preliminary analyses are discussed below. For Word Attack (all items), Elision, Word and Non-word Reading Efficiency, and the three rapid naming measures, all analyses (including preliminary and main analyses addressing question three) were again performed several times, namely (a) with uncorrected variables, (b) with outliers excluded, and (c) with corrected variables. Because the pattern of results remained unchanged, the findings reported are all based on the analyses performed on the uncorrected variables.

Assumption of Linear Relationship between Covariate and Dependent variables

An ideal situation is when a covariate is significantly correlated with the dependent variable (Stevens, 1996). Additionally, the relationship between the dependent variable and covariate should be linear.

Covariate and reading measures. As a first step, correlations between the ADHD index raw score and all dependent variables were inspected using the SPSS Bivariate Correlation analysis program. Results indicated that the ADHD index was significantly correlated with Elision, r (50) = -.30, $p < .05$. No significant relationship was found between the ADHD index raw score and Word Attack (all items), r (50) = -.14, $p = .35$, Word Attack (polysyllables), *r* (50) = -.09, *p* = .53, Word Reading Efficiency, *r* (50) = -.24, *p* = .09, and Non-Word Reading Efficiency, *r* (50) = -.18, *p* = .21. To ensure that the assumption of linearity was not violated, the relationship between the ADHD index and the aforementioned measures was inspected using simple scatter plots that were created with the SPSS Graph function. No violations of linearity assumption (e.g., presence of curvilinear relationship) were observed.

Covariate and rapid naming measures. Among the three rapid naming measures, the ADHD index was only correlated with Rapid Digit Naming, *r* (50) = .42, *p* < .01, but not with Rapid Letter, *r* (50) = .27, *p* = .06, or Rapid Object Naming, *r* (50) = .25, *p* = .08. Nonetheless, inspections of simple scatter plots indicated no violation of linearity assumption.

Covariate and motor/cerebellar measures. Finally, the analyses revealed no significant correlations between the ADHD index and the postural stability component, *r* (50) = .14, *p* = .35, toe tapping component, *r* (50) = .25, *p* = .08, motor component, *r* (50) = .22, *p* = .13, or muscle tone ratio, *r* (50) = .19, *p* = .18. Observing simple scatter plots indicated that the assumption of linearity was not violated.

Assumption of Homogeneity of the Regression Slopes

According to this assumption, the slope of the regression line should be the same in each group for the covariate (Howitt & Cramer, 2008; Newton & Rudestam, 1999; Stevens, 1996; Tabachnik & Fidell, 2007). This assumption was inspected by investigating whether there was an interaction between the covariate and each of the dependent measures that were used in the analyses of covariance (Howitt & Cramer, 2008). This was performed using the SPSS GLM univariate program. Eleven analyses were conducted, each with the ADHD index entered as the covariate. There was one between-subjects factor with three levels in each of the analyses (i.e., Group: dyslexia vs. reading-age match vs. chronologically-age match).

The results of these analyses revealed that the interaction effects were not significant between ADHD index and any of the dependent measures (Word Attack – all items F (2, 44) = 1.53, $p = .23$; Word Attack – polysyllables F (2, 44) = 1.16, $p = .32$; Elision F (2, 44) = .27, $p = .76$; Word Reading Efficiency F (2, 44) = 2.53, $p = .09$; Non-word Reading Efficiency F (2, 44) = 1.01, $p = .37$; Rapid Digit Naming, F (2, 44) = 1.36, $p = .27$; Rapid Letter Naming F (2, 44) = 1.73, $p = .19$; Rapid Object Naming F (2, 44) = 1.65, $p = .20$; postural stability component F (2, 44) = .22, $p = .80$; toe tapping component, F (2, 44) = .01, $p = .99$; motor component, F (2, 44) = 1.08, $p = .35$; and muscle tone ratio, F (2, 44) = 1.05, $p = .36$). This indicated that for these measures the slope of the regression line was similar within the three subgroups for the ADHD index.

Addressing Question 3: Group Difference after Adjustment for Attention

As indicated earlier in this section, the third research question investigated whether there were any group differences in performance on the reading and cerebellar related measures when the effect of attention was controlled statistically. Twelve one-way analyses of covariance were performed with respective Conner's ADHD Index scores as the covariate in each case to address this question. There was one between-subjects factor with three levels (Group: dyslexia vs. reading-age match vs. chronologically-age match) in each of the analyses. The results of the analyses of covariance are displayed in Table 16. For Word Attack (polysyllables) variable, the analysis of covariance was performed on ranked data, a procedure known as non-parametric analysis of covariance (Olejnik & Algina, 1984, 1985). As was indicated earlier, the distribution of this variable was markedly different for the dyslexia subgroup as compared to the other two subgroups and no transformation or correction could be performed on this variable.

Table 16

Mean and Standard Deviations on Word Attack, Elision, TOWRE Word and Non-word Reading, and Rapid Naming Raw Scores and Motor/Cerebellar Components by Group After Adjustment for ADHD Index

Source	Dyslexia		RA-matched		CA-matched		df	F	p	Source of Effect
	M	SD	M	SD	M	SD	Between Subjects			
Word Attack (all items)	10. 50	5.42	14.86	8.50	25.81	6.06	2	20.36	.00	(1 = 2) < 3
Word Attack (polysyllables)	.37	1.13	1.59	1.77	4.73	2.36	2	23.65[a]	.00	1 < 2 < 3
Elision	7.28	2.92	8.66	4.52	12.01	4.36	2	5.87	.005	(1 = 2) < 3
Word Reading Efficiency	37.43	15.31	32.24	14.20	56.29	11.23	2	14.32	.00	(1 = 2) < 3
Non-word Reading Efficiency	13.67	9.23	12.34	7.85	26.73	10.40	2	12.03	.00	(1 = 2) < 3
Rapid Digit Naming	44.01	12.64	48.28	12.53	41.32	8.65	2	1. 84	.17	1 = 2 = 3
Rapid Letter Naming	47.40	12.77	47.99	11.62	47.34	16.63	2	.01	.99	1 = 2 = 3
Rapid Object Naming	69.12	14.45	69.00	9.51	68.07	13.33	2	.03	.97	1 = 2 = 3

Table 16 (cont'd)

Source	Dyslexia		RA-matched		CA-matched		df	F	p	Source of Effect
	M	SD	M	SD	M	SD	Between Subjects			
Postural Stability	-.001	.88	-.26	1.05	-.26	.81	2	.38	.67	1 = 2 = 3
Toe Tapping	.06	.70	-.31	1.02	.05	1.06	2	.82	.45	1 = 2 = 3
Motor	.45	.91	-.39	1.00	.22	.87	2	3.28	.05	1 > 2 = 3
Muscle Tone Ratio	-.09	1.04	.13	.81	-.15	1.18	2	.35	.71	1 = 2 = 3

[a] *F* statistic for Word Attack (polysyllables) variable derived from non-parametric ANCOVA.

Non-Word Reading and Phonological Awareness

In the first three analyses of covariance, Word Attack (all items), Word Attack (polysyllables), and Elision scores were entered as the dependent variables, respectively. Type I error was controlled with the Bonferroni method to adjust the α level to .016 (α of .05 divided by 3). For Word Attack (polysyllables), the analysis of covariance was performed on ranked data, a procedure known as non-parametric ANCOVA (Olejnik & Algina, 1984, 1985). To perform the non-parametric ANCOVA, data for the dependent variable and covariate were ranked separately across subgroups. The ranked data was then analyzed using the same procedure as those used with parametric ANCOVA (Oljenik & Algina, 1984).

After adjustment for ADHD index, the results for Word Attack (all items) were similar to the previous analysis of variance that was uncorrected for attention. That is, the main effect of group remained significant, F (2, 45) = 20.36, $p < .001$. Simple comparisons also indicated that the significant main effect of group for Word Attack (all items) was due to the fact that the dyslexia and RA-match subgroups performed significantly poorer as compared to the CA-match subgroup. However, the dyslexia and RA-match subgroups did not differ in their non-word reading performance.

For Word Attack (polysyllables), the results of the non-parametric analysis of covariance remained similar to those obtained from non-parametric tests performed prior to adjustment for attention, indicating a significant main effect of group, F (2, 46) = 25.82, $p < .001$. Simple comparisons also indicated that the dyslexia and RA-match subgroups performed significantly poorer on decoding polysyllable non-words as compared to the CA-match subgroup. Performance of the dyslexia subgroup on polysyllable non-words was also significantly poorer than that of the RA-match control.

Finally for Elision, the main effect of group was significant, (*F* (2, 46) = 5.87, *p* < .01). Nonetheless, unlike findings from previous non-parametric analyses, phonological awareness skills of the dyslexia subgroup did not differ significantly from that of the RA-match control after attention was controlled. However, performance of both the dyslexia and RA-match subgroups was still significantly poorer than that of the CA-match subgroup.

Rapid Naming Skills

The next three analyses of covariance included Rapid Letter, Digit, and Object Naming as their respective dependent variables. Type I error was controlled with the Bonferroni method to adjust the α level to .016 (α of .05 divided by 3). Similar to previous analyses which were uncorrected for attention, none of the present analyses revealed significant main effects for group (Rapid Digit Naming *F* (2, 46) = 1.84, *p* = .17; Rapid Letter Naming *F* (2, 46) = .01, *p* = .99; Rapid Object Naming *F* (2, 46) = .03, *p* = .97). In other words, after adjustment for attention, the three subgroups still did not differ in the time they rapidly named digits, letters, and objects.

Word Reading Efficiency

Two analyses of covariance investigated whether the subgroups differed in their fluency in word and non-word reading efficiency after adjustment for the ADHD index scores. Type I error was controlled with the Bonferroni method to adjust the α level to .025 (α of .05 divided by 2). Similar to findings obtained in the analysis of variance that was uncorrected for attention, a significant main effect of group was obtained for both analyses of covariance (Word Reading Efficiency *F* (2, 46) = 14.32, *p* < .001; Non-word Reading Efficiency, *F* (2, 46) = 12.03, *p* < .001). Comparisons indicated that after attention was controlled, the dyslexia subgroup’s word and non-word reading efficiency still did not differ from that of the RA-match control subgroup. However, both subgroups performed significantly poorer as compared to the CA-match control subgroup.

Motor and Cerebellar Measures

The final four analyses of covariance were conducted to investigate if any differences were found among the three subgroups on motor and cerebellar measures after adjustment for the ADHD index scores. The postural stability component, toe tapping component, motor component, and muscle tone ratio were the dependent measures in each of the analyses. Type I error was controlled with the Bonferroni method to adjust the α level to .0125 (α of .05 divided by 4). The pattern of results for all four analyses remained similar to those obtained in the analyses of variance that were uncorrected for attention. No significant main effects of group were found for the postural stability component, F (2, 46) = .38, p = .67, toe tapping component, F (2, 46) =.82, p = .45, and muscle tone ratio, F (2, 46) =.35, p = .71. This indicated that after adjustment for attention, the subgroups still did not differ in their performance in any of these measures.

Similar to results from analysis of variance, the main effect of group for the motor component remained significant at the .05 level , F (2, 46) = 3.28, p = .05, but not at the Bonferroni adjusted α level (i.e., α = .0125). This indicated that after adjustment for attention, on average participants in the dyslexia subgroup were still better in their performance on motor tasks as compared to those in the RA-matched control. Similar to the pattern of results obtained from the previous analysis of variance, no differences were found on motor tasks between the dyslexia and CA-match subgroups as well as between the RA- and CA-match subgroups.

Magnitude of Effects in Cohen's d

Similar to previous analyses, the magnitude of effects were explored in these 12 analyses of covariance using effect size ratios. For each measure, Cohen's *d* effect sizes were calculated to contrast the dyslexia subgroup's performance with the RA and CA-match controls respectively. Cohen's *d* effect sizes were also calculated to contrast performance of the RA-match control with the CA-match control subgroup. The effect sizes calculated for the sample are shown in Table 17.

Table 17

Mean Effect Sizes Measured in Cohen's d for the Dyslexia Subgroup versus the Reading- and Chronologically-Age Match Controls and for the Reading-Age Match versus the Chronologically-Age Match Control After Adjustment for ADHD Index

Variable	Dyslexia vs.		RA-matched vs.
	RA-matched	CA-matched	CA-matched
Word Attack (all items)	.63	2.75	1.52
Word Attack (polysyllables)	.81	2.46	1.56
Elision	.37	1.32	.78
Word Reading Efficiency	-.36	1.44	1.93
Non-Word Reading Efficiency	-.16	1.37	1.62
Rapid Digit Naming	.35	.25	.66
Rapid Letter Naming	.05	.004	.05
Rapid Object Naming	.01	.08	.08
Postural Stability	.28	.32	0.00
Toe Tapping	-.44	-.01	.36
Motor	-.90	-.26	.67
Muscle Tone Ratio	-.24	.05	.29

Following Olejnik and Algina's (2000) recommendations for measuring Cohen's *d* effect size in an analysis of covariance, the means in contrasts were the adjusted means rather than the actual pre-adjusted means. Throughout Table 17, a negative score indicates poorer performance. As indicated in the table, for contrasts made between the dyslexia subgroup and both the RA- and CA-matched controls, the pattern of effect sizes for non-word reading (including all items and polysyllables) remained similar to the previous analyses that were uncorrected for attention. However, there seemed to be a small increase in the effect size for non-word reading for contrast of the dyslexia subgroup with the RA-match control for both non-word reading measures (i.e., from .59 to .63 for all items; from .74 to .81 for polysyllables). Examining the mean for non-word reading in both subgroups indicated that attention might have partly explained performance on this task for some participants in the RA-match control subgroup since the overall mean for this subgroup increased slightly (i.e., from 14.76 to 14.86 for all items; from 1.53 to 1.59 for polysyllables) after adjustment for attention. For Elision, the pattern of effect sizes remained large for contrasts of the dyslexia subgroup with the CA-match control and contrast of the RA-match with the CA-match control subgroups. Nonetheless, a noticeable decrease in effect size was observed for contrast of the dyslexia subgroup with the RA-match control (i.e., from .57 to .37). Examining the mean for Elision in both subgroups indicated that attention might have partly explained the poor performance on this task for some participants in the dyslexia subgroup. More specifically, the overall mean for the dyslexia subgroup was 7.28 after adjustment for attention. However, prior controlling for the effect of attention, the mean for the dyslexia subgroup was 6.82, indicating that the presence of attention problems may be the reason for the lower overall group mean.

The pattern of effect sizes among the rapid letter and object naming measures remained similar to those obtained for the analyses of variance that were uncorrected for attention. For Rapid Digit Naming, the effect size for the contrast of the dyslexia subgroup with the CA-match control seemed to decrease from moderate (.64) to small (.25). Additionally, there was a slight increase in the effect size for the contrast of the dyslexia subgroup with the RA-match control (i.e., from .01 to .25). Once again, attention seemed to partly explain the changes observed in these effect sizes. That is, the group mean for the dyslexia subgroup indicated a somewhat better performance in the digit naming task after adjustment for attention (i.e., less time taken to complete the task) (i.e., from 46.71 seconds to 44.01 seconds).

Among motor and cerebellar measures, a reduction in effect size was found in contrast of the dyslexia subgroup with the RA-match control (i.e., from .35 to .28) as well as in contrast of the dyslexia subgroup with the CA-matched control (i.e., from .48 to .42) on postural stability. Examining the means obtained in postural stability for the subgroups indicated that attention might have partly explained performance on postural stability in some participants in the dyslexia and RA-match subgroups, as the overall mean for the subgroups indicated somewhat better performance on the postural stability tasks after adjustment for attention (i.e., lower deviation scores) (i.e., from .04 to -.001 in the dyslexia subgroup, from -.29 to -.26 in the RA-match control). Finally, for the motor component, while the pattern of effect size for contrasts of the dyslexia subgroup with both the RA- and CA-match controls remained unchanged, there was some degree of reduction in both effect sizes (i.e., -.99 to -.90 for the first contrast and from -.41 to -.26 for the second contrast). Examining the means for the subgroups indicated that attention might have partly explained performance on motor tasks for some participants in the RA-match control subgroup since the overall mean for this subgroup increased to some extent (i.e., from -.42 to -.39) after adjustment for attention.

Question 4:

Does the cerebellar deficit provide a good explanatory model at the individual level?

To investigate individual differences in performance in reading and cerebellar measures in the dyslexia and RA-match subgroups, the procedure formerly used by White et al. (2006) was followed. As discussed in the Introduction (pp. 87-91), the present study differed in some respects from White et al.'s.

Calculating Summary Scores

A final step prior to individual analysis in White et al.'s (2006) study was creating summary factors for tasks in their study that belonged to one modality. The factors created for the present study corresponded to the tasks that were used in this project. Where applicable, summary factors were computed, similar to White et al.'s study in that standardized scores were averaged (calculated in relation to performance of the reading-age match subgroup) for each participant on each group of tasks. As indicated previously, unlike White et al.'s study, in which standard scores were used to create reading related factors, in this study raw scores were used, since comparisons made in a reading-level design are based on using raw scores. Prior to calculating summary scores, signs for *z*-standardized scores had to be reversed for rapid naming, postural stability, toe tapping, peg moving, and muscle tone measures in order to ensure that positive scores indicated good performance and negative scores indicated poor performance in this study as in White et al.'s study. This was because in these tasks, larger raw scores (i.e., positive *z*-standardized scores) indicated poorer performance, namely (a) longer time taken to complete rapid naming, toe tapping, and peg moving tasks; (b) greater amount of deviation upon administering the balance challenge in postural stability task; and (c) larger muscle tone ratios reflecting poorer muscle tone. Negative scores, however, indicated better performance, namely (a) less time taken to complete rapid naming, toe tapping, and peg moving tasks; (b) less deviation upon administering the balance challenge in the postural stability task; and (c) a smaller muscle tone ratio reflecting stronger muscle tone.

Reading factor. To illustrate that the dyslexia and reading-age match subgroups were matched on their reading level, the reading age on the Word Identification task was saved as standardized *z*-scores for use in individual analysis. Positive scores indicated good performance and negative scores indicated poor performance.

Rapid naming summary factor. Given that three rapid naming measures were used in this study and the results of the analyses of variance did not reveal differences in findings for these measures, one summary factor was calculated for the purposes of individual analysis. An alpha-numeric rapid naming summary factor was calculated. Raw scores for alpha-numeric rapid naming scores were directly saved as *z*-standardized scores. Upon reversing signs for standardized *z*-scores, an average of the *z*-standardized scores on Rapid Letter and Digit Naming was calculated across both the dyslexia and RA-match subgroups. Positive scores indicated good performance and negative scores indicated poor performance.

Non-word decoding and phonological awareness. To investigate individual differences between participants in the dyslexia subgroup and those in the RA-match control on non-word decoding and phonological awareness, raw scores on Word Attack (all items), Word Attack (polysyllables), and Elision were each saved as *z*-standardized scores and directly used in individual analyses. For all tasks, positive scores indicated good performance and negative scores indicated poor performance.

Word and non-word reading efficiency. Individual differences between the dyslexia and RA-match subgroups on word and non-word reading efficiency was investigated, whereby raw scores on the TOWRE word and non-word reading tasks were each saved as *z*-standardized scores and directly used in individual analyses. For both tasks, positive scores indicated good performance and negative scores indicated poor performance.

Postural stability summary factor. In order to exhibit individual differences in postural stability, a postural stability summary factor was created. That is, raw scores obtained on the two postural stability tasks were first directly saved as *z*-standardized scores. Upon reversing the signs of *z*-standardized scores, the two postural stability tasks were averaged for each participant across the dyslexia and RA-match subgroups. Positive scores on the postural stability summary factor corresponded to good performance and negative scores to poor performance.

Toe tapping summary factor. Similar to the postural stability summary factor, raw scores on both toe tapping measures were first directly saved as standardized *z*-scores. Upon reversing all positive and negative signs on toe tapping *z*-scores, a toe tapping summary factor was created whereby standardized *z*-scores on both tasks were averaged across individuals in both subgroups. Positive scores corresponded to good performance and negative scores indicated poor performance.

Motor summary factor. A motor summary factor was created by first saving raw scores on the bead threading and peg moving tasks directly as *z*-standardized scores. Upon reversing signs for the peg moving *z*-standardized scores, the *z*-scores on both tasks were averaged for each participant across the dyslexia and RA-match subgroups. Positive scores indicated good performance and negative scores indicated poor performance.

Muscle tone ratio. To display individual differences in muscle tone ratio, raw scores obtained on this measure were directly saved as *z*-standardized scores and used in the individual analysis upon sign reversal. Positive scores indicated good performance and negative scores indicated poor performance.

Controlling for Attention

As indicated earlier, an added element in this study was to investigate individual differences in reading and cerebellar measures when attentional difficulties were controlled. To remove the effect of attention from the measures used for the individual analyses, 10 regressions were conducted using the SPSS Linear Regression Analysis program. In all regressions the *z*-standardized ADHD index score across participants in the dyslexia and RA-match subgroups was entered as the independent variable. *Z*-Standardized Word Attack (all items), Word Attack (polysyllables), Elision, TOWRE Word and Non-word reading scores, alpha-numeric RAN summary factor, postural stability summary factor, toe tapping summary factor, motor summary factor, and *z*-standardized muscle tone ratios across the two subgroups were the respective dependent measures in each regression. Using the save function in the SPSS Linear Regression program, unstandardized residuals were saved for each participant in the data.

Addressing Question 4: Individual Differences

Similar to White et al. (2006), in order to investigate individual differences, the individual data were displayed in figures that were created using the Excel program. Nineteen figures were created. As indicated earlier, the cut-off point for identifying extreme poor cases in this study was -.67 instead of -1.65. This cut-off point (i.e., -.67) is shown by a broken line in the figures and participants below this line are considered outliers and labeled. On all figures, values on the y-axis are *z*-scores. Additionally, some lines in the figures illustrate performances that were either equal or very close. When performances were equal, the lines that represented performance of several participants remained similar in thickness to those lines that represented performance of only one participant. When performances were close, lines representing performance of participants became thicker than the other lines. For clarity, all lines that corresponded to performance of several participants (i.e., either equal or close performances) are labeled with case numbers to indicate the number of participants represented by those lines. Figure 19 on the next page represents the reading factor.

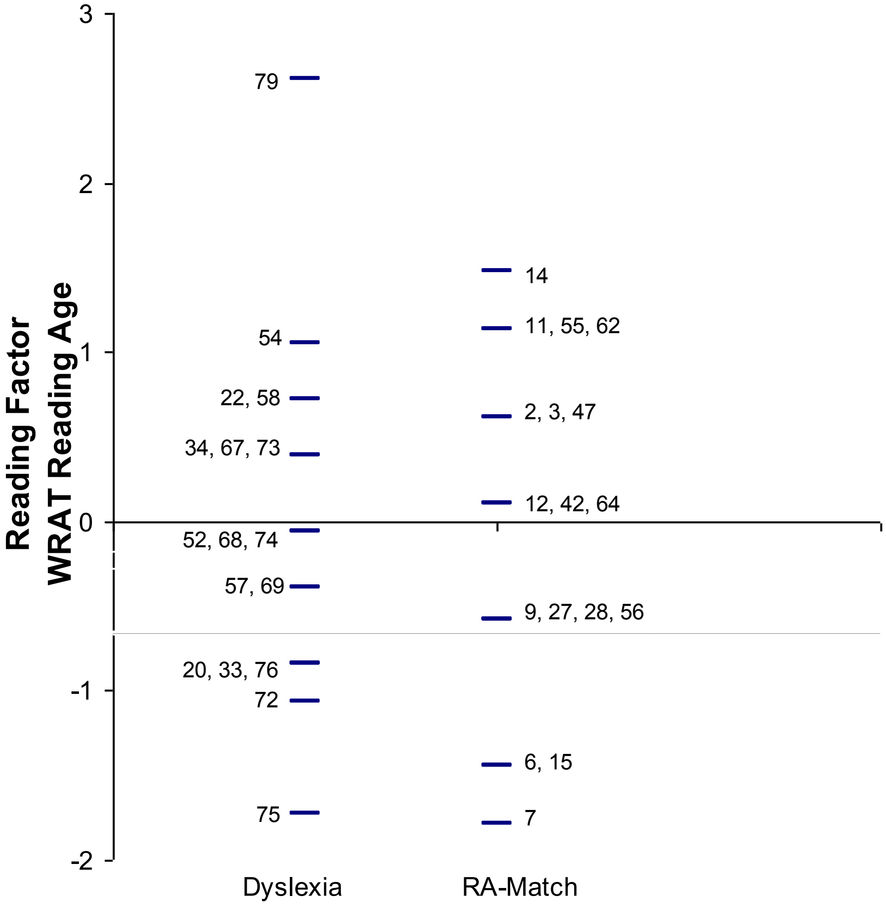
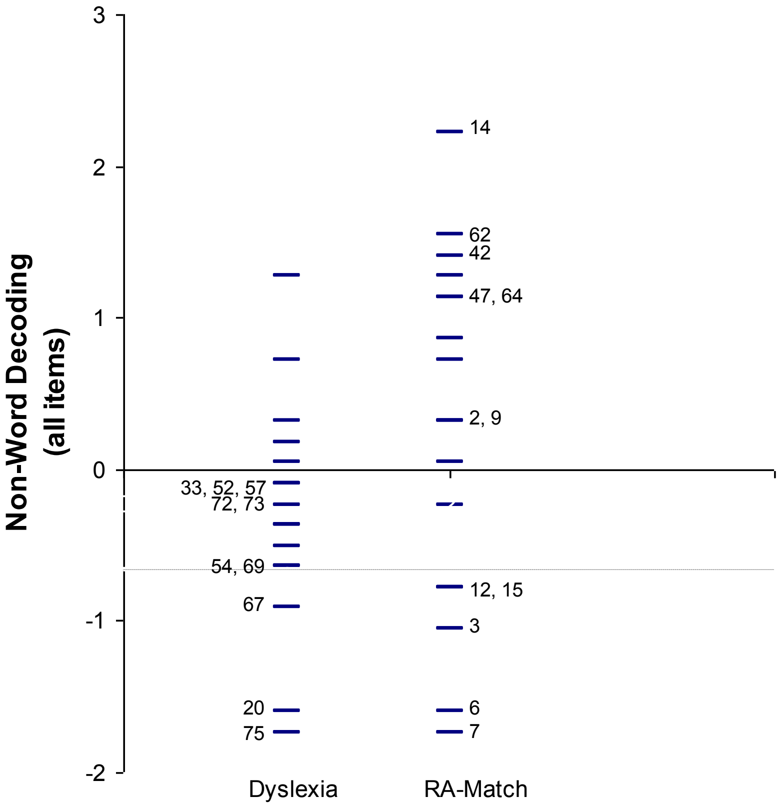


Figure 19. Graph of individual reading ages on single word reading.

As indicated earlier, the dyslexia and RA-match subgroups were matched on their reading level, hence, prior to investigating individual differences on the outcome measures, the purpose of this figure was to illustrate that the *z*-standardized reading age scores on the Word identification task for the dyslexia and RA-match subgroups were comparable at an individual level. Consequently, the distribution of scores should be relatively similar, as is the case here. As shown in Figure 19, only one case (i.e., 79) in the dyslexia subgroup seems to have the highest score on the word reading task. Among participants in the RA-match control subgroup, three cases (i.e., 6, 7, and 15) seem to have the poorest performance on the word reading task with scores falling below the .67 cut-off line.

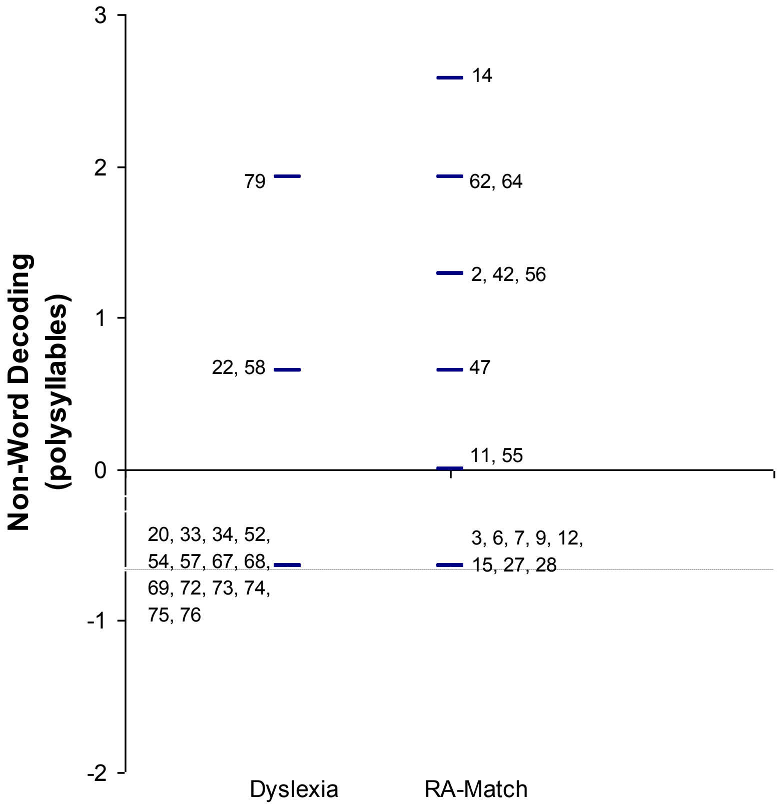
On the next page, Figures 20 and 21 represent individual differences on non-word decoding prior to controlling for attention.



3 < -.67 vs. 5 < -.67

20

Figure 20. Graph of individual performance for for non-word decoding (all items).



0 < -.67 vs. 0 < .67

21

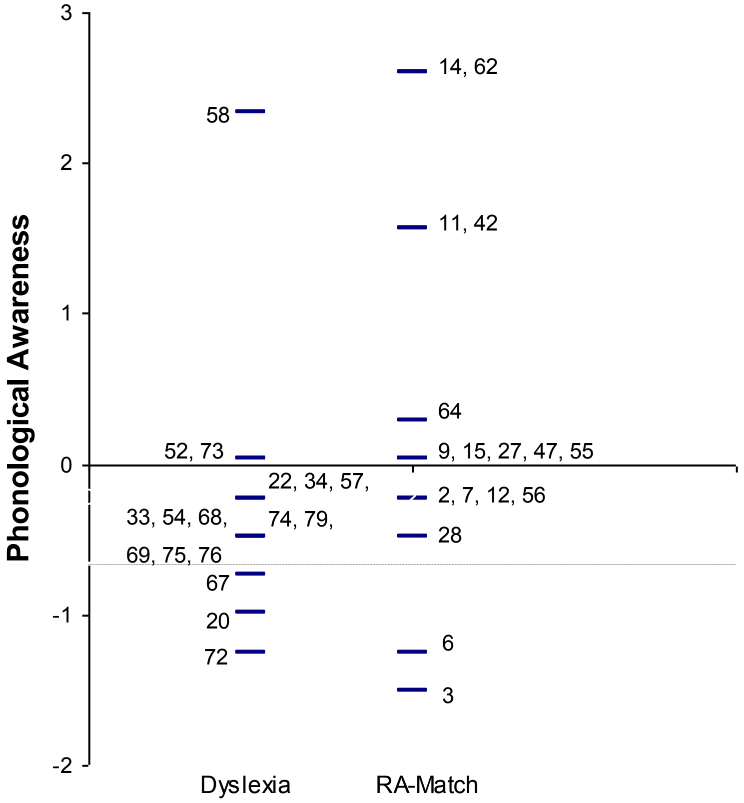
Figure 21. Graph of individual performance for polysyllable non-words.

Individual Performances Prior to Control for Attention

Non-word decoding. As shown in Figure 20, non-word decoding, as measured by the Word Attack, does not appear to distinguish well between participants in the dyslexia subgroup and those in the RA-match control at an individual level. In contrast to 3 participants in the dyslexia subgroup, 5 in the RA-match control are identified as extreme poor cases. Three of these five cases (i.e., 6, 7, and 15) were also outliers on the Word Identification task, as was shown in Figure 19. In both subgroups, two cases have non-word decoding scores falling more than 1 SD below the mean. The dispersion of scores within the two subgroups seems to be somewhat different, with more participants in the RA-match control as compared to the dyslexia subgroup lying above the mean of zero. More specifically, within the dyslexia subgroup, scores range from -1.73 to 1.26. Within the RA-match control subgroup, scores seem to be more dispersed, ranging from -1.73 to 2.22.

In Figure 21, performance on polysyllable non-words of the Word Attack are displayed. As shown in the figure, none of the participants in the dyslexia and RA-match subgroup have scores below the .67 cut-off line. Nonetheless, the dispersion of scores within the two subgroups is different and more distinct than the dispersion of scores on all non-word items shown in Figure 20. For the polysyllable items, only 3 participants in the dyslexia subgroup as compared to 9 in the RA-match control have obtained scores above the mean of zero.

On the next page, Figure 22 displays individual differences in phonological awareness prior to controlling for the effect of attention.

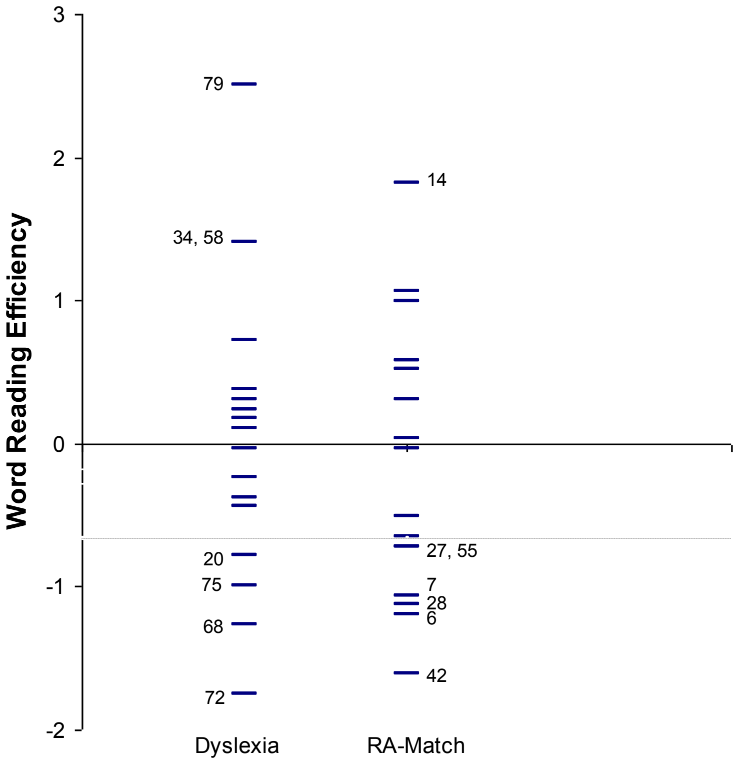


3 < -.67 vs. 2 < .67

Figure 22. Graph of individual performance for phonological awareness.

Phonological awareness. As shown in Figure 22, phonological awareness, as assessed by the Elision subtest, seems to distinguish between the two subgroups to some extent at an individual level. Among the dyslexia subgroup, Elision *z*-standardized scores range from -1.25 to 2.33. Among the RA-match subgroup, this range is from -1.50 to 2.56. While the range of scores is similar within both subgroups, they seem to cluster somewhat differently. Within the dyslexia subgroup, there are only three cases with Elision scores falling below the -.67 cut-off line. However, the remaining cases seem to cluster more between the mean and the cut-off point. There is only one case (i.e., 58) with an Elision *z*-standardized score of 2.18 which seems to be the outlying case within participants in the dyslexia subgroup. Within the RA-match control subgroup, two cases (i.e., 3, 6) have obtained Elision scores below the cut-off point. One of these cases (i.e., 6) was an outlier on the word reading task in Figure 19. In contrast to the dyslexia subgroup, the remaining cases in the RA-match subgroup seem to be more dispersed between *z*-standardized scores of -.48 to 2.60.

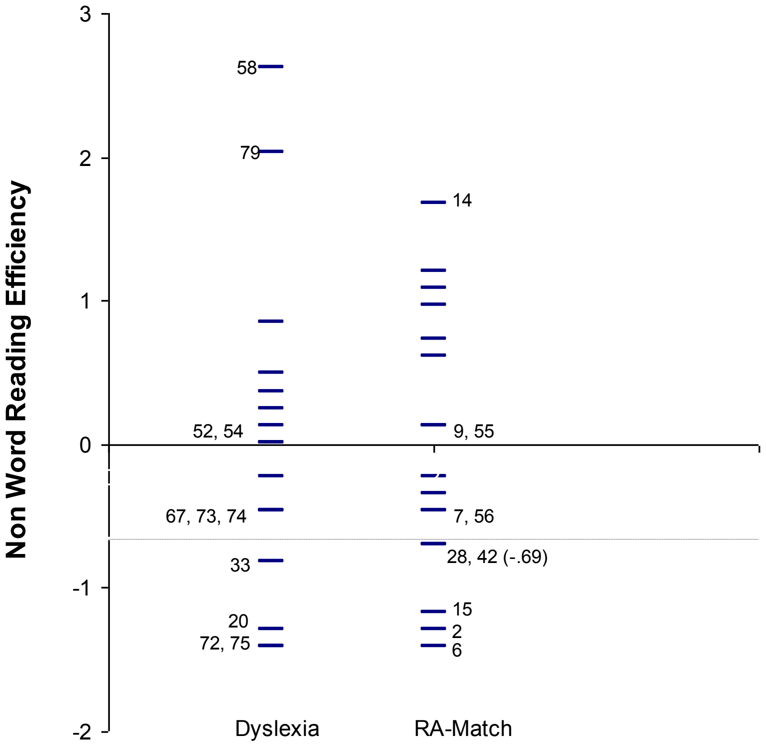
On the next page, Figure 23 and 24 display individual differences in word and non-word reading efficiency prior to controlling for the effect of attention.



4 < -.67 vs. 6 < -.67

23

Figure 23. Graph of individual performance for word reading efficiency.



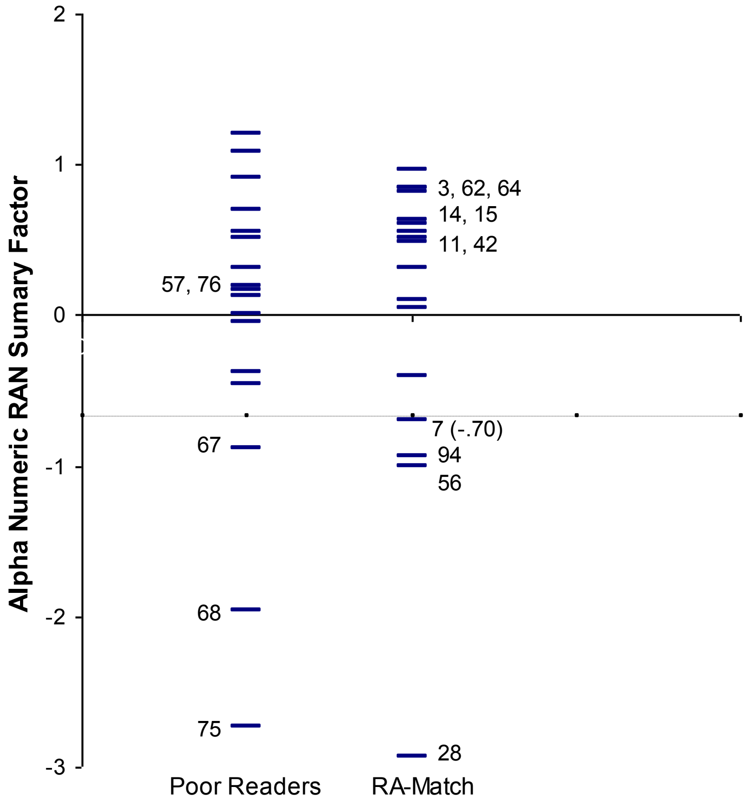
4 < -.67 vs. 5 < .67

24

Figure 24. Graph of individual performance for non-word reading efficiency.

Word and non-word reading efficiency. As shown in Figures 23 and 24, performance on word and non-word reading efficiency did not seem to differentiate between the dyslexia and the RA-match subgroups. In both subgroups, a comparatively similar proportion of participants seems to fall below and above the .67 cut-off point. Within the extreme poor cases in the RA-match subgroup, two (i.e., 6 and 7) on the word reading efficiency and two (i.e., 6 and 15) on the non-word reading efficiency measures are the outlying cases on the word reading task shown in Figure 19. However, even if these cases are ignored, the two measures do not seem to distinguish well between the two subgroups.

Figure 25 displays individual differences in alpha-numeric rapid naming measures prior to controlling for the effect of attention.

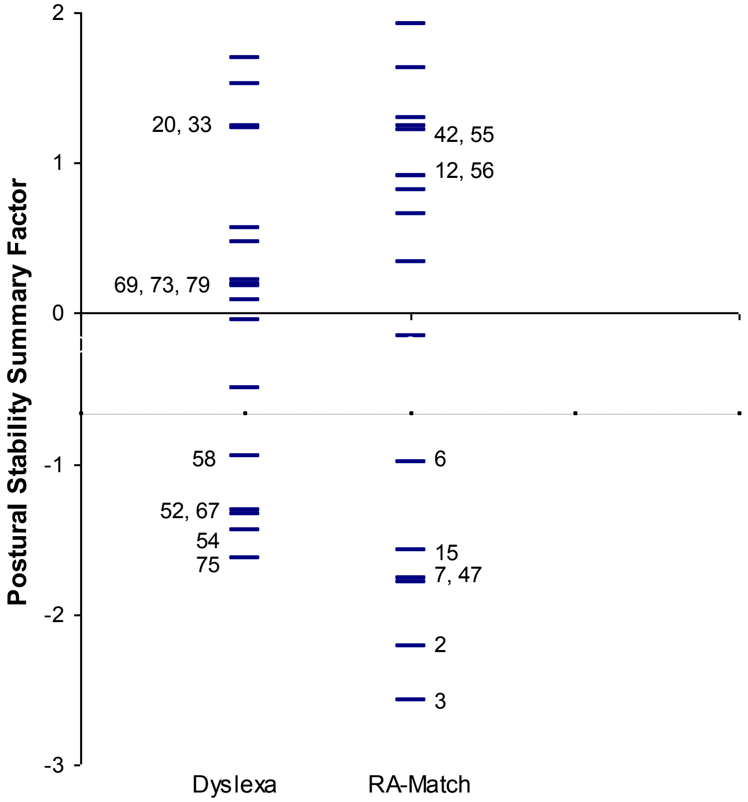


3 < -.67 vs. 4 < -.67

Figure 25. Graph of individual performance for alphanumeric rapid naming summary factor.

Alpha numerical rapid naming summary factor. Performance on alpha numeric rapid naming tasks, shown in Figure 25, did not seem to differentiate between the participants in the dyslexia and RA-match subgroups. In both subgroups, a similar proportion of children fell above the .67 cut-off point. The proportion of the cases falling below .67 cut-off point was also similar (i.e., 3 in the dyslexia vs. 4 in RA-match). One of the extreme poor cases in the RA-match control subgroup (i.e., 7) was also an outlier on the word reading task in Figure 19. Nonetheless, even if this case is ignored, the alphanumeric rapid naming summary factor does not seem to differentiate between the two subgroups.

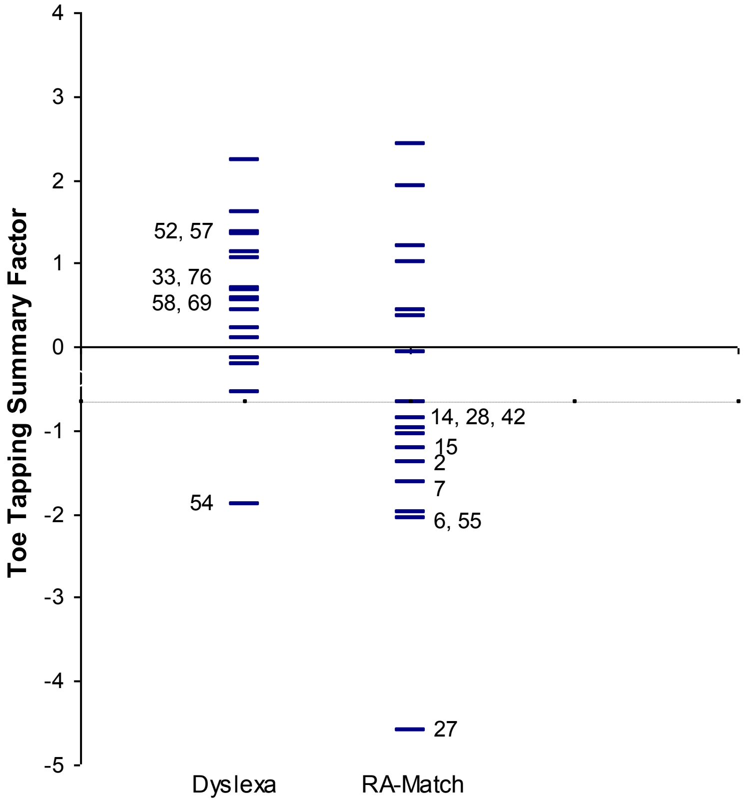
On the next page, Figures 26 and 27 present individual differences on postural stability and toe tapping summary factors prior controlling for the effect of attention.



5 < -.67 vs. 6 < -.67

26

Figure 26. Graph of individual performance for postural stability summary factor.



1 < -.67 vs. 9 < -.67

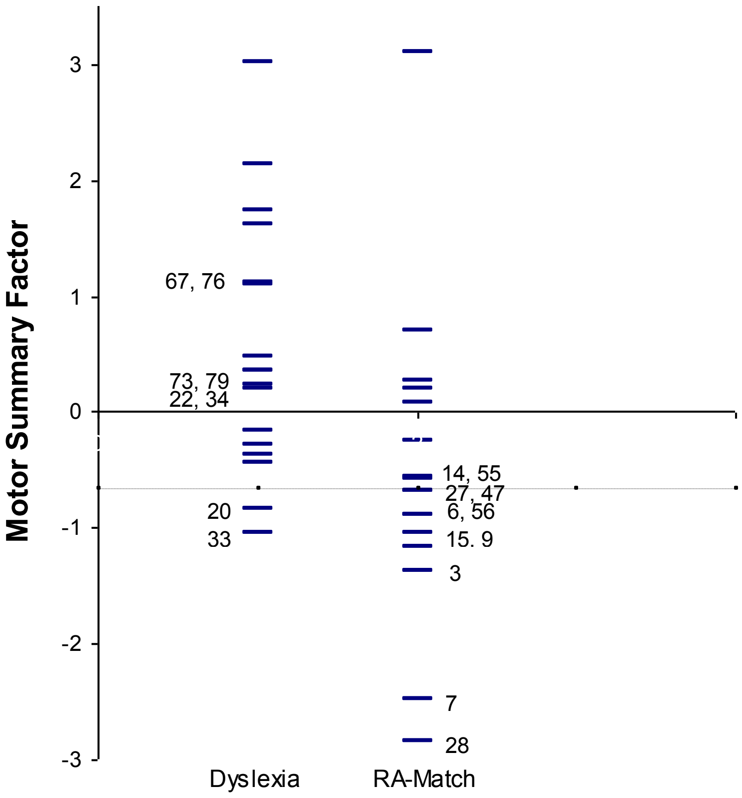
27

Figure 27. Graph of individual performance for toe tapping summary factors.

Postural stability summary factor. Figure 26 illustrates individual performances on the postural stability tasks. As seen in the pattern of performances, the postural stability summary factor does not seem to distinguish between the two subgroups. Similar to the alphanumeric rapid naming summary factor, in the postural stability summary factor a similar proportion of children in each subgroup fell below the .67 cut-off point (i.e., 5 in dyslexia vs. 6 in RA-match). Postural stability *z*-scores in extreme poor cases in the dyslexia subgroup ranged from -.95 to -1.63, and those in the RA-match control ranged from -.99 to -2.58. Within the RA-match control, three cases (i.e., 6, 7, and 15) are the outlying cases on the word reading task shown in Figure 19. If these cases are ignored, the postural stability summary factor seems to improve in its ability to differentiate between the two subgroups as the proportion of the extreme poor cases in the dyslexia subgroup increases comparatively (i.e., 5 in dyslexia vs. 3 in RA-match). Yet, only two of the five extreme poor cases in the dyslexia subgroup were also identified as extreme poor cases on other literacy measures.

Toe tapping summary factor. Figure 27 illustrates individual performance on the toe tapping summary factor. Consistent with group findings, the pattern of *z*-scores in the figure indicates that in all, participants in the dyslexia subgroup performed better on the toe tapping task as compared to those in the RA-match control. In contrast to only one extreme poor case in the dyslexia subgroup with a *z*-score of -1.88, there are nine extreme poor cases in the RA-match control with *z*-scores ranging from -.87 to -4.58. Within the RA-match control subgroup, three of the extreme poor cases (6, 7, and 15) are among the outlying cases on the word reading task shown in Figure 19. However, even if these cases are ignored, the pattern appearing in the toe tapping summary factor remains unchanged.

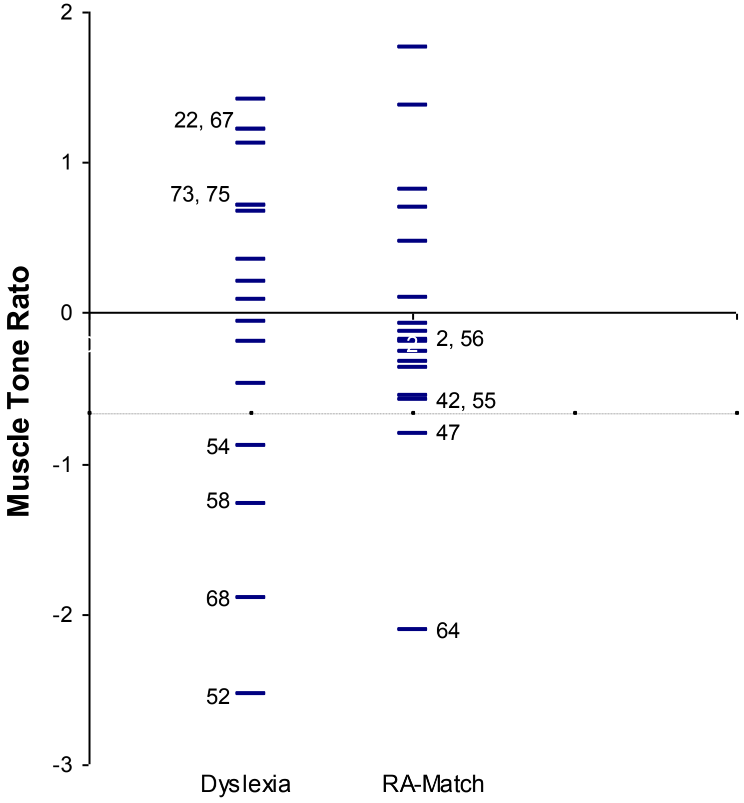
On the next page, Figures 28 and 29 display individual differences on the motor summary factor and muscle tone ratio prior to controlling for attention.



2 < -.67 vs. 9 < -.67

28

Figure 28. Graph of individual performance for motor summary factor.



4 < -.67 vs. 2 < -.67

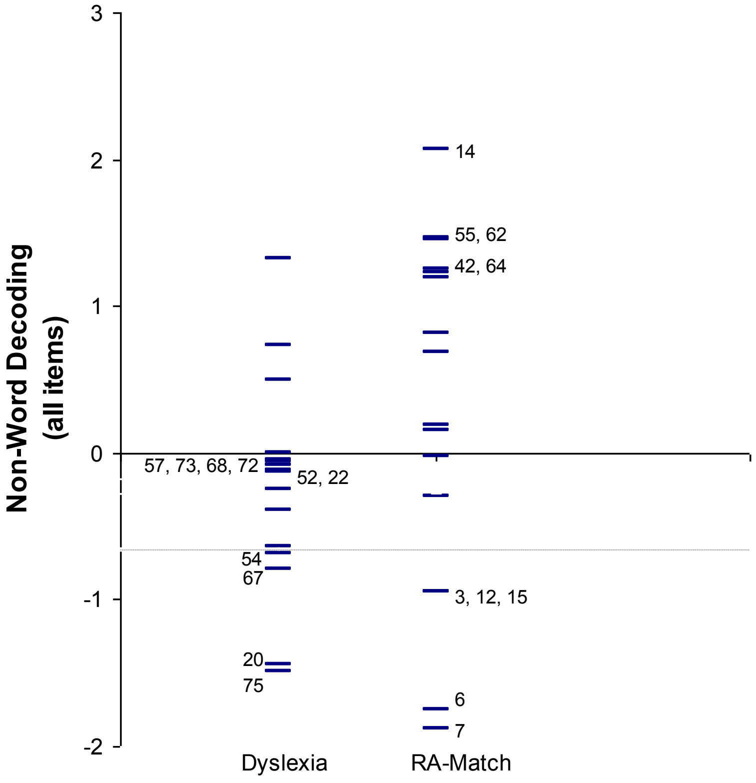
29

Figure 29. Graph of individual performance for muscle tone ratio.

Motor summary factor. Figure 28 shows individual performances on the motor summary factor. Similar to the toe tapping summary factor, the trend in the motor summary factor seems to indicate better overall performance in the dyslexia subgroup as compared to the RA-match control. Although only two cases in the dyslexia subgroup fell below the .67 cut-off point (*z*-scores -.84, -1.05), nine cases in the RA-match control were identified as extreme poor cases with *z*-scores ranging from -.68 to 2.85). Once again, cases 6, 7, and 15 are among the extreme poor cases in the RA-match control. However, the trend seen in the motor summary factor does not change even if these cases are ignored.

Muscle tone. Figure 29 represents individual performances on the final cerebellar measures, namely the muscle tone. As seen in the figure, the pattern of performances seems to suggest a trend. That is, while four cases in the dyslexia subgroup fell below the .67 cut-off point, only two in the RA-match control were identified as extreme poor cases. Nonetheless, a closer look at the extreme poor cases in both subgroups indicates that except for the case (i.e., 68) in the dyslexia subgroup which was also an extreme poor case on the alphanumeric task and word and non-word reading efficiency, none of the other cases were identified as extreme poor cases on any of the literacy factors.

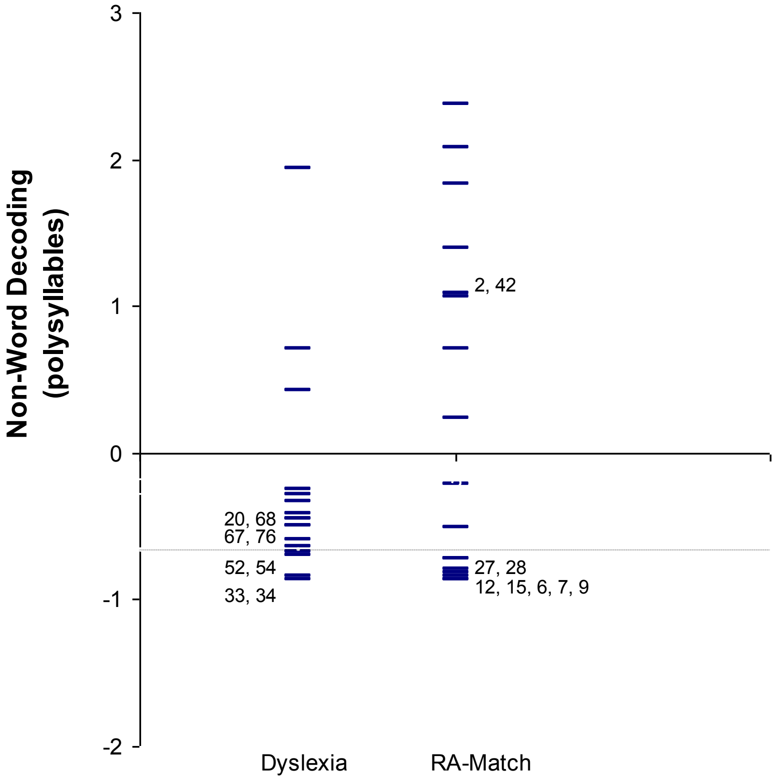
Figures 30 and 31 on the next page display individual differences on non-word decoding after the effect of attention was controlled statistically.



4 < -.67 vs. 5 < -.67

30

Figure 30. Graph of individual performance for non-word decoding (all items) with ADHD index partialled out.



4 < -.67 vs. 7 < -.67

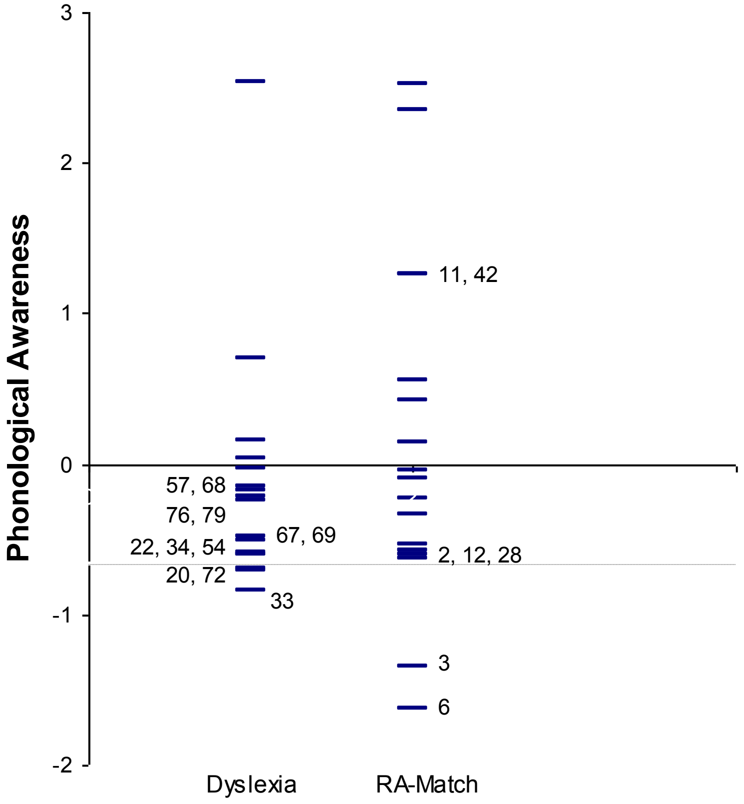
31

Figure 31. Graph of individual performance for polysyllable non-words with ADHD index partialled out.

Individual Performances Upon Control for Attention

Non-word decoding. As shown in Figure 30, upon adjustment for attention, non-word decoding still does not seem to distinguish between the two subgroups. Similar to Figure 20 discussed earlier, a similar proportion of participants in both subgroups are identified as extreme poor cases (i.e., 4 in the dyslexia and 5 in the RA-match). The dispersion of scores in Figure 30 remains similar to Figure 20 (prior controlling for attention). In other words, upon adjustment for attention, more participants in the RA-match control as compared to the dyslexia subgroup seem to have non-word decoding scores falling above the mean. The overall pattern of performance on polysyllable non-words after adjustment for attention, as shown in Figure 31, also remains similar to Figure 21 (prior controlling for attention). Fewer participants in the dyslexia subgroup (i.e., 3) as compared to the RA-match control (i.e., 8) have scores falling above the mean.

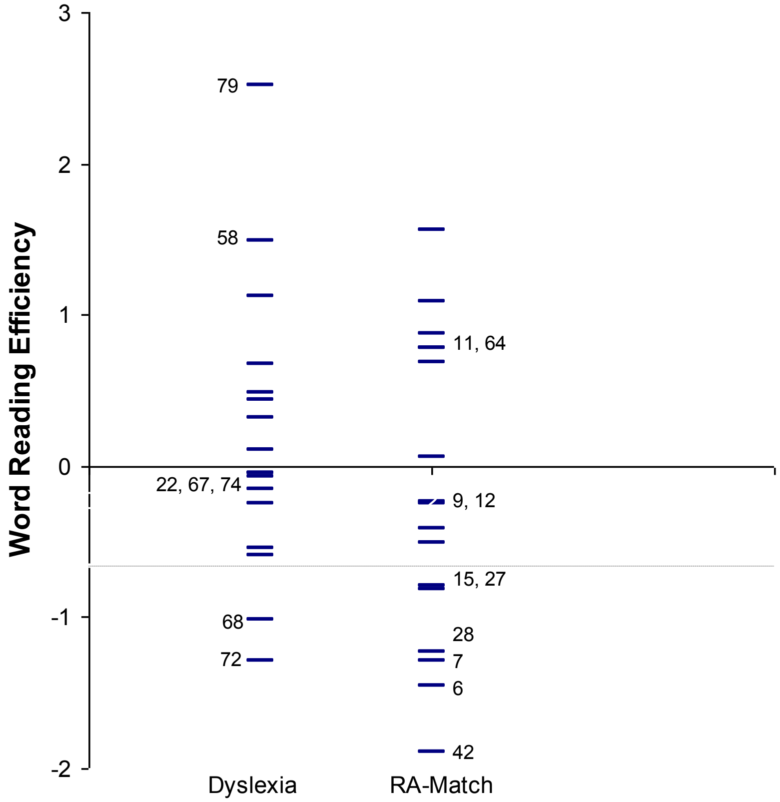
Figure 32 displays individual differences on phonological awareness after the effect of attention was controlled statistically.



3 < -.67 vs. 2 < -.67

Figure 32. Graph of individual performance for phonological awareness with ADHD index partialled out.

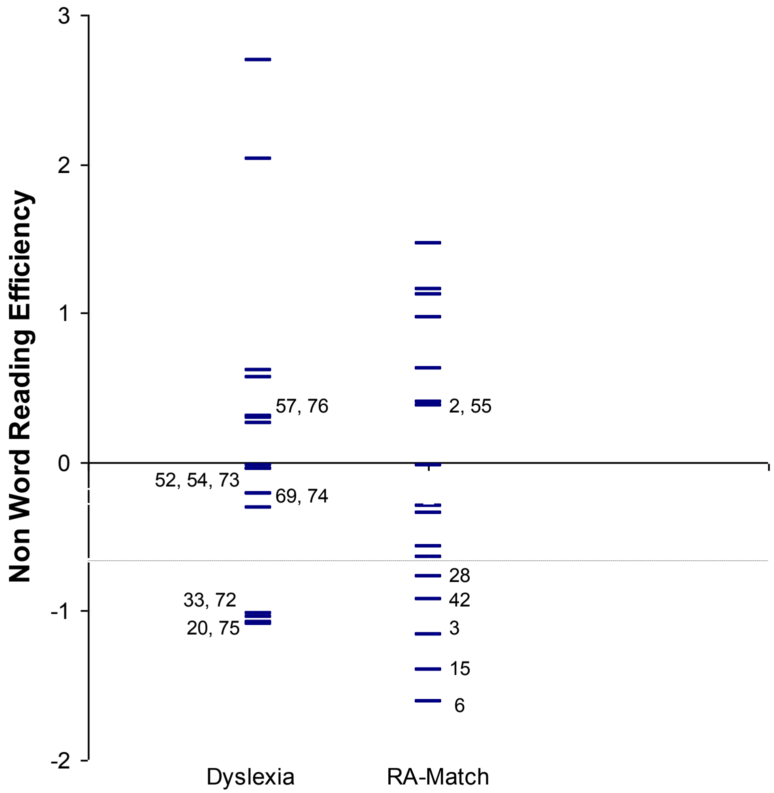
Phonological awareness. As shown in Figure 32, attention seems to explain the performance of some cases among the dyslexia subgroup. All three cases with Elision scores of more than 1 SD below the mean (previously identified in Figure 22) seemed to have improved upon adjustment for attention, as can be seen in Figure 32. Among participants in the RA-match control subgroup, the two cases (i.e., 3, 6) which were identified as extreme poor cases prior to adjustment for attention have remained below the .67 cut-off point. Nonetheless, even after controlling for attention, the scores' dispersion within the two subgroups remains somewhat similar to Figure 22. That is, while scores among participants in the dyslexia subgroup cluster somewhat more between the mean and the .67 cut-off point, those in the RA-match subgroup are more dispersed. Figures 33 and 34 on the next page display individual differences on word and non-word reading fluency after the effect of attention was controlled statistically.



2 < -.67 vs. 6 < -.67

33

Figure 33. Graphs of individual performance for word reading efficiency with ADHD index partialled out.



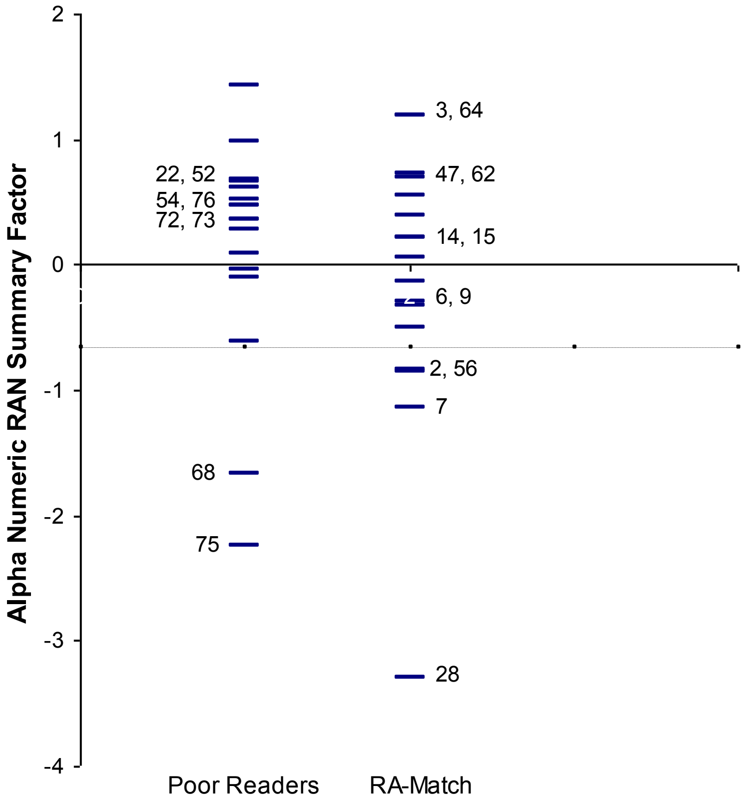
4 < -.67 vs. 5 < -.67

34

Figure 34. Graphs of individual performance for non-word reading efficiency with ADHD index partialled out.

Word and non-word reading efficiency. As shown in Figure 33 and 34, attention seems to explain the performance of some participants in the dyslexia subgroup for both word and non-word reading efficiency. For word reading efficiency in Figure 33, two participants who were identified as extreme poor cases prior to controlling for attention are no longer outliers upon adjustment for attention. However, the number of extreme poor cases in the RA-match subgroup remains unchanged. Consequently, upon adjustment for attention, more participants in the RA-match as compared to the dyslexia subgroup are identified as extreme poor cases (i.e., 6 in RA-match vs. 2 in dyslexia). Similarly, for non-word reading efficiency in Figure 34, it seems that the performance of 3 participants (i.e., 20, 72, and 75) in the dyslexia subgroup has improved upon controlling for the effect of attention. Nonetheless, all three cases still remain below the .67 cut-off line.

Figure 35 on the next page represents individual differences on alpha-numeric rapid naming measures after the effect of attention was controlled statistically.

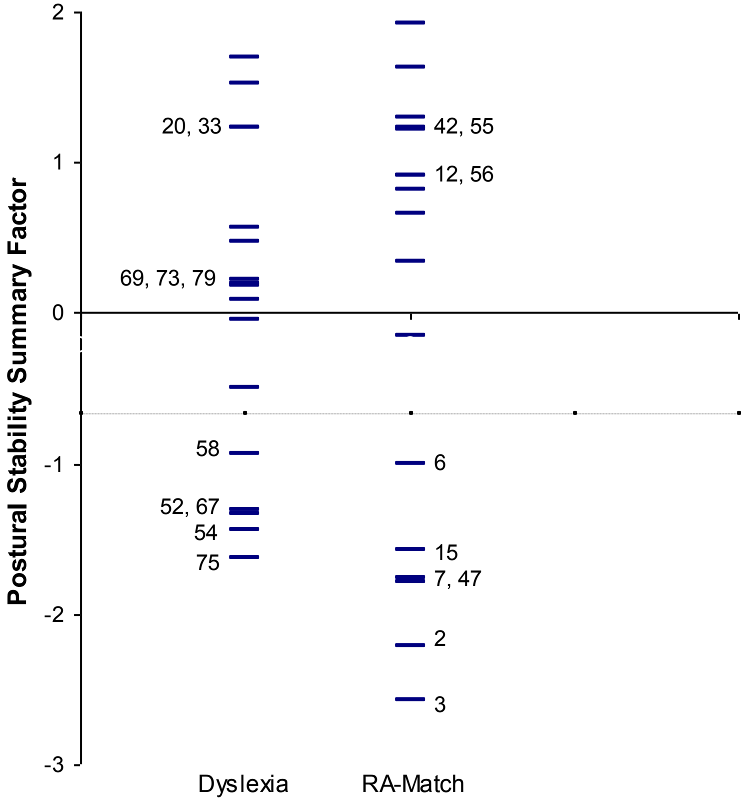


3 < -.67 vs. 4 < -.67

Figure 35. Graph of individual performance for alphanumeric rapid naming summary factor with ADHD index partialled out.

Alpha numeric rapid naming summary factor. Figure 35 illustrates individual performances on the alpha numeric rapid naming measures after statistical adjustment for attention. The pattern of performance indicated that even after adjustment for attention, alpha numeric rapid naming tasks did not seem to be able to distinguish between the dyslexia and RA-match subgroups at an individual level. Similar to Figure 25 shown previously, a similar proportion of cases fell below and above the cut-off point (i.e., 3 in dyslexia vs. 4 in RA-match are below .67 cut-off point). Except for one new case (i.e., case number 56) in the RA-match control subgroup, the other extreme poor cases among participants in the dyslexia and RA-match subgroups were the same as those that were also extreme poor cases prior to the statistical adjustment of attention. Overall, attention did not seem to explain performance on the alphanumeric rapid naming tasks.

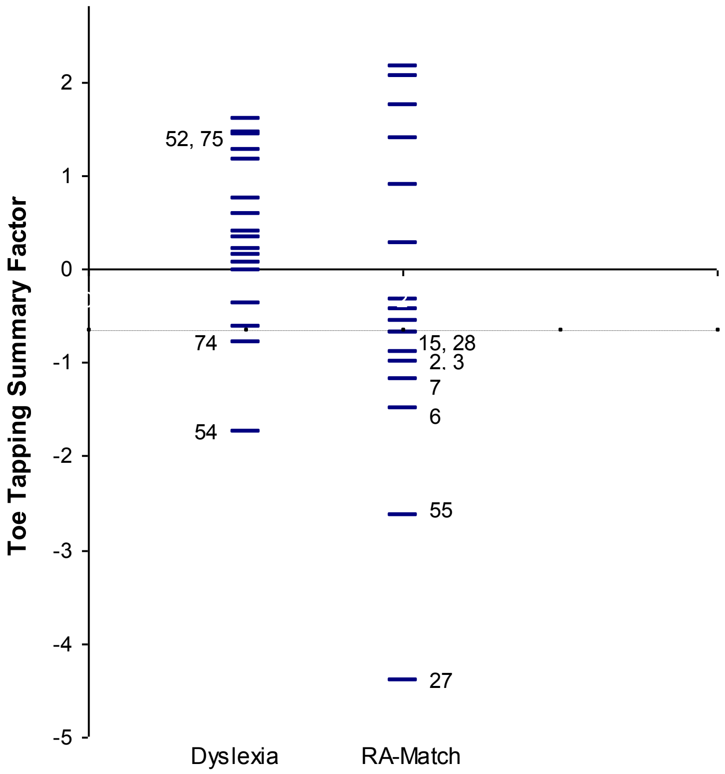
Figures 36 and 37 on the next page display individual differences on postural stability and toe tapping summary factors after the effect of attention was controlled statistically.



5 < -.67 vs. 6 < -.67

36

Figure 36. Graphs of individual performance for postural stability summary factor with ADHD index partialled out.



2 < -.67 vs. 8 < -.67

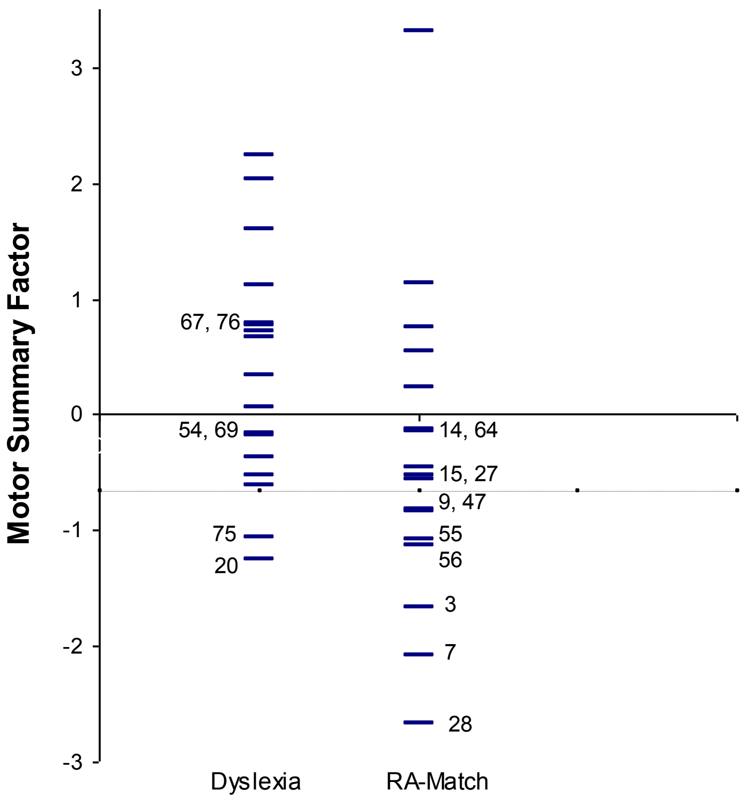
37

Figure 37. Graphs of individual performance for toe tapping summary factor with ADHD index partialled out.

Postural stability summary factor. As seen in Figure 36, the pattern of individual performances on postural stability tasks remained similar to when the effect of attention was not removed. There was also no change in the proportion of participants falling below or above the cut-off line in both subgroups. This means that attention did not explain performance on these tasks.

Toe tapping summary factor. Figure 37 illustrates individual performance on the toe tapping summary factor after the effect of attention was removed statistically. Similar to the postural stability summary factor, the pattern of individual performance on the toe tapping summary factor did not seem to change after the effect of attention was removed. The trend observed prior to statistical adjustment for attention also did not change. That is, there were still fewer extreme poor cases among participants in the dyslexia subgroup as opposed to those in the RA-match control (i.e., 2 vs. 7). Additionally, except for a few individual extreme poor cases whose performance seemed to have somewhat improved after statistical adjustment of attention, altogether attention did not seem to explain performance on toe tapping summary factor.

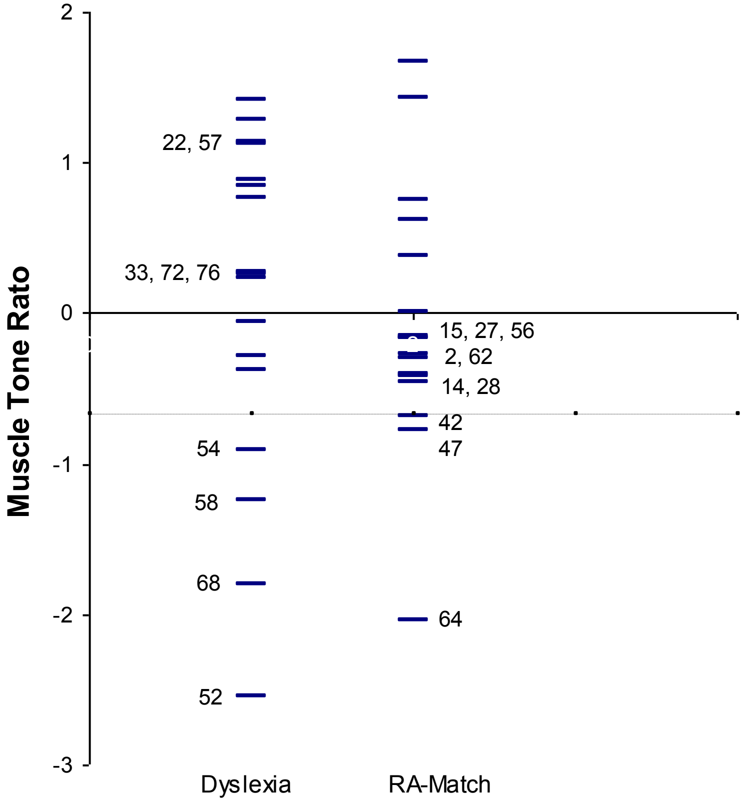
On the next page, Figures 38 and 39 display individual differences on motor summary factor and muscle tone ratio after the effect of attention was controlled statistically.



2 < -.67 vs. 7 < -.67

38

Figure 38. Graph of individual performance for motor summary factor with ADHD index partialled out.



4 < -.67 vs. 3 < -.67

39

Figure 39. Graph of individual performance for muscle tone ratio with ADHD index partialled out.

Motor summary factor. As seen in Figure 38, even after statistically controlling for attention, the trend of individual performances remained the same on the motor summary factor and this summary factor does not seem to distinguish among the two subgroups. Similar to toe tapping summary factor, the trend observed in the motor summary factor prior to adjustment of attention seems to persist (i.e., less extreme poor cases among participants in the dyslexia subgroup as compared to those in the RA-match control).

Muscle tone. Figure 39 illustrates individual performances on the final cerebellar measure after statistical control for the effect of attention. The pattern of individual performances remained similar to that observed prior adjustment of attention. Hence, attention did not seem to explain performance on the muscle tone task.

Summary

Altogether, prior to statistically removing the effect of attention, the individual analysis on all literacy measures revealed five extreme poor cases among the dyslexia subgroup on the reading factor (i.e., word reading task) as compared to three in the RA-match control. Among the five cases in the dyslexia subgroup, four were also identified as extreme poor cases on other literacy measures. However, only one of these five cases was identified as an extreme poor case on the rapid naming summary factor and only two of the five cases were identified as extreme poor cases on motor and cerebellar measures. More specifically, one of these two cases was identified as an extreme poor case on postural stability and the other on the motor summary factor.

In sum, both prior to and after controlling for the effect of attention, among the literacy measures, phonological awareness and polysyllable non-words as compared to non-word decoding that represented performance on all items of the Word Attack seemed to be only somewhat more successful in differentiating between the dyslexia and the RA-match subgroups at an individual level. Generally, the pattern of scores' dispersion for phonological awareness seemed to suggest that scores within the dyslexia subgroup clustered somewhat more between the mean and the .67 cut-off line while those in the RA-match control were more dispersed. Additionally, one of the extreme cases on phonological awareness in the RA-match control was also an outlier on the word reading task. Ignoring this case would also result in an increase in the proportion of extreme poor cases in the dyslexia subgroup as compared to the RA-match control both prior and after controlling for the effect of attention (i.e., 3 in dyslexia vs. 1 in RA-match). For both phonological awareness and polysyllable non-words, fewer participants in the dyslexia subgroup as compared to the RA-match control had scores falling above the mean of zero. In the case of the non-word decoding measure that represented all items of the Word Attack (i.e., including mono- and polysyllables), the differentiation between the two subgroups seemed to be poor at an individual level. However, more participants in the RA-match control as compared to the dyslexia subgroup seemed to fall above the mean of zero (i.e., 11 in RA-match vs. 5 in dyslexia). The other literacy measures (i.e., word and non-word reading efficiency and alpha-numeric rapid naming measures) also did not seem to differentiate well between the two subgroups at an individual level both prior to and after controlling for attention.

Among the motor and cerebellar measures, muscle tone seemed to illustrate a trend at first glance in which more extreme poor cases were identified in the dyslexia subgroup as compared to the RA-match control both prior to and after adjustment for attention (i.e., 4 in dyslexia v. 2 in RA-match). Nonetheless, none of the extreme poor cases in the dyslexia subgroup seemed to have difficulties in any of the literacy measures except for one case that was identified as an extreme poor case on alpha-numeric rapid naming, as well as word and non-word reading efficiency measures. Furthermore, at first glance the postural stability measure did not seem to distinguish well between the two subgroups at an individual level both prior to and after adjustment for attention. However, this pattern changed moderately when the three extreme poor cases (i.e., 6, 7, and 15), which were outliers on word reading in the RA-match subgroup were ignored. In that case, more participants in the dyslexia subgroup as compared to the RA-match control were identified as extreme poor cases on the postural stability summary factor both prior to and after adjustment for attention (i.e., 5 in dyslexia vs. 3 in RA-match). Further exploration of the five extreme poor cases on postural stability in the dyslexia subgroup indicated that only two were also identified as extreme poor cases on other literacy measures. Only one of these cases was among the extreme poor cases originally identified in the word reading task.

CHAPTER 5

Discussion

This thesis addressed four questions. The results for each of these questions are discussed in this chapter. Prior to addressing the implications of main questions of this study, a preliminary step was taken to reduce the motor and cerebellar measures into clusters. The results related to this preliminary step are therefore reviewed before addressing the implications of the main findings of the thesis.

Preliminary Analyses of the Component Structure of Motor and Cerebellar Tasks

The motor and cerebellar measures used in this study included peg moving, bead threading, two postural stability tasks, muscle tone, and two toe tapping tasks. Altogether, three Principal Component analyses were performed to investigate how these measures loaded together. This led to extraction of three clear components which together explained a total of 80.52% of the variance. The three components were (a) a motor component comprised of bead threading and peg moving, (b) a postural stability component, and (c) a toe tapping component. In addition, all analyses confirmed that muscle tone did not load strongly together with any of the other tasks including postural stability. Consequently, muscle tone was retained for subsequent analyses as a distinct variable outside of the three components identified from data reduction.

This pattern of loadings for variables seems to provide some empirical confirmation of Fawcett and Nicolson's (1995c) earlier assumption that bead threading and peg moving tasks load together closely and are somewhat separate from postural stability. Peg moving has been widely considered to be a fine motor task and used to assess motor impairment in eye-hand coordination and motor speed as well as uni-manual and bimanual finger and hand dexterity (Annett et al., 1974; Bishop, 1990; Strauss et al., 2006). As previously discussed, the two other components extracted from the exploratory Principal Component analysis were postural stability and toe tapping. Of the three tasks including toe tapping, postural stability, and muscle tone, the latter two, namely postural stability and muscle tone, have been considered by some researchers to be “most clearly cerebellar-based” (Fawcett & Nicolson, 1999, p. 70). Arguably then, the first result here suggests that the assumption that postural stability and muscle tone are closely associated is not borne out in the principal components analyses conducted in this study. Indeed, they are not even correlated in this sample (see Table 11 on pp. 178-179).

More generally, perhaps the most important aspect of these results is to suggest that, in behavioral terms, the battery of motor tasks used here are not convergent but instead are divergent in terms of their underlying statistical structure. As all data reduction analyses undertaken identified at least two components and an independent muscle tone variable. It is quite possible that the measures reflect distinct underlying behavioral constructs. This impression is also supported by the finding that the three latent variables derived from the data reduction analysis were uncorrelated. This diversity has not been demonstrated in previous work in this area, as data reduction has not been applied to these variables to date. This analysis confirms the importance of establishing the structure of motor task variables before considering their role in reading. In this way the fullest possible picture of motor-literacy skill associations can now be evaluated by exploring the role of these latent variables in reading acquisition.

Although speculative at this stage, it is also possible that the distinct loadings for the range of motor measures used here may also be an indication that these three tasks are modulated by different cerebellar zones or areas, especially considering the complex structure of the cerebellum. The cerebellum is considered to be one of the three most important areas that contribute to coordination and movement (Shumway-Cook & Woollacott, 2001). Cerebellar function is partly determined by its neuronal circuitry as well as its input and output connections (Shumway-Cook & Woollacott, 2001). A detailed discussion of the anatomy of the cerebellum, its different zones, and functions that are regulated by these zones is beyond the scope of this thesis. In simplified terms, however, the cerebellum can be divided by function into three zones, as illustrated in Figure 40. The oldest zone, in terms of evolutionary development, is functionally related to the vestibular system and the areas more recently developed are the vermis (in the medial zone) and the intermediate part of the hemispheres, and finally, the lateral hemispheres (Dobkin, 2003; Shumway-Cook & Woollacott, 2001).

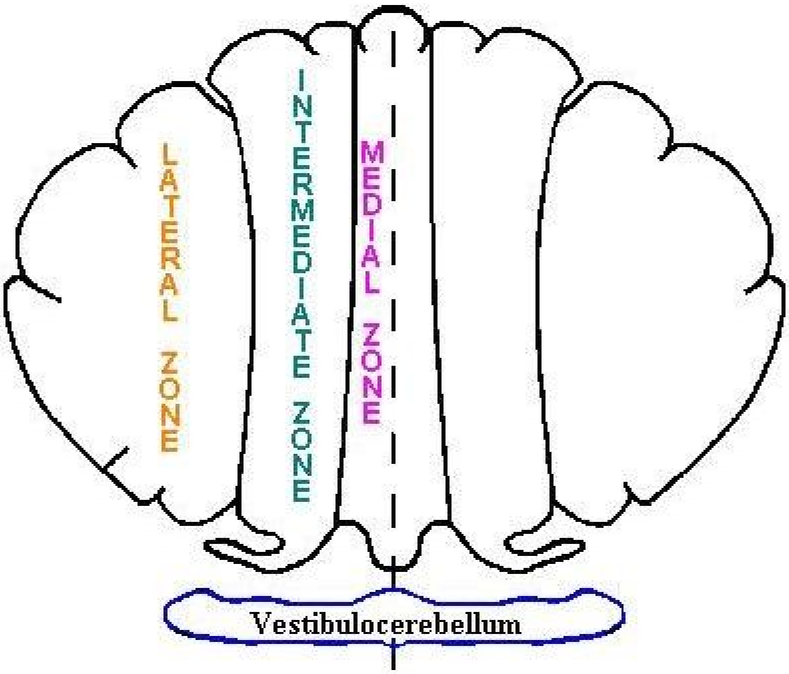


Figure 40. The basic anatomy of the cerebellum.

It is well established that the vestibular system is the area that provides sensory information regarding head movement and position with reference to gravity. This system regulates stabilization of gaze, posture, and balance and it contributes the conscious sense of orientation. Consequently, pathology within the vestibular system produces problems related to gaze, stabilization, posture, and balance, as well as vertigo or dizziness (Dobkin, 2003; Shumway-Cook & Woollacott, 2001). On the other hand, the vermis and intermediate areas appear to regulate muscle tone. For instance, lesions in these areas are associated with a drop in muscle tone or hypotonia (Shumway-Cook & Woollacott, 2001). Finally, the lateral areas are suggested to be involved in higher level functions involving both motor and non-motor skills. For example, there is evidence suggesting that lateral areas are involved in the preparation of movements and coordination of ongoing movements (i.e., programming the execution of movement) (Shumway-Cook & Woollacott, 2001). Furthermore, it is suggested that the speed, rhythm, and smoothness of repetitive movements, such as hand patting or toe tapping are more affected by the dysfunction of the lateral zones of the cerebellum (Granacher, 2003; Walker, 1990). Considering this simplified explanation of the different areas in the cerebellum and their functions, the separate loadings of the measures toe tapping, postural stability, and muscle tone may not be surprising. However, in light of the complex structure of the cerebellum, no definite conclusions can be drawn in this regard.

Any model of the neurological underpinning of these patterns is highly speculative at this stage and naturally the real picture may be much more complicated than illustrated here. Furthermore, there may be other explanations of the same patterns of statistical effects reported here. As Fawcett and Nicolson (1999) have pointed out, no behavioral measure is a pure measure of brain functioning. The involvement of other brain systems in any experimental task is to be expected. Consequently, although the three tasks of toe tapping, muscle tone, and postural stability are considered to be sensitive to the cerebellar dysfunction, they may not be necessarily *specific* to the cerebellum (Fawcett & Nicolson, 1999). This fact in itself may explain why the three measures, or at least muscle tone and postural stability, did not load together in the analysis. Finally, there may be entirely non-neurological explanations of the pattern of results reported. In particular, individual tests such as those of postural stability were marked by low reliability and this artifact may have contributed to or even explained the distinct patterns reported above. Some of the preliminary analyses conducted in the present study produced a 2-factor rather than a 3-factor solution (with muscle tone as a distinct additional factor in each case), a finding less tidily associated with the tripartite division of cerebellar function described above. Although the 3-factor models were clearly a better fit than the 2-factor models, there is an element of researcher judgment that comes into play in making decisions about how to best interpret results of data reduction techniques.

Question 1:

Is there a relationship between word reading as measured by word identification and (a) phonological awareness, (b) reading fluency and rapid automatized naming, and (c) purported cerebellar processing tasks?

According to the proponents of the cerebellar deficit theory, a mild congenital deficit in the cerebellum is proposed to give rise to a series of impairments (e.g., deficits in central processing speed and automatization of elementary auditory and articulatory skills) that eventually lead to deficits in rapid naming, phonological processing and finally reading (Fawcett et al., 2001; Nicolson & Fawcett, 1999; Nicolson et al., 2001). Based on this assumption, correlations between motor-cerebellar measures and measures related to reading, rapid naming, and phonological processing were expected.

The results of this study did not support this prediction. To address these findings, results for phonological, word reading efficiency, and rapid naming measures are briefly discussed first and then followed by results for motor and cerebellar tasks and their implications considered.

Relationship between reading, phonological and rapid naming measures. Similar to other repeatedly reported findings in the literature, (e.g., Berninger et al., 1987; Katzir et al., 2006; Share, 1995; Siegel, 1993; Snowling, 1995; Swanson, Trainin, Necoechea, & Hammill, 2003; White et al., 2006), findings from the present investigation indicated significant relationships between word reading and measures of non-word reading, word reading efficiency and phonological awareness. These results show that the preliminary findings of this research are quite typical of many of those in the literature in confirming that among a range of variables, measures related to phonological awareness seem to have the strongest association with reading ability.

Moderately strong relationships were also found between word reading and all three measures of rapid naming indicating that in the present sample those with strong word reading skills also seemed to be more efficient in all three rapid naming measures and vice versa. A significant relationship between rapid naming measures and word reading has also been reported by many other researchers (e.g., Ackerman, Dykman, & Gardner, 1990; Bowers, Steffy, & Swanson, 1986; Swanson et al., 2003). Consistent with the results of this study, some other reported findings have also shown that relative to the relationship between word reading and phonological awareness, the link between word reading and rapid naming measures may be more modest in size (e.g., Cardoso-Martins & Pennington, 2004; Savage & Frederickson, 2006). Furthermore, results of this study indicated that all three rapid naming measures were related to each other. Significant but modest relationships were also found between rapid digit naming and phonological awareness and stronger associations between all rapid naming measures and measures of non-word reading, word reading efficiency, and phonological awareness. Generally, these findings were consistent with other evidence (e.g., Katzir et al., 2006; Wolf et al., 2002).

Relationship between word reading and motor/cerebellar measures. Contrary to the predictions made by the cerebellar deficit hypothesis, no significant relationships were revealed in the present sample among any of the reading related measures and latent measures of postural stability, toe tapping, and muscle tone. Interestingly, and consistent with the notion of developmental maturity, the significant but modest relationship found between measures of chronological age and the toe tapping component and the significant strong relationship revealed between chronological age and the motor component in this study indicate better performances on both of these components with increasing age. Additionally, significant relationships that were found between the motor component and three reading related measures (i.e. single word reading, and word and non-word reading efficiency) were moderated once chronological age was partialled out. In fact, the motor component was no longer significantly correlated with single word reading and word reading fluency.

In contrast to the results of this study, there have been some studies that have found a relationship between measures of reading and motor/cerebellar variables. For example, White et al. (2006) found a moderate relationship between the motor variable in their study (which included postural stability measures, bead threading as well as finger to thumb tasks) and both literacy and phonology variables in their sample of readers with dyslexia and age-match controls. Nonetheless, as White et al. (2006) explained, these correlations seemed to largely reflect the influence of two children with dyslexia in their sample that were considered extreme poor cases (i.e., outliers in White et al.'s terms) on literacy, phonology, and motor domains. Once these two apparent outliers were removed from their sample, the correlations no longer persisted.

Savage et al. (2005b) have also found a significant but modest relationship between postural stability and non-word reading in their study, mentioned earlier, where relationships between rapid naming, phonological awareness, nonsense word reading, rapid perception and motor balance automaticity were examined among average, below average and above average readers. Their findings also indicated that postural stability loaded modestly on the same component as phonological processing and rapid naming. Nonetheless, postural stability was not found to be a significant predictor of word reading, nor did this measure reliably distinguish between the groups in their sample. The results from the majority of research studies in a systematic review conducted by Hammill (2004) of some 450 concurrent and longitudinal studies over the past 30 years suggested that motor measures were generally very poor predictors of reading ability. Findings of this study are consistent with this general pattern. Overall, the lack of evidence for a significant relationship between measures of literacy including phonological processing and motor/cerebellar measures, especially when accompanied by evidence of strong relationships between measures of literacy and phonological processing seem to provide little support for the possible link between cerebellar and reading processes or of a link between cerebellar processing and reading processes via phonological awareness. The present study adds to knowledge here by showing that even with using robust latent variables of motor abilities rather than individual motor variables to measure a range of motor skills and despite precise and calibrated measurements of postural stability and muscle tone, none are significant associates of variability in literacy in school-based samples of typical readers.

Interestingly, findings of this research also did not reveal a significant relationship between attention and any of the motor and cerebellar measures. This was surprising considering some of the findings that seem to suggest an association (e.g., Denckla et al., 1985; Raberger & Wimmer, 1999; Raberger & Wimmer, 2003; Ramus, Pidgeon et al., 2003) as well as the evidence indicating that attention may be one of the potential variables to modulate the effect size in differences found on cerebellar tasks between average and poor readers (Rochelle & Talcott, 2006). A possible reason for these findings may be that there were few participants in the present sample with clinically significant attention problems. Alternatively it may be that attention-motor associations reported frequently in clinical samples are just not evident at a group level in school-based samples of typical readers.

Question 2 and 3:

The second question investigated whether a subgroup of children with dyslexia selected from the sample differed in their performance on any of the motor, cerebellar, reading, phonological, and rapid naming related measures when compared to two control subgroups that were selected from the same sample and matched to the dyslexia subgroup based on (a) their reading level, and (b) chronological age.

The third question investigated whether any group differences in performance on the above reading and cerebellar related measures emerged when the effect of attention is controlled statistically.

The cerebellar deficit hypothesis predicts that children with dyslexia should display deficits in motor- and cerebellar-related tasks in relation to a RA-match and a CA-match control group. Moreover, group differences in the motor- and cerebellar measures should survive the covariance of attention because according to the proponents of the cerebellar deficit hypothesis (Fawcett et al., 1996; Fawcett et al., 2001) cerebellar impairment in children with dyslexia is independent from the presence of attention difficulties. None of these predictions were supported by the present results. To address these results, again findings pertaining to phonological, word reading efficiency and rapid naming measures are discussed first in this section and they are then followed by discussion of the results related to motor and cerebellar measures.

Non-word reading and phonological awareness. Findings of the present study revealed that as compared to participants in both the RA- and CA-match subgroups, those in the dyslexia subgroup performed significantly poorer in phoneme deletion and polysyllable non-word decoding consistent with a phonological deficit account of dyslexia. When all non-word items (including both mono- and polysyllables) were considered in the analysis, the dyslexia subgroup's performance did not differ significantly from the RA-match control subgroup. Group differences were also explored using Cohen's *d* effect size ratios, especially considering that noteworthy effects with clinical and practical significance can be obtained despite non-significant results (Durlak, 2009; Thompson, 2002). Indeed, for both phoneme deletion and non-word reading (all items), the effect sizes found for contrasts between the dyslexia subgroup and the RA-match control were moderate, and they were accompanied by large effect sizes for the contrasts between the dyslexia subgroup and the RA-match control with the CA-match control. Of note was the large effect size that was found for polysyllable non-words for the contrast between the dyslexia subgroup and RA-match control.

The modest (.59) and large (.74) effect sizes for non-word reading obtained in this study are comparable to the overall combined effect size of .65 reported in the systemic review for non-word reading deficit by Herman et al. (2006). Van IJzendoorn and Bus (1994) have reported a somewhat smaller overall combined effect size of .48 in their systematic review for the non-word reading deficit. There are several factors that seem to be relevant to the interpretation of the effect sizes obtained in this study. One issue is the reading variable upon which groups have been matched. According to Herrman et al. (2006), effect sizes reported for non-word reading in reading-level designs vary systematically depending on the reading variable used for group matching procedures. Herrman et al. (2006) have reported smaller effect sizes in studies that have used passage reading tests to match groups compared to those studies that have used word-level reading accuracy to match groups. The pattern of effect sizes found for non-word reading in the present study, which used word-reading to match subgroups and reported both a moderately large and a large effect size (i.e., .59, .74) for non-word reading deficit is, in this sense, consistent with Herrman et al.'s (2006) report.

Van IJzendoorn and Bus (1994) have also reported that studies with a better match on verbal intelligence have yielded larger effect sizes because of the fact that the results are less affected by a general language deficit. Although the subgroups in this study were matched on verbal intelligence, the effect size obtained for non-word reading (all items) was in the moderate range. This could be due to the choice of the verbal test used for matching the subgroups. As van IJzendoorn has noted, purely verbal intelligence tests such as the Peabody Picture Vocabulary Test (PPVT) (Dunn & Dunn, 2007) have shown to result in larger effect sizes (ranging from .76 to 1.03) compared to mixed verbal/performance tests such as those used in this study.

According to Rack et al. (1992), the age of normal readers in the RA-match control group may also affect the magnitude of the effect size obtained for non-word reading deficit. As Rack et al. (1992) note, it may be too early to expose young 7-year-old readers to non-word decoding tests since these readers may experience developmentally normal difficulty with reading non-words. This in turn can result in obtaining a smaller effect size for non-word reading deficit. In the present study, the magnitude of effect sizes for non-word reading (including all items and polysyllables only) did not seem to be affected by this factor, even though participants in the RA-match control subgroup were young. Inspections for univariate normality for variables within each subgroup indicated that non-word reading variables were normally distributed within the RA-match control subgroup. Although the age of participants in the RA-match control subgroup may not have been a factor in the present study, the difference in mean ages of participants in the RA-match control and those in the dyslexia subgroup may have played a role in the magnitude of the effect size obtained for non-word reading. In their systematic review, Herrmann et al. (2006) found some support for an association between larger group differences in age with smaller non-word reading deficits. According to Herrmann et al. (2006), this may be explained by the fact that older children with reading difficulties may have skills or strategies that could mask group differences on non-word reading tests. For example, they may have more “educational experience” or “cognitive and perceptual maturity” that could result in a smaller effect size for non-word reading tests (Herrmann et al., 2006).

Hermann et al.’s (2006) explanation may apply to less complex non-words, such as monosyllables, as opposed to polysyllable non-words. As Rack et al. (1992) have pointed out in their systematic review of non-word deficit, performance on more complex non-words may be more likely in identifying a deficit. Indeed, in this study the large effect size of .74 found for polysyllable non-words as opposed to the moderately large effect size of .59 found for all non-words, which included both mono- and polysyllables, seemed to support this notion. Nonetheless, some caution is advised since, as was indicated in the Method chapter, only 11 of the 45 items in the Word Attack test are polysyllables and they start from item 19. This means that not all participants had a chance to read every polysyllable non-word depending on when the ceiling was reached. A closer observation of performance on polysyllables indicated that within the dyslexia subgroup, 14 of 17 participants had a score of zero for polysyllable non-words. Seven of these 14 participants did not have the chance to read any of the polysyllable non-words because they reached ceiling prior to item 19. Nonetheless, it was performance on the polysyllable non-words that differentiated between the dyslexia subgroup and the RA-match control, while performance on all items which also included monosyllable non-words did not.

Overall, findings for non-word reading and phoneme deletion in the present research are similar to many other studies that have also used reading-level designs (e.g., Ackerman & Dykman, 1993, Bowey et al., 1992; Bruck, 1992; Bowey & Hansen, 1994, 2005a; Gillon & Dodd, 1994; Savage et al., 2005a). Furthermore, a possible causal interpretation of the positive findings for phonological deficits in the dyslexia subgroup in the present sample is strengthened by some of the more careful approaches followed in sampling participants for this study, which seems to be lacking in some reading-level designs, as discussed in the Chapter 2. For example, as Jackson and Butterfield (1989) have advised, the sample of poor and typical readers in this project was recruited in a comparable fashion from the same public school system. The preliminary analysis also indicated that the three subgroups were matched on their overall IQ level. They were also matched on the proportions of their first language spoken. Most importantly, as was indicated earlier, the sample seemed to generally represent English-dominant bilinguals who were concurrently exposed to English and another language. Thus language did not seem to explain differences observed in reading processes. These controls for potential extraneous variables strengthen findings of this study. Generally the pattern of findings for non-word reading, especially polysyllable non-words, and phoneme deletion as explored via effect size ratios seemed to provide support for the presence of decoding and phonological deficits in participants in the dyslexia subgroup in the sample when compared to both same-aged and younger RA-match subgroups.

The findings for non-word reading also persisted even after the effect of attention was controlled. However, it was interesting to find that for phonological awareness, the pattern of results changed dramatically after the effect of attention was controlled. That is, participants in the dyslexia and RA-match subgroups no longer differed in their phonological awareness skills once the results were adjusted for attention. This was also confirmed when the effect sizes were explored using Cohen's *d*. This seemed to be because attention partly explained the poor performance on phonological awareness task (i.e., Elision) for some participants in the dyslexia subgroup. In fact, Palacios and Semrud-Clikeman (2005) have also reported a negative relationship between phonological awareness and ADHD scores. Similarly, in this study there seemed to be a trend in the dyslexia subgroup indicating somewhat lower scores on Elision tasks in children with higher ADHD scores.

Nevertheless, no conclusions can be drawn considering the small number of participants with higher ADHD scores in this study. Evidently, it is also not really clear whether the cases with higher ADHD index scores in the dyslexia subgroup in the present research sample truly represent coexisting attention deficit hyperactivity disorder since a diagnosis cannot rely merely on behavior ratings. Although some participants in the dyslexia subgroup may truly represent coexisting attentional problems, it is also possible that some of the children with reading difficulties show ADHD-like symptoms in reaction to their reading problems (Pennington, Groisser, & Welsh, 1993).

The particular question, however, is why attention problems in this sample specifically explained worse performance on the phoneme deletion (i.e., Elision) task and not on the non-word reading task (i.e., Word Attack). One interpretation may be related to the nature of the task, since between Elision and non-word decoding tasks, the former may be more complex. As opposed to non-word reading, the Elision task is auditory, requiring the child to hear, encode and hold verbal information (e.g., “cat”) in auditory working memory, and then separate and manipulate phonemic units and articulate a new word (e.g., take away “c” and say “at”) (Plaza, 2003). Therefore, it may be more likely for attentional problems to interfere at some level in performance on the Elision task to a greater extent than they would in performance on the non-word decoding task. One plausible idea is that the Elision task involves more auditory working memory resources than does reading of non-words. Some research findings suggest that children who exhibit coexisting reading and attentional difficulties seem to have more difficulties in working memory (Bental & Tirosh, 2007; Savage, Lavers, & Pillay, 2007). Working memory involvement may thus be a possible explanation for findings related to the Elision task in this study.

Regardless of the specific processes in the Elision task, which may be negatively affected by attentional problems, a phonological deficit as opposed to attentional difficulties seems to remain the primary potential explanation for the difficulties experienced in this task in participants in the dyslexia subgroup in this sample. Indeed, while the mean performance for this subgroup increased to some extent (i.e., from 6.82 to 7.28) after attention was controlled, it still remained below average.

Rapid naming skills. In contrast to the results for phonological processes, the main findings for rapid naming measures indicated that participants in the dyslexia subgroup did not differ significantly from those in the RA- and CA-matched controls in their performance on any of the rapid naming tasks. Although the main analyses for rapid digit naming were negative as they did not reach an overall significance, the pattern of effect size ratios (in terms of Cohen's *d*) indicated that there was some evidence for group differences in rapid digit naming. More specifically, both the dyslexia subgroup and the younger normal readers performed more poorly than the older normal readers (effect size ratios of .64 for both contrasts).

Overall, while the main findings related to performance on rapid naming tasks in the present study were not consistent with some of the related evidence in literature that has found differences in speeded naming between normal and poor readers (Bowers et al., 1988; Denckla & Rudel, 1976a, 1976b; Meyer, Wood, Hart, & Felton, 1998; Semrud-Clikeman, Guy, & Griffin, 2000), some of these findings may be hard to interpret, especially because of the lack of a reading-level design. As Wolf and Bowers (1999) have indicated, findings related to naming speed differences have been mixed in studies that have used reading-level design. Some reading-level design studies have not found any group differences (e.g., Badian, 1996) and others have found slower performance in naming symbols in readers with dyslexia as compared to an RA-match control subgroup (e.g., Ackerman & Dykman, 1993; Sunseth & Bowers, 2002). There has been speculation that a speeded naming deficit might be specific to IQ-achievement discrepant readers with dyslexia, whereas non-discrepant “garden-variety” poor readers have demonstrated normal performance in speeded naming (e.g., Ackerman & Dykman, 1993; Wolf, 1991; Wolf & Bowers, 1999). However, other researchers have failed to find support for this view (e.g., Hammill, 2004; Metz et al., 2003; Savage, 2007).

The pattern of the present results for rapid digit naming in terms of effect size ratios was in part somewhat similar to the pattern of effect size ratios that was obtained for this measure in a study by Savage et al. (2005a) that also included a reading- and a chronologically-age match control group. In that study, Savage et al. (2005a) did find a significant main effect for rapid digit naming in their sample which was due to poorer performance of poor readers as compared to the CA-match, but not to RA-match control group. However, in comparisons that were made in terms of effect size ratios, large effect sizes were obtained for both comparisons (i.e., 1.85 for poor readers vs. CA-match, and .85 for poor readers vs. RA-match). Although no effect size ratios were reported for contrast between RA- and CA-match controls in Savage et al. (2005a), this effect size would also have been large if calculated (i.e., .95) indicating that consistent with the pattern of findings for this study, the RA-match also seemed to perform poorer than the CA-match group.

For this study, the pattern of effect size ratios for rapid digit naming tasks remained consistent for the contrast of young and same aged normal readers after adjustments were made for attention. However, the effect size decreased from moderate to small (i.e., from .64 to .25) indicating equal performances in participants in the dyslexia subgroup and those in the CA-match control after attention effects were controlled statistically. Therefore, attention seemed to partly explain poorer performance in digit naming for some participants in the dyslexia subgroup. Although equal group performances were found for other rapid naming measures in terms of effect size ratios, the pattern of effect sizes after adjustment for attention were similar to that found for the rapid digit naming task for the contrast of the dyslexia and CA-match subgroups. If the dyslexia subgroup in this study represents some individuals with comorbid attention and reading deficits, then it is possible that slower rapid naming performance may be unique to comorbid cases. This view is supported by some findings in the literature that have indicated slower performance on rapid naming measures in comorbid cases as compared to those with merely reading disability or attentional difficulties (e.g., Bental & Tirosh, 2007; Chan, Hung, Liu, & Lee, 2008; Rucklidge & Tannock, 2002).

Overall, the main findings for rapid letter and digit naming in this study indicated that the underlying processes related to these speed naming tasks may not play a role in causing reading difficulties in the dyslexia subgroup in the sample. While findings for digit naming in this study and in Savage et al.’s (2005a) study may provide some evidence for possible difficulties in speed naming tasks in the dyslexia subgroup, the mixed patterns found in terms of main analysis as well as the pattern and size of effect size ratios, which indicated equal performance in participants in the dyslexia subgroup and those in the RA-match control, seem to suggest a weak effect in relation to findings for phonological processes, possibly an indication of a delay rather than a deficit. Considering some of the research findings on rapid naming tasks in comorbid cases, it is also possible that slower performance in speed naming tasks may be unique to comorbid groups (Bental & Tirosh, 2007; Rucklidge & Tannock, 2002). Caution is advised, however, when interpreting such findings because of the lack of a reading-level design in the studies discussed.

Word reading efficiency. In contrast to findings for phonological awareness and non-word decoding, results indicated that participants in the dyslexia subgroup were not significantly less efficient than those in the RA-match control in timed word and non-word reading. However, both subgroups were significantly poorer in their word and non-word reading efficiency than the CA-match control subgroup. Group comparisons examined in terms of Cohen's *d* effect size ratios confirmed the pattern found in the ANOVA main analyses. The pattern of results also persisted after controlling for attention.

Findings for word and non-word reading efficiency could not be directly compared to other research findings in the literature since studies with a similar context to the present project were not found in which group differences on these specific measures were investigated. Considering the strong relationship between these two measures with word identification, non-word reading, and phonological awareness as demonstrated in the correlation analysis, one might expect difficulties in reading fluency and efficiency, especially because word and non-word reading efficiency rely on both accurate decoding and fluent reading. However, the null finding for the contrast between the dyslexia subgroup and the RA-age match control suggests that participants in the dyslexia subgroup may only be delayed in their word reading efficiency. That is, speeded word and non-word reading does not seem to represent a deficit or potential explanation for difficulties in reading. Considering that older children with reading difficulties may have more educational experience or be more mature cognitively or perceptually as suggested by Herrmann et al. (2006), they may be more likely to compensate for their phonological difficulties by developing other skills or strategies or by using other processes to read more efficiently. However, these strategies or processes may not be effective enough as those used by their same-aged peers who read normally and thus they may lag behind their peers in their reading fluency.

Motor and Cerebellar Measures

Postural stability and muscle tone. As indicated earlier, the cerebellar deficit hypothesis predicts that children with dyslexia should display deficits in motor- and cerebellar-related tasks in relation to a RA-match and a CA-match control group. However, contrary to the expectations derived from the cerebellar deficit hypothesis, the main analyses in the present study did not reveal any group differences on postural stability and muscle tone ratio. For the muscle tone measure, Cohen's *d* effect size ratios were small for all contrasts, indicating that the three subgroups did not really differ on this measure. For postural stability, the Cohen's *d* effect size ratio for the contrast between the dyslexia and RA-match subgroups was small. However, comparisons indicated moderately poorer performance in participants in the dyslexia subgroup as compared to their same aged peers with normal reading skills. This effect size for the contrast between the dyslexia and the CA-match subgroups decreased somewhat after the effect of attention was controlled. This finding contradicts the assumptions made by the proponents of the cerebellar deficit hypothesis (Fawcett et al., 1996; Fawcett et al., 2001) that cerebellar impairment in children with dyslexia is independent from the presence of attention difficulties. Instead the finding is consistent with the evidence that has shown attentional difficulties to be associated with balance impairment (e.g. Raberger & Wimmer, 1999, 2003; Ramus, Pidgeon et al., 2003; Rochelle & Talcott, 2006).

The null finding from the main analysis for postural stability for the dyslexia vs. RA-match control subgroup was similar in some regards to Savage et al.'s (2005a) results. In terms of effect size ratios, the RA-match control subgroup did not perform better than the dyslexia subgroup in this study as Savage et al. (2005a) found in their study. However, similar to Savage et al. (2005a), findings from the present investigation indicated that participants in the dyslexia subgroup were somewhat poorer than their same aged peers in their postural stability. Although this may provide some support for possible motor problems in the dyslexia subgroup, finding of equal performances found in the dyslexia and RA-match subgroups are not consistent with a causal role for cerebellar processing. As indicated by Goswami and Bryant (1989) positive results in a chronologically-age match group are not interpretable because of the different reading levels in the groups. As the authors have suggested, only positive findings for a reading-level group can be interpreted causally.

Furthermore, the comparisons made in this study in terms of effect size metric differed from those reported by Fawcett et al.'s (1996) as the large effect sizes reported in their study (i.e., -1 SD or greater) were not found here. There are some other factors that could also explain the difference in findings of the present study from those obtained in Fawcett et al.'s studies (e.g., 1996, 1999). For example, Fawcett et al. did not control for possible effects of attention in their sample. As findings of this study indicated, even the small group differences found for the dyslexia and CA-match subgroups in terms of effect size ratios seemed to be partially explained by attentional factors. This was consistent with findings from other studies which have indicated that attention seems to explain the differences found between poor and good readers on motor and cerebellar measures (e.g., Denckla et al., 1985; Raberger & Wimmer, 1999, 2003; Ramus, Pidgeon et al., 2003; Rochelle & Talcott, 2006; Wimmer et al., 1999).

Another factor that may explain the difference in the present findings to those obtained by Fawcett et al. (1996), is the way postural stability was measured in the two studies. As indicated earlier, in their early body of research, Fawcett and Nicolson (e.g., Fawcett & Nicolson, 1999; Fawcett & Nicolson, 1995c; Fawcett et al., 1996) used observer ratings and the Likert Scale to investigate postural stability and muscle tone. However, there are problems associated with this type of scale. For example, they may be less sensitive. They may also provide a more subjective estimation by observers, and they are expressed in qualitative units (Shipley & Harley, 1971). In an attempt to improve the sensitivity and reliability of findings, careful scientific measurement of postural stability and muscle tone was undertaken in this study using an accelerometer sensor and goniometer that was described in the Method chapter. It is possible that using more sensitive measurements such as those used in this study show that there may be no real cerebellar deficit. In fact, as was discussed in the second chapter and consistent with the present findings, several studies have failed to report this effect. Stoodley, Fawcett, et al. (2006) also used sensitive measures of cerebellar and motor processing similar to those used in this study, and did not find any balance deficits in a sample of adults with dyslexia. Similar results were reported by Brown et al. (1985) in a sample of children identified with dyslexia using the IQ-achievement discrepancy definition.

However, Moe-Nilssen et al. (2003) did find some evidence for balance deficits in their study using sensitive measures of motor processing, including accelerometer sensors similar to those used in the present study. Nonetheless, the impaired balance found in their study was exclusive to tests of undisturbed balance with eyes open (fixated on a visual target on the wall) and not when eyes were closed. Their findings were interpreted in terms of poor eye movements (i.e., inability to maintain steady fixation), which were suggested to possibly affect the children's ability to adequately take advantage of the visual cue positioned on the wall during the balance task. However, considering some factors in Moe-Nilssen's study that could have affected the interpretation of their results, such as sampling issues and the lack of a reading-level design, more research may be needed to further explore findings related to both challenged and unchallenged postural stability tasks.

Motor and toe tapping components. The main analysis for both motor and toe tapping components indicated significant group differences. However, neither group differences were in the direction predicted by the cerebellar deficit hypothesis. For both motor and toe tapping components, simple comparisons from the main analysis as well as comparisons using Cohen's *d* effect size metric indicated that in fact participants in the dyslexia subgroup in the sample were significantly better than the younger normal readers. The pattern of effect sizes also indicated that participants in the dyslexia subgroup were somewhat better in their motor skills as compared to their same aged peers. The overall pattern of results for motor and toe tapping components remained consistent when the effect of attention was controlled statistically.

The present findings for toe tapping and motor components in this study do not seem to provide support for the cerebellar deficit. In fact, the trend obtained for these two components using Cohen's *d* index of effect sizes was similar to the trend found in Savage et al.'s (2005a) and Savage (2007) studies for postural stability and bead threading measures. The pattern of the present results related to motor skills and toe tapping speed seem to be consistent with the notion of developmental maturity, as was also suggested by Savage et al. (2005a) for their postural stability results. Similar to Savage et al.'s (2005a) study, this notion seemed to be supported not only by the fact that participants in the dyslexia subgroup in the present sample performed equally well as their same aged peers who were normal readers, but also by the somewhat better performance of the latter subgroup as compared to younger normal readers. However, it should be noted that while some (e.g., Goswami & Bryant, 1989) have argued that it is possible to interpret null findings from chronologically-age match studies in causal terms, this must be done with extreme caution as a host of other extraneous factors may potentially explain any null effect (Jackson & Butterfield, 1989), such as using insensitive measures for the dependent variable.

Question 4:

Does the cerebellar deficit provide a good explanatory model at the individual level?

To examine individual differences in performance on reading, motor and cerebellar measures in the dyslexia subgroup and the RA-match control, the procedure formerly used by White et al. (2006) was followed with some differences that were explained earlier in the part 3 of Chapter 2. White et al.'s (2006) findings indicated that compared to cerebellar measures, phonological awareness tasks were more successful in differentiating between participants with dyslexia and their same-aged peers who were normal readers. Considering the alleged link between a cerebellar deficit, phonological processing, rapid naming, and reading, the motor and cerebellar measures were expected to be at least as successful as phonological and rapid naming measures in distinguishing between the participants in the dyslexia subgroup and those in the reading-age match control in this study. As before, interpretation of the findings pertaining to phonological, word reading efficiency, and rapid naming measures are reviewed first prior to considering interpretations of findings for motor-cerebellar variables.

Non-word decoding, phonological awareness, and word reading efficiency. In contrast to White et al.'s findings for phonological measures derived from their case level analysis, the overall pattern of results of the present study indicated that measures related to phonological processing used here did not distinguish well between the dyslexia and the RA-match subgroups at an individual level. A similar pattern was also found for word reading efficiency. The only observable pattern seemed to be that for both of the non-word reading measures (i.e., all items and polysyllables only) and phonological awareness fewer participants in the dyslexia subgroup had scores falling above the mean of zero as compared to those in the RA-match control. The pattern of results observed for all the enumerated measures seemed to persist after attention was controlled statistically.

Considering the overall pattern of findings for both non-word decoding and phonological awareness, it appears that at an individual level of reading, difficulties in the dyslexia subgroup do not seem to be as well-explained by a phonological deficit as they may be at a group level. This finding differs from that obtained in White et al.'s (2006) study. There may be a few reasons for the difference found in findings derived from the individual analysis between the present project and White et al.'s (2006) study. One reason for this difference may be in part related to the fact that White et al. removed extreme poor cases from their control group prior to performing individual analysis. Even the additional case that was identified as an extreme poor case on their literacy factor was also removed from the individual analyses which might have led to having no extreme poor cases on the phonology factor in their study. In this study none of the outliers in the RA-match control were removed from this subgroup. While removing the outliers may not have had a great impact, the pattern of the results might have been somewhat different, at least for phonological awareness.

Another reason that may have led to the different pattern of findings for case-level analysis in the two studies may be due to the type of scores used to display individual differences. Unlike White et al. (2006) who used standard (i.e., norm-referenced age corrected) scores to create their reading related factors, raw scores were used in this study for this purpose. It is possible that using standard scores for reading-related measures would have yielded similar findings to those in White et al.'s study. This is because standard scores are age-corrected scores that estimate whether a child's performance on a given task is average, below or above average as compared to his/her same-aged peers. In this sense, while most participants in the RA-match control subgroup would have obtained average standard scores on phonologically based tasks as compared to their same-aged peers, those in the dyslexia subgroup would have achieved below average scores on the same measures when compared to their same-aged peers. This, in turn, would have been likely to affect the pattern of findings obtained for phonologically based measures in the individual analysis.

In sum, the findings derived from the individual analysis for phonological measures are not entirely consistent with the evidence obtained from the group comparisons (questions 2 and 3 in this thesis have provided more solid evidence for phonological processing deficits in the dyslexia subgroup). It seems that in contrast to findings derived from group analyses, those obtained from the individual level analyses are ambiguous. In fact, these findings suggest that phonological awareness may not be as closely associated with difficulties in reading on a case-by-case level.

Alphanumeric rapid naming summary factor. The alphanumeric rapid naming measures did not seem to successfully distinguish between the dyslexia and RA-match subgroups in the present sample at an individual level. More specifically, in both the dyslexia and RA-match subgroups, a similar proportion of cases were identified as extreme poor cases. This pattern persisted even after statistical adjustments were made for the effect of attention. Overall, the pattern of the individual analysis for rapid naming seemed to follow the results that were obtained from the main analyses for these measures.

The present findings cannot be compared to White et al.'s (2006) since they subsumed their speed naming measures under the phonology factor. Generally, findings from both group and individual analysis seemed to suggest that, at least in this sample, deficits in alphanumeric rapid naming may not be associated strongly with reading difficulties. It also seems somewhat unclear whether rapid naming deficits affect extreme poor cases as suggested by Wolf and Bower (1999) or Wolf et al. (2002), as this pattern was ambiguous in the results of this study at an individual level. Caution may be advised considering the small number of extreme poor cases identified in the dyslexia subgroup in the sample.

Findings of the present study are generally consistent with some other evidence that has found unclear and ambiguous relationships between rapid naming deficits and reading (e.g., Pennington et al., 2001; Savage et al., 2005a). Importantly, the present findings also extend knowledge here. In light of using various statistical analyses (i.e., correlational analysis, group contrasts, and individual analysis) and controlling for possible extraneous factors, even though rapid naming measures seemed to be moderately correlated with measures of word reading and phonological awareness in the full sample, rapid naming measures did not successfully distinguish between the three subgroups of dyslexia versus RA-match versus CA-match at a group level. Only a moderate effect was found for rapid digit naming, indicating moderately poorer performance in the dyslexia subgroup as compared to the CA-match control. The effect was reduced to a small size once attention was controlled statistically. However, this finding is difficult to interpret because of differing reading levels in the two subgroups. Furthermore, rapid naming measures did not successfully distinguish between dyslexia and the RA-match subgroups at an individual level, suggesting that rapid naming is a modest but significant feature of variation in typical reading, and particularly speeded reading, but not a consistent feature of the cognitive profiles of the subgroup with dyslexia.

Motor and Cerebellar Measures

Postural stability summary factor. Postural stability did not distinguish between participants in the dyslexia and RA-match subgroups in the present sample and this pattern persisted after adjustments were made for attention. As the individual analysis indicated, a similar number of cases in both subgroups showed very low performance on the postural stability measure. Additionally, it remains unclear as to how useful postural stability may be as a measure that differentiates between children with dyslexia and normal readers in the sample, especially because only two of the five extreme poor cases on postural stability were identified as outliers on some of the literacy measures. Hence, the overall pattern of findings derived from various analyses appears to suggest that a strong association between postural stability and difficulties in reading at any level may be unlikely.

Muscle tone. Unlike the pattern found in the group level analysis, an interesting pattern was revealed for muscle tone at an individual level. At first glance, this measure seemed to distinguish between the dyslexia and RA-match subgroups to some extent. This pattern persisted after the effect of attention was removed. Despite this moderate success rate in distinguishing between the subgroups, muscle tone did not seem to be related to reading difficulties in the sample. As findings from the individual analysis indicated, except for one case in the dyslexia subgroup which was identified as an extreme poor case on alphanumeric rapid naming and word and non-word reading efficiency measures, none of the cases that were extremely poor on the muscle tone measure (including cases in both dyslexia and the RA-match subgroups) were identified as extreme poor cases on any of the literacy related measures.

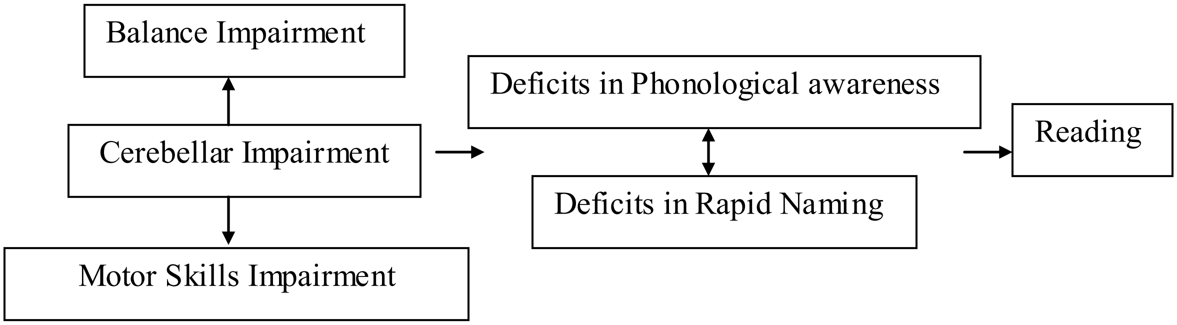
In all, this pattern suggested that poorer performance on muscle tone was not associated with extreme difficulties on literacy measures. Yet, it is difficult to determine whether the extreme poor cases on muscle tone were cases that displayed a true deficit (i.e., hypotonia) on this measure. In some studies that have examined hypertonia (muscle spasticity), ratios of more than 1.6 have been reported for elderly with normal muscle tones (Brown, Lawson, Leslie, MacArthur et al., 1988; Brown, Lawson, Leslie, & Part, 1988). However, a report on normal muscle tone ratio ranges could not be found for the age groups in the present sample. Additionally, while in the presence of hypotonia, higher muscle tone ratios would be expected, a comparison of how high the ratio should be was not available. In the extreme poor cases that were found on the muscle tone measure in both subgroups in the sample, ratios ranged from 1.72 to 2.12, which is close to the ratio that has been suggested as normal muscle tone for the elderly. In view of the fact that an independent measure of comparison was not available for children either typical or atypical, it cannot be assumed that the extreme low scorers in the present sample reflect genuine clinical extreme poor cases on this measure.

Toe tapping and motor summary factors. Findings for toe tapping and motor skills at an individual level were consistent with the pattern found in the group analyses. In other words, in line with those analyses, fewer children in the dyslexia subgroup as compared to RA-match control demonstrated deficits in toe tapping and motor summary factors. This pattern persisted even after adjustments were made for attention. As in the case of postural stability, the association between motor and toe tapping components with literacy measures did not seem to be strong. Specifically, only the two outlier cases on the motor summary factor were among the cases which had also demonstrated difficulties in some reading related measures. However, one of these cases was no longer an outlier on the motor summary factor after the effect of attention was controlled.

Overall, the pattern of findings for motor and cerebellar measures in this study was similar to White et al.'s (2006) findings. Consistent with their findings, results from the present investigation indicated that none of the motor and cerebellar measures used here successfully explain the difficulties in reading or distinguish between the dyslexia and RA-match subgroups in the sample, at both the group and individual level. Generally, the overall findings from group comparisons (i.e., positive findings for non-word decoding and phonological awareness and null finding for difficulties related to motor, and cerebellar measures) do not seem to provide strong support for a cerebellar deficit account of dyslexia and are more consistent with a phonological deficit account of this disorder as in White et al.'s (2006) findings. At an individual-level analysis, a few participants with motor/cerebellar difficulties in the dyslexia subgroup seem to display difficulties on other literacy measures. However, the general trend in findings, as in White et al.'s study, does not provide strong support for a causal relationship between a cerebellar deficit and reading difficulties. The use of a reading-level design, which was not a feature of White et al.'s study, as well as careful measures that were taken in sampling of this study adds further strength to the findings for the present study.

Summary of Findings

This thesis critically evaluated the cerebellar deficit theory of dyslexia, which as its authors claim provides a parsimonious explanation for the two existing cognitive deficit theories of dyslexia, namely the phonological and the speed naming deficit theories (Fawcett et al., 2001; Nicolson et al., 2001). Four major questions regarding the cerebellar deficit theory were addressed to investigate the presumed link, depicted below and discussed in Chapter 2, between the cerebellar deficit, motor and balance impairment, phonological awareness, rapid naming and reading.



The Link between a Cerebellar Deficit, Phonological Awareness, Rapid Naming and Reading

In the first question, it was investigated whether there is a relationship between word reading and (a) phonological awareness and reading fluency, (b) rapid automatized naming, and (c) purported cerebellar processing tasks. Contrary to what the cerebellar deficit theory suggests, no relationship existed among any of the literacy measures and those related to cerebellar and motor tasks. Instead, findings for this study suggest strong relationships among single word reading and measures of phonological awareness, phonological recoding, and word reading efficiency. Moderately strong relationships were also found among rapid naming measures and single word reading, phonological recoding and word reading efficiency.

Cerebellar Deficit as an Explanatory Model of Dyslexia at Group Level

In the second and third question, it was investigated whether (a) a subgroup of children with dyslexia selected from the sample differed in their performance on any of the motor, cerebellar, reading, phonological, and rapid naming related measures when compared to RA- and CA-match control subgroups selected from the same sample; and (b) if any group differences on the above measures emerged when the effect of attention was controlled statistically. Contrary to what is claimed in the cerebellar deficit hypothesis (Nicolson & Fawcett, 2006), the static cerebellar tasks, including postural stability and muscle tone, did not successfully distinguish between the subgroups in the present sample. According to Fawcett et al. (2001), while the incidence of deficits in phonological processing, speed naming, and motor tasks are high in both IQ-discrepant and IQ-consistent (i.e., garden-variety) poor readers, the extended difficulties in static cerebellar tasks are only present in the IQ-discrepant group. Hence, the authors claim that postural stability and muscle tone measures have the unique diagnostic power to discriminate between the IQ-discrepant and IQ-consistent poor readers. Even though this claim does not seem to be consistently supported by research findings (Savage, 2007), one may argue that the sample in the present study may not show typical deficits in static cerebellar tests since a discrepancy-based definition was not used to identify dyslexia in participants. Nonetheless, as indicated earlier in Chapter 2, most children in the dyslexia subgroup of the present sample did in fact seem to meet the Quebec definition of dyslexia, which is based on the discrepancy-based criterion. This means that for most of these children, reading performance seemed to be below what would be expected from their intellectual functioning. Consequently, if deficits in postural stability and muscle tone are unique to discrepancy-based dyslexia, they should have been present in most children in the dyslexia subgroup of the sample. In fact, exploring differences in terms of effect size ratios did indicate moderately poorer postural stability in participants in the dyslexia subgroup as compared to their same-aged peers who were normal readers.

Nonetheless, this difference could not be interpreted in causal terms because of the differing reading levels in the two subgroups. Additionally, consistent with some previous bodies of research and reviews (e.g., Raberger & Wimmer, 2003; Rochelle & Talcott, 2006; Wimmer et al., 1999), the degree of this difference decreased to some extent after the effect of attention was controlled.

Evidently, a difference between the dyslexia and the RA-match subgroups on the static cerebellar tests of postural stability and muscle tone would have been more meaningful especially since most of the previous investigations of the cerebellar deficit theory lack a reading-level design (e.g., Fawcett & Nicolson, 1995c, 1999; Fawcett et al., 2001; Moe-Nilssen et al., 2003; Nicolson & Fawcett, 1990; Raberger & Wimmer, 2003; Wimmer et al., 1999; White et al., 2006) and a demand for it has been voiced in the literature (Goswami, 2006). This was, however, not the case in the present study even though carefully-scaled scientific measurements instead of Likert scale measurements were used.

By using a reading-level design that includes both a RA- and a CA-match control group, findings for the motor and toe tapping components were more consistent with the notion of developmental maturity than the cerebellar deficit theory. This interpretation was supported by the results of the correlational analysis, which indicated that better performance on motor and toe tapping components was associated with increasing age. Additionally, findings from group contrasts indicated that the dyslexia subgroup in the sample performed equally well on motor and toe tapping measures as their same aged peers who were normal readers, but both subgroups were significantly better than the younger normal readers on these measures.

Unlike motor and cerebellar measures, phonologically based measures seemed to differentiate successfully between the subgroups in the sample. As findings indicated, both the dyslexia and the RA-match subgroups performed significantly poorer than the CA-match subgroup in non-word decoding and phoneme deletion. Additionally, group contrasts explored in terms of effect size ratios confirmed moderately poorer performances on both measures in the dyslexia subgroup as compared to the younger normal readers. More importantly, the effect size for the contrast between the dyslexia subgroup and the RA-match control increased from moderate to large for performance on polysyllable non-words only. This pattern persisted for phonological recoding but not for phoneme deletion after the effect of attention was controlled. However, regardless of whether the attentional difficulties were primary or secondary to reading impairment, a phonological deficit seemed to remain the best supported explanation for difficulties experienced in phoneme deletion for the dyslexia subgroup considering that the mean performance on this task remained below average even after attention was controlled.

The findings of the present study also indicated that despite being associated with word reading and phonologically based measures, measures related to speed naming and fluency were not successful in differentiating between the subgroups in the sample. Findings from the main analysis indicated that participants in the dyslexia subgroup were not significantly less efficient in timed word and non-word reading as compared to younger normal readers, while both subgroups were significantly less efficient than older normal readers. This pattern was also confirmed by the comparisons using effect size ratios and it persisted after the effect of attention was controlled. Furthermore, rapid naming tasks were also not successful in distinguishing between participants in the dyslexia subgroup and those in the RA- and CA-matched controls. Although comparisons using effect size ratios provided some support for possible difficulties in rapid digit naming, these findings were weak and not readily interpretable considering that the differences were found between the dyslexia and CA-matched subgroups who differed in reading level. Findings also indicated that attention seemed to explain poorer performance in rapid digit naming for some participants in the dyslexia subgroup.

Cerebellar Deficit as an Explanatory Model of Dyslexia at Individual Level

Finally, the fourth question of this thesis investigated whether a cerebellar deficit provided a good explanatory model at the individual level. Considering the predicted link between a cerebellar deficit, phonological and rapid naming processes and reading, the motor and cerebellar-related measures should be successful in distinguishing between participants in the dyslexia and the RA-match subgroups at an individual level. This was, nonetheless, not supported by findings of the present research. At an individual level, the two measures of muscle tone and postural stability seemed to be moderately successful in differentiating between participants in the dyslexia and RA-match subgroups. However, further exploration of the extreme poor cases indicated that poor performances on these measures may not necessarily be accompanied by poor performance on literacy measures. The pattern of findings obtained for motor and toe tapping factors also seemed to follow those obtained from the main analyses. Overall, deficits in motor and cerebellar measures did not seem to be common among participants in the dyslexia subgroup in the present sample. That is, among the five extreme poor cases identified on the reading factor (i.e., word reading task), only one displayed extremely poor performance on postural stability and another on the motor factor which were accompanied with some other literacy measures. The overall pattern of these findings was similar to those obtained by White et al. (2006) for their motor and cerebellar measures. What is noteworthy is the fact that individual level comparisons in the present study were performed using an RA-match control group as opposed to a CA-match control group used by White et al. (2006) and findings were still replicated. Considering the null findings for the group analyses and the small effect sizes obtained for simple group contrasts in this study, it is unlikely that individual level comparisons using a CA- instead of an RA-match control group would have yielded a different pattern of results for motor and cerebellar measures.

Interestingly, findings derived from the individual analysis indicated that phonologically based tasks may not be more successful than other measures used in this study to distinguish between participants in the dyslexia subgroup and the RA-match control. Specifically, the overall pattern of the present findings seemed to suggest that in contrast to White et al.'s (2006) study which compared the dyslexia group against a CA-match control group, comparisons against a RA-match control group lead to more ambiguous findings for phonologically based measures at a case-level analysis. Evidently, the pattern for phonological measures obtained in the present study might have been different if comparisons were made against a CA-match control group. Unlike motor and cerebellar measures, large effect size ratios were found for phonological measures for the contrasts between the dyslexia and CA-match subgroups in this study. Consequently, it is likely that a case-level analysis between these two subgroups would have yielded a similar pattern of findings to those in White et al.'s study. Nevertheless, findings derived from such comparisons would be difficult to interpret in causal terms since the groups differ in their reading level (Goswami & Bryant, 1989). In sum, while the present findings at the group level seem to be more consistent with a wide body of convergent evidence that has pointed to both the central place of phonological awareness in explaining variation in typical reading, as well as of an underlying phonological deficit explaining the reading difficulties of most poor readers and children with dyslexia (Bradley & Bryant, 1983; Snowling, 1998; Vellutino et al., 2004; Vellutino et al., 1996; Vellutino, Scanlon, Small, & Fanuele, 2006; Wagner & Torgesen, 1987), the results for individual level comparisons do not seem to follow this trend.

Similar to findings derived from group analysis, the individual-level findings also indicated that measures related to speed naming and fluency were not successful in distinguishing between the dyslexia and RA-match subgroups. Moreover, alphanumeric difficulties also did not seem to be as common among the extreme poor cases identified in the dyslexia subgroup since the results seemed to be more ambiguous in this sense. That is, among the five extreme poor cases in the dyslexia subgroup on word reading, only one was among the extreme poor cases identified on the rapid naming summary factor that also displayed non-word decoding difficulties. There was one other case with rapid naming difficulties with accompanied non-word decoding and phonological awareness difficulties. However, this case did not belong to the original extreme poor cases on the word reading task. This case was also no longer an outlier on phonological awareness after adjustments were made for the effect of attention.

Conclusion

In considering the four questions related to the cerebellar deficit theory of dyslexia (Nicolson & Fawcett, 1999; Nicolson et al., 2001), this thesis considered data drawn from a single database of school-sampled children analyzed completely or in part using a range of techniques. These techniques were: (a) an individual differences analysis of the full sample of typical children (n = 85) using correlational techniques, (b) principled group contrasts of a dyslexia subgroup versus reading-age match and chronologically-age match controls (n = 17 in each case), (c) principled group contrasts mentioned above with control for extraneous attention factors, and (d) analysis of the deficits at an individual child level for the dyslexia subgroup above compared to the reading-age match controls with and without control for attention. One of the general advantages this complete set of designs *potentially* offers is the convergence or consistency of findings at different grains of analysis. An actual strength of the findings of the present thesis that follows from such potential methodological strength was the clarity of the findings reported across analyses for the cerebellar deficit hypothesis. More importantly, the fact that the present study included reading-level controls is crucial for drawing causal interpretations, as has been suggested by Goswami (2006).

In sum, what the present findings indicated is that motor and cerebellar measures did not stand up well in *any* grain of analysis in techniques one to four as described above as descriptors of reading ability, typical or otherwise. Contrary to what would be expected considering the claims made by the cerebellar deficit theory of dyslexia, no relationship was found between any of the motor and cerebellar measures and those related to reading, phonological awareness, and rapid naming. Motor and cerebellar tasks also did not successfully differentiate between participants in the dyslexia subgroup and those in the RA- and CA-match controls both at a group and at an individual level. Indeed, findings appear to indicate that phonological and/or speed-naming difficulties exist independently from motor and cerebellar problems. As findings specified, in contrast to motor and cerebellar measures, phonologically based measures stood up relatively well in all but one grain of analysis (i.e., points 1 to 3) as described above as descriptors of reading ability, typical or otherwise. The positive findings for the dyslexia and RA-match subgroups suggest that a phonological deficit, independent from a cerebellar deficit, may provide the best-supported description for reading difficulties in the dyslexia subgroup. Moreover, the individual-level findings seemed to confirm this independence since cases displaying difficulties related to motor and cerebellar measures accompanied by literacy measures were rare.

In all, the overall pattern of the results of this study indicates that a cerebellar deficit does not provide an explanation for phonological awareness, rapid naming, and reading difficulties. These findings appear to be consistent with some other research that has failed to find strong support for a cerebellar deficit account of dyslexia (e.g., Raberger & Wimmer, 2003; Ramus, Pidgeon et al., 2003; Savage et al., 2005a; Needle et al., 2006; White et al., 2006). The present results, however, differ from some other studies that have found support for the cerebellar deficit theory (e.g., Fawcett & Nicolson, 1995c, 1999, 1996; Moe-Nilssen et al., 2003). Nevertheless, the differences in these findings are not related to the sample in this study or statistical power to detect meaningful differences. Given the fact that meaningful differences were detected for phonologically based measures in the sample, such differences should have also been detected for motor and cerebellar measures if they were present. Indeed as indicated earlier, moderate differences were detected for postural stability in contrast against the CA-match control subgroup. But these differences were not only hard to interpret due to the groups' differing reading levels, they were also diminished once attention was controlled statistically. Furthermore, moderate and large differences were detected for top tapping and motor skills, respectively. However, the pattern of these differences did not support a cerebellar deficit but rather the notion of developmental maturity. In short, the cerebellar deficit theory was not supported by findings derived from the present study despite using a range of statistical techniques. More importantly, findings from this study were strengthened by the careful measures that were taken to improve upon several methodological shortcomings that seemed to explain the inconsistencies across research findings for the cerebellar deficit theory. Specifically, the present findings did not provide strong support for a cerebellar deficit account of dyslexia despite (a) statistical control for attentional factors, (b) use of wider range of motor and cerebellar measures along with using more sensitive measures for postural stability and muscle tone tasks which yielded continuous data, (c) including a reading-level design, and (d) use of a more homogeneous sample that included non-clinical, mainstream children.

Limitations of the Study

Despite the strengths of this study enumerated in the second chapter and earlier in this section, this study also included some limitations.

Limitations Related To Sample.

Although some level of homogeneity was present across the three subgroups that were selected from the sample, there may be some biases associated with the sample in general. The offers that were made to schools and parents, such as short reports on children's performance on reading related measures or packages including intervention strategies, may have enticed specific schools or parents to participate. For example, the process of screening for reading abilities which was a part of this study as well as the short reports prepared for children may have made this study more attractive to schools that were experiencing a shortage of school psychologists and were in need of assessment.

Furthermore, while the three subgroups were matched on their intellectual functioning and homogenous in the proportion of first language spoken, they were spread across three different schools. Consequently, the method of instruction used in these schools could have also had an effect on the children's reading performances. It should be noted that the principal researcher orally inquired about the method of instruction in the three schools. Except for one school in which the teachers were not available to discuss their method of instruction, a few teachers reported using more or less similar strategies. Nonetheless, these were only short oral reports provided by a few teachers and this study did not control for possible differences that may have existed in the method of instruction across these schools. As was indicated earlier in Chapter 2, this limitation is commonly shared by many studies on dyslexia, including those reviewed in this thesis. Indeed, it has only been recently that the cognitive characteristics of children who are known to have not responded well to intensive intervention are beginning to be studied (Deault & Savage, in press).

Moreover, as was indicated in the recruitment process of the Method chapter, this study was at first conducted simultaneously with another project that focused on eye movements during reading. In hindsight, it might have been more effective if the present research was conducted independently from the start especially considering the rate of participant attrition due to technical difficulties related to the eye movement equipment. Conducting both studies together might have also had a negative impact on recruiting participants because the proposals sent to schools included a larger number of tasks, thus demanding increased testing time and larger space to fit the equipment related to the eye-movement project. Consequently, schools might have been more reluctant to participate considering that testing time interfered with teaching periods and the fact that many schools had only limited space available for testing.

In addition to the enumerated complications that might have affected participant recruitment, there were other factors that played an important role in limiting the number of potential Anglophone or English-dominant bilingual participants available to this study. Bi/multilingualism is a common phenomenon in the city where the study took place. Subsequently, it was nearly impossible to recruit a sample that included only monolingual English speaking children unlike in the US or the UK where a monolingual English-speaking sample would be readily available. Nonetheless, it is crucial to note that bi/multilingualism is different within the context of Montreal from what it would often be, for example, in the USA where a child grows up learning his/her mother tongue as the first language (e.g., Spanish) and then learns English later as a second language. Most children who grow up in bilingual homes in Montreal can probably be considered equal status bilinguals at birth because they grow up equally strong in both languages. The important issue regarding sampling for this study was to target English-dominant bilinguals and this limited the number of potential participants available to this project. As was indicated earlier in the Method section, participant recruitment was limited only to schools in which English was the leading language since this guaranteed that the sample would generally represent genuine English-dominant bilinguals. It was also indicated that among the schools with English as the leading language, selection had to be limited only to those with English as the leading language during the elementary years. This further limited the recruitment of potential participants in this study. Finally, as noted earlier in the Method chapter, the closure of schools associated with the English School Boards was another factor that also affected participant recruitment.

The wide age range in this study may also be another limitation as it was a criticism addressed in White et al.'s study. Controlling for this factor is crucial in any study that includes a wide age range and this was the case in this research. Strong and potentially important effects in correlation analysis with statistical controls for chronological age could still be demonstrated. In another sense, however, the advantages and disadvantages of a wide chronological age range of participants across this research had to be weighed. A reasonably wide age range in the general sample was required in order to construct a reading-level design that incorporated both a reading-age and a chronologically-age match control. Sampling so as to be able to construct this design was particularly desirable, as discussed earlier, because of the potentially causal power of such designs. Thus, a principled decision was made that this was more desirable than the execution of a study of individual differences with *n* = 85 participants, all of which fell within a narrow age range.

Limitations Related to Measures

Measure of attention. As a result of time limitations, the study included only a behavioral rating scale to assess possible attentional difficulties in children as in many other studies. Behavioral ratings, used alone, are not the most reliable measures to assess attention or other behavioral or emotional difficulties, as they can be biased not only in general but especially in the cultural diversity that represents the nature of the population in the Greater Montreal area. Ideally, a multimodal assessment would have increased the reliability of findings related to the presence or absence of ADHD in participants but time did not allow for this. Asking both parents and teachers to complete the behavior rating scales would have also been another more optimal way to generate more reliable results. Considering teachers’ work overload and their time shortage, this was also not possible. Nonetheless, as was explained earlier in the Method chapter, some of the behavior rating scales were in fact completed by teachers instead of by parents. This might also have affected the data generated on ADHD since parents and teachers can have differing perceptions of child’s behavior given the fact that they interact with them in different contexts. However, examinations did indicate high correlations between the responses derived from parents and those derived from teachers which provided some level of reliability for the data that was obtained from these two different sources.

Measure of intelligence. Because of time limitations, a brief measure of intelligence was used in this research, which provided an estimate of intellectual functioning. While this brief measure is highly correlated with the full scale measure as reported in the Method chapter, a more comprehensive scale of intelligence would have probably provided us with more accurate verbal and nonverbal skills. However, the measure has been reported to be reliable and correlations between this abbreviated scale and the comprehensive test of cognitive functioning in the Wechsler series, namely the Wechsler Intelligence Scale for Children-Third Edition (WISC-III) are reported to be high (i.e., .82 for Verbal IQ, .76 for Performance IQ, and .82 for the 2-subtest Full Scale IQ) (Saklofske et al., 2000). More importantly, IQ measure was not the main focus of this study since a body of literature, as discussed in the Introduction, has pointed to its irrelevance (e.g., Fletcher et al., 1992; Siegel, 1988; Siegel, 1989; Siegel, 1992).

Measure of non-word decoding. As noted previously, the Word Attack task used to assess non-word decoding ability included both mono- and polysyllable non-words, as this is a psychometric test. The test is not specifically designed to assess ability on decoding polysyllable non-words. In this sense, the analysis on polysyllable items in this study was limited. The Word Attack test itself may also be limited considering that Rack et al.'s (1992) review of non-word reading deficit points to the important role that the level of complexity of non-words may play in identifying individuals with dyslexia. Nevertheless, despite these limitations, the large effect size found in this study for polysyllable non-words replicated previous findings reviewed by Rack et al. (1992).

Measure of postural stability. There were also some limitations associated with the postural stability task. One of the limitations was related to Fawcett and Nicolson's (1996) balance tester device that was used to administer the balance challenge to participants. Although the device was calibrated on a kitchen scale as recommended by Fawcett and Nicolson (1996), to control the amount of pressure applied to the back in order to challenge balance the balance tester did not seem to be the best method for administering a challenge since it still involved some amount of human error. Moe-Nilssen et al. (2003), for example, controlled for this factor by using external forces to administer the balance challenge. However, their method might not be readily available to use by others since it involves acquiring specific equipment for this purpose which was not feasible for this study.

Additionally, I controlled for one of the physical factors, namely participants' weight, to ensure that this did not have an affect on the degree of sway that occurred after administering the balance challenge. However, there are other physical factors that may also affect the postural stability task which were not controlled in this study, such as participants' height. The postural stability task in this study also followed Fawcett and Nicolson's (1996) DST, which included measuring postural stability upon administering a balance challenge with eyes closed and hands at sides or stretched out in front. There are, however, other forms of postural stability tasks as well as measures of postural stability which were not included in this study, such as measures of postural stability without administration of a balance challenge (e.g., heel-to-toe) or measure of gait (Moe-Nilssen, 2003), or the amount of time taken to regain balance upon administration of the balance challenge as was recommended by Dr. P. Stapley (personal communication, Oct. 6, 2008).

As indicated in the Method chapter, the reliabilities obtained for the postural stability measures were overall not as high as the reliabilities obtained for the other measures used in this study. This was despite using a sensitive device to measure this task. Nonetheless, this may also raise some questions as to how reliable this task may be considering the overall inconsistent evidence and often skewed data derived for the postural stability task using both sensitive and subjective measures. As Ramus et al. (2006) have also noted, the possible inadequacy of current methods to measure cerebellar skills may be a potential reason for the null findings often reported for this measure. Ramus et al. advise that the onus is on proponents of sensorimotor theories to improve their methods.

Measure of muscle tone. Some limitations were also associated with the Pendulum test of muscle tone measure. For example, as indicated in the Method chapter, the level of relaxation in participants is important in the Pendulum test in order to ensure that participants don't assist or resist leg swings. Despite efforts made to achieve a level of relaxation in participants prior to performing the actual task, a few participants in the sample still seemed to assist during the task. While this could be controlled to some extent, the level of relaxation in participants has been reported as one of the limitations of the Pendulum test. Other factors that have also been reported by some researchers to affect the muscle tone task are posture, muscle length, and starting angle (Fowler et al. 1998; Amman et al., 2008). A few researchers have also questioned the level of sensitivity that the Pendulum test may show to truly measure both hypotonia and hypertonia (e.g., Fowler et al., 2000).

Sensitivity of motor/cerebellar measures to detect gross damage. Generally, the motor and cerebellar measures are indirect behavioral tests listed to assess gross damage to the cerebellum. The question remains as to how sensitive these tests may be to assess possible cerebellar deficits, which are suggested to be subtle in the case of dyslexia. Moreover, a lack of standardization for samples or a basis for comparison also limits the findings related to these subtle deficits to some extent.

Lack of test-retest reliability for some measures. Unfortunately, because of time limitations obtaining a test-retest reliability for some of the measures including the rapid naming tasks as well as toe tapping and bead threading measures was not possible since they were only administered once. Consequently, no reliability for these measures was presented in this study. Nonetheless as discussed in the Method chapter, the published test-re-test reliabilities reported for rapid naming and bead threading measures are generally high.

Overall, considering the methodological limitations here along with some of the structural limitations associated with designs such as reading-age match studies as discussed in the literature review earlier, all results described here should be interpreted with appropriate caution, and particularly, in isolation, as strong causal evidence of deficits in poor reading.

Implications of Findings and Directions for Future Research

Overall, the present findings did not provide support for a cerebellar deficit account of difficulties in reading but rather for cognitive models of dyslexia that emphasize phonological processing as a potential cause of poor reading and of variation in typical reading. These findings may have both clinical and theoretical implications.

Generally, findings of this study may have important clinical implications as to what specific skills may be crucial as screening tools for identifying children at risk for reading difficulties among their same-aged peers. Considering the large effect sizes that were found for the contrast between the dyslexia subgroup and the CA-match control for phonological awareness, non-word decoding, and reading fluency, these tasks seem to be appropriate screening tools for identifying children at risk for reading impairment among their same-aged peers. Including such tasks in comprehensive assessments may also be crucial for a more accurate diagnosis of dyslexia. In this sense, as opposed to findings related to phonologically based tasks, the results related to motor and cerebellar measures seem to suggest that processes underlying these measures may not be either concurrently associated with or causally related to reading. Thus, despite Nicolson and Fawcett's (2006) claim about the diagnostic power of the extended difficulties in static cerebellar tasks (including balance and muscle tone), such measures do not appear to have the same success rate as phonologically based tasks in identifying at-risk readers or in distinguishing between poor and good readers. These findings may have importance in ensuring good practices in light of the limited resources and funds that may be available in schools both for assessment and identification. They may also be important for implementing interventions directed toward prevention and improvement of reading difficulties.

While the present project was not an intervention study, the results provided indirect support for evidence derived from these studies which have shown that early training that focuses on improving phonologically based processes may be crucial for prevention of future reading difficulties and for improving reading practices in poor readers. In contrast, intervention techniques that focus on motor skills and postural stability such as the remedial techniques in the Dore Achievement Centers (Reynolds & Nicolson, 2007) may not have similar benefits as they do not seem to address the cause underlying reading difficulties. Hence, prior to implementing major changes in practices related to dyslexia based on findings related to a specific hypothesis, researchers hold an important responsibility to ensure that future research is clear about what disorder is being treated, which aspects of the disorder are being affected by certain treatments, or what the treatment is (see Irannejad & Savage, 2009).

This is certainly not to encourage abandoning research on motor and cerebellar deficits or other processes that may be involved in reading. It is important to consider the possibility that there may not be a single cause for dyslexia and that limiting it to a single cause may also be just as harmful, since it could ignore those who do not fit under the category of reading disabilities as defined by that cause. Indeed findings derived from individual level analysis here suggest that neither the motor and cerebellar measures, nor the measures related to phonological awareness, speed naming and fluency differentiated well between participants in the dyslexia subgroup and younger normal readers. In other words compared to group-level comparisons, findings became more ambiguous at the individual level. Further, while none of the measures included in this study successfully distinguished between the subgroups at an individual level, there were participants, although few in number, who seemed to display difficulties other than those related to phonological awareness. Evidently, further investigation using a much larger sample size is needed to draw firm conclusions as other important processes may either be directly related to reading or merely coexist in individuals with reading difficulties. As van IJzendoorn and Bus (1994) have indicated, experiences in the early stages of learning to read may also be crucial. This is especially important considering that one of the major limitations in the majority of studies on dyslexia (including the present study and those reviewed here) is a lack of knowledge on the educational history of the samples as well as a lack of knowledge on how the samples involved respond to intervention (Deault & Savage, in press).

Nonetheless, what the present findings do seem to point out is that motor and cerebellar measures or processes may not necessarily be directly associated with reading difficulties. Wider research suggests that while they seem to co-exist in a group of readers with dyslexia, in some of these cases, these processes may be associated with or explained by other factors such as attention and developmental maturation, which were demonstrated in this and other studies (Ramus, Pidgeon et al., 2003; Savage et al., 2005a) or a developmental coordination disorder as suggested by (Herrmann et al., 2006). Consequently, it is important to exert some level of caution in drawing conclusions regarding these deficits from published studies, especially if careful measures have not been taken in sampling, design, and methodology and in the level of control for possible confounding factors that may affect findings reported.

In terms of future research, larger and well-designed reading-age match studies as well as intervention-based studies and longitudinal studies are needed to establish causal models (Goswami & Bryant, 1989). Convergent evidence from these three types of studies is more likely to establish a causal role of variables. The best evidence may come from longitudinal studies that include both individuals with dyslexia and normal readers in order to establish: (a) whether there is any link between the cerebellum and tasks involving phonological processing; (b) whether such a link, if any, is more evident at earlier ages when these skills are developing in a child; and (c) how such links may change as these skills become more automatized. Additionally, studies could also investigate how learning to read and practice in reading may change the possible involvement of the cerebellum in this process. It is also crucial to take into consideration that just as certain brain structures and their abnormalities may be related to some problems in reading, an individual's reading habits and his or her environmental experiences may equally affect the brain structures. Therefore, how can we be sure that abnormalities in the cerebellum or other areas of the brain are causing the reading difficulties and not vice versa? There is increasing evidence that brain structure can be modified in many ways with experience. For example, functional changes in brain activity (essentially, normalization) have in fact been observed after short or longer periods of training in phonological awareness and decoding (Aylward, et al., 2003; Eden, et al., 2004; Richards, et al., 2000; Shaywitz et al., 2004; Simos et al., 2002). These changes at the neurological level are reliably accompanied by improvement in the cognitive processes targeted by the training. Findings from these intervention studies are intriguing and they need to be explored further using longitudinal studies

As discussed earlier in the literature review, generally evidence related to the cerebellar deficit seems to be inconsistent even in studies that have used more sensitive measures. The inconsistencies in findings seem to also be partially related to different methodologies used or the way that motor and cerebellar tasks are assessed or measured. While use of a variety of tasks is important in determining possible subtle deficits that may exist, the use of more standardized, common, and unified practices to assess tasks related to motor and cerebellar deficits in larger scale studies that include both clinical and non-clinical samples may also be required in order to improve the interpretations of findings.

Finally, more sophisticated models would arguably consider the possibility that dyslexia and its manifestations, both in terms of observable symptoms and brain structures involved, could be a result of an interaction between biological factors within individuals (such as hereditary factors) as well as the environmental factors affecting the individual such as opportunities for learning, method of instruction, cultural and language factors, or the individual's own reading habits and compensation methods used to facilitate reading. It is crucial for studies to investigate such issues using longitudinal studies to compare changes in cognitive processes and brain structures, such as the cerebellum, in normal readers and individuals with dyslexia in conjunction with well-designed intervention studies exploring the impacts of phonological, speeded naming and motor interventions (Irannejad & Savage, 2009).

References

Ackerman, P. T., & Dykman, R. A. (1993). Phonological processes, confrontational naming, and immediate memory in dyslexia. *Journal of Learning Disabilities, 26*, 597-609.

Ackerman, P. T., Dykman, R. A., & Gardner, M. Y. (1990). Counting rate, naming rate, phonological sensitivity, and memory span: Major factors in dyslexia. *Journal of Learning Disabilities, 23*, 325-327.

Ackermann, H., Graber, S., Hertrich, I., & Daum, I. (1997). Categorical speech perception in cerebellar disorders. *Brain and Language, 60*, 323-331.

Ackermann, H., Graber, S., Hertrich, I., & Daum, I. (1999). Cerebellar contributions to the perception of temporal cues within the speech and nonspeech domain. *Brain & Language, 67*, 228-241.

American Psychiatric Association. (1987). *Diagnostic and statistical manual of mental disorders* (3rd ed.). Washington DC: American Psychiatric Association.

Ammann, C. M., Kawanami, L. M., Giratalla, M. M., Hoetmer, R. A., Rodriguez, V. J., & Munro, K. K. (2005). Choosing a spasticity outcome measure: A review for the neuromodulation clinic. *University of Alberta Health Sciences Journal, 2*, 29-32.

Annett, M., Hudson, P., & Turner, A. (1974). The reliability of differences between the hands in motor skill. *Neuropsychologia 12*, 527-531.

Aylward, E. H., Richards, T. L., Berninger, V. W., Nagy, W. E., Field, K. M., Grimme, A. C., et al. (2003). Instructional treatment associated with changes in brain activation in children with dyslexia. *Neurology, 61*, 212-219.

Ayres, A. J. (1973). *Sensory integration and learning disorders*. Los Angeles: Western Psychological Services.

Ayres, A. J. (1978). Learning disabilities and the vestibular system. *Journal of Learning Disabilities, 11*, 30-41.

Backman, J. (1983). The role of psycholinguistic skills in reading acquisition: A look at early readers. *Reading Research Quarterly, 18*, 466-479.

Backman, J. E., Mamen, M., & Ferguson, H. B. (1984). Reading level design: Conceptual and methodological issues in reading research. *Psychological Bulletin, 96*, 560-568.

Baddeley, A. D., Gathercole, S. E., & Papagno, C. (1998). The phonological loop as a language learning device. *Psychological Review, 105*, 158-173.

Baddeley, A. D., Logie, R. H., & Ellis, N. C. (1988). Characteristics of developmental dyslexia. *Cognition, 29*, 197-228.

Baddeley, A. D., Thomson, N., & Buchanan, M. (1975). Word length and the structure of short-term memory. *Journal of Verbal Learning & Verbal Behavior, 14*, 575-589.

Badian, N. A. (1993). Predicting reading progress in children receiving special help. *Annals of Dyslexia*, 90-109.

Badian, N. A. (1996). Dyslexia: A validation of the concept at two age levels. *Journal of Learning Disabilities, 29*, 102-112.

Baker, C. (1993). *Foundations of bilingual education and bilingualism*. Clevedon, UK: Multilingual Matters.

Beaton, A. A. (2002). Dyslexia and the cerebellar deficit hypothesis. *Cortex, 38*, 479-490.

Beaton, A. A. (2004). *Dyslexia, reading, and the brain: A sourcebook of psychological and biological research* (pp. 3-24). New York: Psychology Press.

Bental, B., & Tirosh, E. (2007). The relationship between attention, executive functions and reading domain abilities in attention deficit hyperactivity disorder and reading disorder: A comparative study. *Journal of Child Psychology & Psychiatry, 48*, 455-463.

Berninger, V. W., Thalberg, S. P., DeBruyn, I., & Smith, R. (1987). Preventing reading disabilities by assessing and remediating phonemic skills. *School Psychology Review, 16*, 554-565.

Biel, A. (2005). *Trail guide to the body* (3rd ed., pp. 231-272). Boulder, CO: Books of Discover.

Bishop, D. V. (2006). Dyslexia: What's the problem? *Developmental Science, 9*, 256-257.

Bishop, D. V. M. (1990). *Handedness and developmental disorder.* England: Mac Keith Press.

Boada, R., & Pennington, B. F. (2006). Deficient implicit phonological representations in children with dyslexia. *Journal of Experimental Child Psychology, 95*, 153-193.

Bohannon, R. W. (1987). Variability and reliability of the pendulum test for spasticity using a Cybex II Isokinetic Dynamometer. *Physical Therapy, 5*, 659-661.

Bowers, P. G. (1995). Tracing symbol naming speed's unique contributions to reading disabilities over time. *Reading and Writing: An Interdisciplinary Journal, 7*, 189-216.

Bowers, P. G., & Newby-Clark, E. (2002). The role of naming speed within a model of reading acquisition. *Reading & Writing, 15*, 109-126.

Bowers, P. G., Steffy, R., & Tate, E. (1988). Comparison of the effects of IQ control methods on memory and naming speed predictors of reading disability. *Reading Research Quarterly, 23*, 304-319.

Bowers, P. G., Steffy, R. A., & Swanson, L. B. (1986). Naming speed, memory, and visual processing in reading disability. *Canadian Journal of Behavioural Science, 18*, 209-223.

Bowers, P. G., & Swanson, L. B. (1991). Naming speed deficits in reading disability: Multiple measures of a singular process. *Journal of Experimental Child Psychology, 51*, 195-219.

Bowers, P. G., & Wolf, M. (1993a, March). *A double-deficit hypothesis for developmental reading disorders.* Paper presented at the Biennial Meeting of the Society for Research in Child Development, New Orleans, LA.

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

Bowey, J. A., Cain, M. T., & Ryan, S. M. (1992). A reading-level design study of phonological skills underlying fourth-grade children's word reading difficulties. *Child Development, 63*, 999-1011.

Bowey, J. A., & Hansen, J. (1994). The development of orthographic rimes as units of word recognition. *Journal of Experimental Child Psychology, 58*, 465-488.

Bradley, L., & Bryant, P. (1983). Categorizing sounds and learning to read: A causal connection. *Nature 301*, 419-421.

Brookes, R. L., & Stirling, J. (2005). The cerebellar deficit hypothesis and dyslexic tendencies in a non-clinical sample. *Dyslexia: An International Journal of Research and Practice, 11, 174-185*.

Brown, B., Haegerstrom-Portnoy, G., Herron, J., Galin, D., Yingling, C. D., & Marcus, M. (1985). Static postural stability is normal in dyslexic children. *Journal of Learning Disabilities, 18*, 31-34.

Brown, R. A., Lawson, D. A., Leslie, G. C., MacArthur, A., MacLennan, W. J., McMurdo, M. E., et al. (1988). Does the Wartenberg pendulum test differentiate quantitatively between spasticity and rigidity? A study in elderly stroke and Parkinsonian patients. *Journal of Neurology, Neurosurgery & Psychiatry, 51*, 1178-1186.

Brown, R. A., Lawson, D. A., Leslie, G. C., & Part, N. J. (1988). Observations on the applicability of the Wartenberg pendulum test to healty, elderly subjects. *Journal of Neurology, Neurosurgery & Psychiatry, 51*, 1171-1177.

Bruck, M. (1992). Persistence of dyslexics' phonological awareness deficits. *Developmental Psychology, 28*, 874-886.

Bryant, P., & Goswami, U. (1986). Strengths and weaknesses of the reading level design: A comment on Backman, Mamen, and Ferguson. *Psychological Bulletin, 1*, 101-103.

Campbell, D. T., & Stanley, J. C. (1963). *Experimental and quasi-experimental designs for field settings*. Chicago: Rand McNally College.

Cardoso-Martins, C., & Pennington, B. F. (2004). The relationship between phoneme awareness and rapid serial naming skills and literacy acquisition: The role of developmental period and reading ability. *Scientific Studies in Reading, 8*, 27-52.

Catts, H. W., Gillispie, M., Leonard, L. B., Kail, R. V., & Miller, C. A. (2002). The role of speed of processing, rapid naming, and phonological awareness in reading achievement. *Journal of Learning Disabilities, 35*, 510-525.

Chan, W. S. R., Hung, S. F., Liu, S. N., & Lee, C. K. K. (2008). Cognitive profiling in Chinese developmental dyslexia with attention-deficit/hyperactivity disorders. *Reading & Writing, 21*, 661-674.

Chiappe, P., Stringer, R., Siegel, L. S., & Stanovich, K. E. (2002). Why the timing deficit hypothesis does not explain reading disability in adults. *Reading & Writing, 15*, 73-107.

Clay, M. M. (1987). Learning to be learning disabled. *New Zealand Journal of Educational Studies, 22*, 155-173.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). New York: Academic Press.

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112*, 155-159.

Conners, C. K. (1997). *Manual for the Conners' rating scales - revised*. North Tonawanda, NY: Multi-Health Systems.

Cossu, G. (2003). The role of output speech in literacy acquisition: Evidence from congenital anarthria. *Reading & Writing, 16*, 99-122.

Cossu, G., Shankweiler, D., Liberman, I. Y., Tola, G., & Katz, L. (1988). Awareness of phonological segments and reading ability in Italian children. *Applied Psycholinguistics, 9*, 1-16.

de Jong, P. F., & Share, D. L. (2007). Orthographic learning during oral and silent reading. *Scientific Studies of Reading, 11, 55-71.*

Deault, L., & Savage, R. S. (in press). Understanding and supporting children experiencing dyslexia and ADHD: The challenge of constructing models incorporating constitutional and classroom influences. In K. Littleton, C. Wood & J. K. Staarman (Eds.), *New perspectives on learning and teaching*. London: Emerald.

Denckla, M. B., & Rudel, R. G. (1976a). Naming of objects by dyslexic and other langauge-disabled children. *Brain and Language, 3*, 1-15.

Denckla, M. B., & Rudel, R. G. (1976b). Rapid "automatized" naming (R.A.N): Dyslexia differentiated from other learning disabilities. *Neuropsychologia, 14*, 471-479.

Denckla, M. B., Rudel, R. G., Chapman, C., & Krieger, J. (1985). Motor proficiency in dyslexic children with and without attentional disorders. *Archives of Neurology, 42*, 228-231.

Diener, H. C., Dichgans, J., Guschlbauer, B., Bacher, M., Rapp, H., & Klockgether, T. (1992). The coordination of posture and voluntary movement in patients with cerebellar dysfunction. *Movement Disorders, 7*, 14-22.

Diener, H. C., Hore, J., Ivry, R., & Dichgans, J. (1993). Cerebellar dysfunction of movement and perception. *Canadian Journal of Neurological Sciences, 20*, 62-69.

Dobkin, B. H. (2003). Organizational plasticity in sensorimotor and cognitive networks *The clinical science of neurologic rehabilitation* (2nd ed., pp. 3-75). New York: Oxford University Press.

Dow, R. S., & Moruzzi, G. (1958). *The physiology and pathology of the cerebellum*. Minneapolis: University of Minnesota Press.

Duncan, L. G., & Johnston, R. S. (1999). How does phonological awareness relate to nonword reading skill amongst poor readers? *Reading and Writing: An Interdisciplinary Journal, 11*, 405-439.

Dunn, L. M., & Dunn, L. M. (2007). *Examiner's manual for the PPVT-IV: Peabody Picture Vocabulary Test Fourth Edition*. Circle Pines, MN: American Guidance Service.

Durlak, J. A. (2009). How to select, calculate, and interpret effect sizes. *Journal of Pediatric Psychology, 34*, 917-928.

Eden, G. F., Jones, K. M., Cappell, K., Gareau, L., Wood, F. B., Zeffiro, T. A., et al. (2004). Neural changes following remediation in adult developmental dyslexia. *Neuron, 44*, 411-422.

Eden, G. F., & Moats, L. (2002). The role of neuroscience in the remediation of students with dyslexia. *Nature, 5*, 1080-1084.

Fawcett, A. J. (2002). Dyslexia, the cerebellum and phonological skill. In E. Witruk, A. D. Friederici & et al. (Eds.), *Basic functions of language, reading and reading disability: Neuropsychology and cognition* (pp. 265-279). Dordrecht, Netherlands: Kluwer Academic Publishers.

Fawcett, A. J., & Lynch, L. (2000). Systematic identification and intervention for reading difficulty: Case studies of children with EAL. *Dyslexia: The Journal of the British Dyslexia Association, 6*, 57-71.

Fawcett, A. J., & Nicolson, R. I. (1994). Naming speed in children with dyslexia. *Journal of Learning Disabilities, 27*(10), 641-646.

Fawcett, A. J., & Nicolson, R. I. (1995a). The dyslexia early screening test. *Irish Journal of Psychology, 16*, 248-259.

Fawcett, A. J., & Nicolson, R. I. (1995b). Persistence of phonological awareness deficits in older children with dyslexia. *Reading and Writing: An Interdisciplinary Journal, 7*, 361-376.

Fawcett, A. J., & Nicolson, R. I. (1995c). Persistent deficits in motor skill of children with dyslexia. *Journal of Motor Behavior, 27*, 235-240.

Fawcett, A. J., & Nicolson, R. I. (1996). *The dyslexia screening test: Manual*. London: Harcourt Brace.

Fawcett, A. J., & Nicolson, R. I. (1999). Performance of dyslexic children on cerebellar and cognitive tests. *Journal of Motor Behavior, 31*, 68-78.

Fawcett, A. J., & Nicolson, R. I. (2002). Children with dyslexia are slow to articulate a single speech gesture. *Dyslexia, 8*, 189-203.

Fawcett, A. J., Nicolson, R. I., & Dean, P. (1996). Impaired performance of children with dyslexia on a range of cerebellar tasks. *Annals of Dyslexia, 46*, 259-283.

Fawcett, A. J., Nicolson, R. I., & Maclagan, F. (2001). Cerebellar tests differentiate between groups of poor readers with and without IQ discrepancy. *Journal of Learning Disabilities, 34*, 119-135.

Felton, R. H., & Wood, F. B. (1992). A reading level match study of nonword reading skills in poor readers with varying IQ. *Journal of Learning Disabilities, 25*, 318-326.

Fiez, J. A., Petersen, S. E., Cheney, M. K., & Raichle, M. E. (1992). Impaired non-motor learning and error detection associated with cerebellar damage: A single case study. *Brain, 115*, 155-178.

Fletcher, J. M., Foorman, B. R., Boudousquie, A., Barnes, M. A., Schatschneider, C., & Francis, D. J. (2002). Assessment of reading and learning disabilities: A research-based intervention-oriented approach. *Journal of School Psychology, 40*, 27-63.

Fletcher, J. M., Francis, D. J., Rourke, B. P., Shaywitz, S. E., & Shaywitz, B. A. (1992). The validity of discrepancy-based definitions of reading disabilities. *Journal of Learning Disabilities, 25*, 555-561.

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

Fowler, A. E. (1991). How early phonological development might set the stage for phoneme awareness. In S. A. Brady & D. P. Shankweiler (Eds.), *Phonological processes in literacy: A tribute to Isabelle Y Liberman* (pp. 97-117). Hillsdale, NJ: Lawrence Erlbaum.

Fowler, E. G., Nwigwe, A., & Wong Ho, T. (2000). Sensitivity of the pendulum test for assessing spasticity in persons with cerebral palsy. *Developmental Medicine & Child Neurology, 42*, 182-189.

Frank, J., & Levinson, H. (1973). Dysmetric dyslexia and dyspraxia: Hypothesis and study. *Journal of the American Academy of Child Psychiatry, 12*, 690-701.

French, M. P., Opatrny, C., & Cochran, L. (2008). The power of phonological awareness as a predictor of basic reading skill. Retrieved June 30, 2008, from http://www.americanreadingforum.org/Yearbooks/01_yearbook/pdf/08_French.pdf

Frith, U. (1997). Brain, mind and behavior in dyslexia. In C. Hulme & M. Snowling (Eds.), *Dyslexia: Biology, cognition and intervention*. London: Whurr.

Georgiou, G. K., Parrila, R., Kirby, J. R., & Stephenson, K. (2008). Rapid naming components and their relationship with phonological awareness, orthographic knowledge, speed of processing, and different reading outcomes. *Scientific Studies in Reading, 12*, 325-350.

Gillon, G., & Dodd, B. J. (1994). A prospective study of relationship between phonological, semantic, and syntactic skills and specific reading disability. *Reading and Writing: An Interdisciplinary Journal, 6*, 321-345.

Gillon, G. T. (2004). *Phonological awareness: From research to practice*. New York: Guilford Press.

Gonzalez, J. E. J. (1997). A reading-level match study of phonemic processes underlying reading disabilities in a transparent orthography. *Reading and Writing: An Interdisciplinary Journal, 9*, 23-40.

Goswami, U. (2006). Sensorimotor impairments in dyslexia: Getting the beat. *Developmental Science, 9*, 257-259.

Goswami, U., & Bryant, P. (1989). The interpretation of studies using the reading level design. *Journal of Reading Behavior, 4*, 413-424.

Grainger, J., Bouttevin, S., Truc, C., Bastien, M., & Ziegler, J. (2003). Word superiority, pseudoword superiority, and learning to read: A comparison of dyslexic and normal readers. *Brain & Language, 87*, 432-440.

Granacher, R. P. (2003). *Traumatic brain injury: Methods for clinical and forensic neuropsychiatric assessment*. Boca Raton, Fl: CRC Press.

Gustafson, S., & Samuelsson, S. (1999). Intelligence and dyslexia: Implications for diagnosis and intervention. *Scandinavian Journal of Psychology, 40*, 127-134.

Haggard, P., Miall, R. C., Wade, D., Fowler, S., Richardson, A., Anslow, P., et al. (1995). Damage to cerebellocortical pathways after closed head injury: A behavioural and magnetic resonance imaging study. *Journal of Neurology, Neurosurgery & Psychiatry, 58*, 433-438.

Hammill, D. L. (2004). What we know about correlates of reading. *Exceptional Children, 70*, 453-469.

Hatcher, P. J., Hulme, C., & Ellis, A. W. (1994). Ameliorating early reading failure by integrating the teaching of reading and phonological skills: The phonological linkage hypothesis. *Child Development, 65*, 41-57.

Herrmann, J. A., Matyas, T., & Pratt, C. (2006). Meta-analysis of the nonword reading deficit in specific reading disorder. *Dyslexia: An International Journal of Research and Practice, 12, 195-221*.

Holmes, G. (1917). The symptoms of acute cerebellar injuries due to gunshot injuries. *Brain, 40*, 461-535.

Holmes, G. (1939). The cerebellum of man. *Brain, 62*, 1-30.

Howell, D. C. (2010). *Statistical methods for psychology* (7th ed.). Belmont, CA: Cengage Wadsworth.

Howitt, D., & Cramer, D. (2008). *Introduction to SPSS in psychology: For version 16 and earlier*. England: Prentice Hall.

International Dyslexia Association. (2002). Promoting literacy through research, education, and advocacy. Report retrieved February 10, 2010, from http://www.interdys.org/FAQWhatIs.htm

Irannejad, S., & Savage, R. (2009). The cerebellar deficit theory of developmental dyslexia: Evidence and implications for intervention. In C. Wood & V. Connelly (Eds.), *Contemporary perspectives on reading and spelling* (pp. 254-270). London: Oxford Brookes University.

Ito, M. (1984). *The cerebellum and neural control*. New York: Raven Press.

Ito, M. (1993). Movement and thought: Identical control mechanisms by the cerebellum. *Trends in Neurosciences, 16*, 448-450.

Ito, M. (2002). Historical review of the significance of the cerebellum and the role of Purkinje cells in motor learning. *Annals of the New York Academy of Sciences, 978*, 273-288.

Ivry, R. B., & Gopal, H. S. (1993). Speech production and perception in patients with cerebellar lesions. In D. E. Meyer & S. Kornblum (Eds.), *Attention and performance XIV: Synergies in experimental psychology, artificial intelligence, and cognitive neuroscience* (pp. 771-802). Cambridge, MA: The MIT Press.

Ivry, R. B., & Keele, S. W. (1989). Timing functions of the cerebellum. *Journal of Cognitive Neuroscience, 1*, 136-152.

Jackson, N. E., & Butterfield, E. C. (1989). Reading-level match designs: Myths and realities. *Journal of Reading Behavior, 4*, 387-412.

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

Kasselimis, D. S., Margarity, M., & Vlachos, F. (2008). Cerebellar function, dyslexia and articulation speed. *Child Neuropsychology, 14*, 303-313.

Katzir, T., Kim, Y., Wolf, M., O'Brien, B., Kennedy, B., Lovett, M., et al. (2006). Reading Fluency: The whole is more than the parts. *Annals of Dyslexia, 56*, 51-82.

Kelly, S. W., Griffiths, S., & Frith, U. (2002). Evidence for implicit sequence learning in dyslexia. *Dyslexia, 8*, 43-52.

Kingsley, R. E. (2000). *Concise text of neuroscience* (2nd ed.). Philadelphia, PA: Lippincott Williams & Wilkins.

Kirby, J. R., Desrochers, A., Roth, L., & Lai, S. S. V. (2008). Longitudinal predictors of word reading development. *Canadian Psychology, 49*, 103-110.

Korhonen, T. T. (1991). Neuropsychological stability and prognosis of subgroups of children with learning disabilities. *Journal of Learning Disabilities, 24*, 48-57.

Korhonen, T. T. (1995). The persistence of rapid naming problems in children with reading disabilities: A nine-year follow-up. *Journal of Learning Disabilities, 28*, 232-239.

Kozey, M., & Siegel, L. S. (2008). Definitions of learning disabilities in Canadian provinces and territories. *Canadian Psychology, 49*, 162-171.

Lachmann, T. (2002). Reading disability as a deficit in functional coordination. In E. Witruk, A. D. Friederici & et al. (Eds.), *Basic functions of language, reading and reading disability: Neuropsychology and cognition* (pp. 165-198). Dordrecht, Netherlands: Kluwer Academic Publishers.

Lang, C. E., & Bastian, A. J. (2002). Cerebellar damage impairs automaticity of a recently practiced movement. *Journal of Neurophysiology, 87*, 1336-1347.

Leiner, H. C., Leiner, A. L., & Dow, R. S. (1989). Reappraising the cerebellum: what does the hindbrain contribute to the forebrain? *Behavioral Neurosciences, 103*, 998-1008.

Leiner, H. C., Leiner, A. L., & Dow, R. S. (1993). Cognitive and language functions of the human cerebellum. *Trends in Neurosciences, 16*, 444-447.

Leslie, G. C., Muir, C., Part, N. J., & Roberts, R. C. (1992). A comparison of the assessment of spasticity by the Wartenberg pendulum test and the Ashworth grading scale in patients with multiple sclerosis. *Clinical Rehabilitation, 6*, 41-48.

Levinson, H. N. (1988). The cerebellar-vestibular basis of learning disabilities in children, adolescents and adults: Hypothesis and study. *Perceptual & Motor Skills, 67*, 983-1006.

Liberman, A. M. (1982). On finding that speech is special. *American Psychologist, 37*, 148-167.

Liberman, A. M., & Mattingly, I. G. (1985). The motor theory of speech perception revised. *Cognition, 21*, 1-36.

Liberman, I. Y., & Shankweiler, D. (1991). Phonology and beginning: A tutorial. In L. Rieben & C. A. Perfetti (Eds.), *Learning to read: Basic research and its implications* (pp. 3-17). Hillsdale, NJ: Erlbuam.

Locke, J. L. (1983). *Phonological acquistion and change*. New York: Academic Press.

Lopez, M. R., & Gonzalez, J. E. (2000). IQ vs phonological recoding skill in explaining differences between poor readers and normal readers in word recognition: Evidence from a naming task. *Reading & Writing, 12*, 129-142.

Lundberg, I. (1999). Towards a sharper definition of dyslexia. In I. Lundberg, F. E. Toennessen & I. Austad (Eds.), *Dyslexia: Advances in theory and practice* (pp. 9-29). Dordrecht, Netherlands: Academic Publisher.

Lundberg, I., Frost, J., & Petersen, O. (1988). Effects of an extensive program for stimulating phonological awareness in preschool children. *Reading Research Quarterly, 23*, 263-284.

Lundberg, I., & Hoien, T. (1990). Patterns of information processing skills and word recogntion strategies in developmental dyslexia. *Scandinavian Journal of Educational Research, 34*, 231-240.

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

Lyon, G. R., Shaywitz, S. E., & Shaywitz, B. A. (2003). A definition of dyslexia. *Annals of Dyslexia, 53*, 1-14.

Manis, F. R., Doi, L. M., & Bhadha, B. (2000). Naming speed, phonological awareness, and orthographic knowledge in second graders. *Journal of Learning Disabilities, 33*, 325-333.

McPhillips, M. (2003). A commentary on an article published in the February 2003 edition of 'Dyslexia', 'Evaluation of an exercise-based treatment for children with reading difficulties' (Reynolds, Nicolson, & Hambly). *Dyslexia: An International Journal of Research & Practice, 9*, 161-163.

Medland, S. E., Geffen, G., & McFarland, K. (2002). Lateralization of speech production using verbal/manual dual tasks: Meta-analysis of sex differences and practice effects. *Neuropsychologia, 40*, 1233-1239.

Messaoud-Galusi, S., & Marshall, C. R. (2010). Exploring the overlap between dyslexia and SLI: The role of phonology. *Scientific Studies in Reading, 14*, 1-7.

Metz, U., Marx, P., Weber, J., & Scheider, W. (2003). Overachievement in reading and spelling: Consequences for discrepancy definition of dyslexia. *Zeitschrift fur Entwicklungspsychologie und Padagogische Psychologie, 35*, 127-134.

Meyer, M. S., Wood, F. B., Hart, L. A., & Felton, R. H. (1998). Selective predictive value of rapid automatized naming in poor readers. *Journal of Learning Disabilities, 31*, 106-117.

Miall, R. C., & Christensen, L. O. (2004). The effect of rTMS over the cerebellum in normal human volunteers on peg-board movement performance. *Neuroscience Letters, 371*, 185-189.

Ministère de l'Éducation, du Loisir et du Sport. (2009). Instruction in English in Québec. Retrieved November 20, 2009, from http://www.mels.gouv.qc.ca/daasa/rens/banque/Fiches/F95a.htm#03

Moe-Nilssen, R., Helbostad, J. L., Talcott, J. B., & Toennessen, F. E. (2003). Balance and gait in children with dyslexia. *Experimental Brain Research, 150*, 237-244.

Moher, D., Schulz, K. F., & Altman, D. G. (2001). The CONSORT statement: Revised recommendations for improving the quality of reports of parallel-group randomized trials. *Lancet, 357*, 1191-1194.

Muter, V., Hulme, C., Snowling, M., & Taylor, S. (1997). Segmentation, not rhyming, predicts early progress in learning to read. *Journal of Experimental Child Psychology, 65*, 376-396.

Muter, V., & Snowling, M. (1998). Concurrent and longitudinal predictors of reading: The role of metalinguistic and short-term memory skills. *Reading Research Quarterly, 33*, 320-337.

Naslund, J. C., & Schneider, W. (1996). Kindergarten letter knowledge, phonological skills, and memory processes: Relative effects on early literacy. *Journal of Experimental Child Psychology, 62*, 30-59.

National Association of School Psychologists. (2007). NASP position statement on identification of students with specific learning disabilities. Report retrieved May 10, 2010, from http://www.caspsurveys.org/NEW/pdfs/nasp12.pdf

National Reading Panel. (2000). Teaching children to read: Reports of the subgroups. Report retrieved December, 19, 2008, from http://www.nichd.nih.gov/publications/nrp/report.htm

Needle, J. L., Fawcett, A. J., & Nicolson, R. I. (2006). Balance and dyslexia: An investigation of adults' abilities. *European Journal of Cognitive Psychology, 18, 909-936.*

Neuhaus, G., Foorman, B. R., Francis, D. J., & Carlson, C. D. (2001). Measures of information processing in rapid automatized naming (RAN) and their relation to reading. *Journal of Experimental Child Psychology, 78*, 359-373.

Neuhaus, G. F., Carlson, C. D., Jeng, W. M., Post, Y., & Swank, P. R. (2001). The reliability and validity of rapid automatized naming scoring software ratings for the determination of pause and articulation component durations. *Educational & Psychological Measurement, 61*, 490-504.

Neuhaus, G. F., & Swank, P. R. (2002). Understanding the relations between RAN letter subtest components and word reading in first-grade students. *Journal of Learning Disabilities, 35*, 158-174.

Newton, R. R., & Rudestam, K. E. (1999). *Your statistical consultant: Answers to your data analysis questions*. Thousand Oaks, CA: SAGE Publications.

Nicholls, M. E. R. (1996). Temporal processing asymmetries between the cerebral hemispheres: Evidence and implications. *Laterality: Asymmetries of Body, Brain & Cognition, 1*, 97-137.

Nicolson, R. I., Daum, I., Schugens, M. M., Fawcett, A. J., & Schulz, A. (2002). Eyeblink conditioning indicate cerebellar abnormality in dyslexia. *Experimental Brain Research, 143*, 42-50.

Nicolson, R. I., & Fawcett, A. J. (1990). Automaticity: A new framework for dyslexia research? *Cognition, 35*, 159-182.

Nicolson, R. I., & Fawcett, A. J. (1994). Comparison of deficits in cognitive and motor skills among children with dyslexia. *Annals of Dyslexia, 44*, 147-164.

Nicolson, R. I., & Fawcett, A. J. (1995). Dyslexia is more than a phonological disability. *Dyslexia, 1*, 19-36.

Nicolson, R. I., & Fawcett, A. J. (1998). *The dyslexia adult screening test*. London: Psychological Corporation.

Nicolson, R. I., & Fawcett, A. J. (1999). Developmental dyslexia: The role of the cerebellum. *Dyslexia: The Journal of the British Dyslexia Association, 5*, 155-177.

Nicolson, R. I., & Fawcett, A. J. (2000). Long-term learning in dyslexic children. *European Journal of Cognitive Psychology, 12*, 357-393.

Nicolson, R. I., & Fawcett, A. J. (2006). Do cerebellar deficits underlie phonological problems in dyslexia? *Developmental Science, 9, 259-262.*

Nicolson, R. I., Fawcett, A. J., & Dean, P. (1995). Time estimation deficits in developmental dyslexia: Evidence of cerebellar involvement. *Proceedings of the Royal Society of London, 259*, 43-47.

Nicolson, R. I., Fawcett, A. J., & Dean, P. (2001). Developmental dyslexia: The cerebellar deficit hypothesis. *Trends in Neurosciences, 24*, 508-511.

Nordmark, E., & Anderson, G. (2002). Wartenberg pendulum test: Objective quantification of muscle tone in children with spastic diplegia undergoing selective dorsal rhizotomy. *Developmental Medicine & Child Neurology, 44*, 26-33.

O'Connor, R. E., Fulmer, D., Harty, K. R., & Bell, K. (2005). Layers of reading intervention in kindergarten through third grade: Changes in teaching and student outcomes. *Journal of Learning Disabilities, 38*, 440-455.

Office Québécoise de la Langue Française. (2008). The charter of French language. Report retrieved December 15, 2008, from http://www.olf.gouv.qc.ca/english/charter/title1chapter5.html

Olejnik, S., & Algina, J. (2000). Measures of effect size for comparative studies: Applications, interpretations, and limitations. *Contemporary Educational Psychology, 25*, 241-286.

Olejnik, S. F., & Algina, J. (1984). Parametric ANCOVA and the rank transform ANCOVA when the data are conditionally non-normal and heteroscedastic. *Journal of Educational Statistics, 9*, 129-149.

Olejnik, S. F., & Algina, J. (1985). Nonparametric alternatives to analysis of covariance. *Evaluation review, 9*, 51-83.

Olson, R., Wise, B., Conners, F., Rack, J., & Fulker, D. W. (1989). Specific deficits in component reading and language skills: Genetic and environmental influences. *Journal of Learning Disabilities, 22*, 339-348.

Orton Dyslexia Society. (1995). Definition of dyslexia: Report from committee of members. *Perspectives, 21*, 16-17.

Palacios, E. D., & Semrud-Clikeman, M. (2005). Delinquency, hyperactivity, and phonological awareness: A comparison of adolescents with ODD and ADHD. *Applied Neuropsychology, 12*, 94-105.

Palliyath, S., Hallett, M., Thomas, S. L., & Lebiedowska, M. K. (1998). Gait in patients with cerebellar ataxia. *Movement Disorders, 13*, 958-964.

Pennington, B. F., Cardoso-Martins, C., Green, P. A., & Lefly, D. L. (2001). Comparing the phonological and double deficit hypothesis for developmental dyslexia. *Reading and Writing: An Interdisciplinary Journal, 14*, 707-755.

Pennington, B. F., Groisser, D., & Welsh, M. (1993). Contrasting cognitive deficits in attention deficit hyperactivity disorder versus reading disability. *Developmental Psychology, 29*, 511-523.

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

Pennington, B. F., Van Orden, G. C., Smith, S. D., Green, P. A., & Haith, M. M. (1990). Phonological processing skills and deficits in adult dyslexics. *Child Development, 61*, 1753-1778.

Plaza, M. (2003). The role of naming speed, phonological processing and morphological/syntactic skill in reading and spelling performance of second-grade children. *Special Issue on Language Disorders and Reading Acquisition, 1*, 1-7.

Plaza, M., & Cohen, H. (2003). The interaction between phonological syntactic awareness, and naming speed in the reading and spelling performance of first grade children. *Brain and Cognition, 53*, 287-292.

Pope, D. J., & Whiteley, H. E. (2003). Developmental dyslexia, cerebellar/vestibular brain function and possible links to exercise-based interventions: A review. *European Journal of Special Needs Education, 18*, 109-123.

Powell, D., Stainthorp, R., Stuart, M., Garwood, H., & Quinlan, P. (2007). An experimental comparison between rival theories of rapid automatized naming performance and its relationship to reading. *Journal of Experimental Child Psychology, 98*, 46-68.

Psychological Corp. (1999). *Wechsler abbreviate scale of intelligence*. San Antonio, TX: Psychological Corporation.

Raberger, T., & Wimmer, H. (1999). Is poor reading caused by an automatisation deficit? *Zeitschrift fur Padagogische Psychologie, 13*, 74-83.

Raberger, T., & Wimmer, H. (2003). On the automaticity/cerebellar deficit hypothesis of dyslexia: Balancing and continuous rapid naming in dyslexic and ADHD children. *Neuropsychologia, 41*, 1493-1497.

Rack, J., Hulme, C., Snowling, M., & Wightman, J. (1994). The role of phonology in young children learning to read words: The direct-mapping hypothesis. *Journal of Experimental Child Psychology, 57*, 42-71.

Rack, J. P., Snowling, M. J., & Olson, R. K. (1992). The nonword reading deficit in developmental dyslexia: A review. *Reading Research Quarterly, 27*, 28-53.

Ramus, F., Pidgeon, E., & Frith, U. (2003). The relationship between motor control and phonology in dyslexic children. *Journal of Child Psychology & Psychiatry & Allied Disciplines, 44*, 712-722.

Ramus, F., Rosen, S., Dakin, S. C., Day, B. L., Castellote, J. M., White, S., et al. (2003). Theories of developmental dyslexia: Insights from a multiple case study of dyslexic adults. *Brain, 126*, 841-865.

Ramus, F., White, S., & Frith, U. (2006). Weighing the evidence between competing theories of dyslexia. *Developmental Science, 9*, 265-269.

Reason, R., Frederickson, N., Hefferman, M., Martin, C., & Woods, K. (1999). *Dyslexia, literacy and psychological assessment*. Leicester, UK: British Psychological Society.

Reid, G. (2003). *Dyslexia: A practitioner's handbook*. West Sussex, England: John Wiley & Sons.

Reid, G., Fletcher, J. M., Shaywitz, S. E., Shaywitz, B. A., Torgesen, J. K., Wood, F. B., et al. (2001). Rethinking learning disabilities. In C. E. Finn Jr, A. J. Rotherham &

C. R. Hokanson Jr (Eds.), *Rethinking special education for a new century* (pp. 259-287). Washington: Progressive Policy Institute/The Thomas B. Fordham Foundation.

Reid, L. G. (1994). *Frames of reference for the assessment of learning disabilities: New views on measurement issues*. Baltimore: Brookes.

Reynolds, D., & Nicolson, R. I. (2007). Follow-up of an exercise-based treatment for children with reading difficulties. *Dyslexia, 13*, 78-96.

Reynolds, D., Nicolson, R. I., & Hambly, H. (2003). Evaluation of an exercise-based treatment for children with reading difficulties. *Dyslexia: The Journal of the British Dyslexia Association, 9*, 48-71.

Richards, T. L., Corina, D., Serafini, S., Steury, K., Echelard, D. R., Dager, S. R., et al. (2000). Effects of a phonologically driven treatment for dyslexia on lactate levels measured by proton MR spectroscopic imaging. *American Journal of Neuroradiology, 21*, 916-922.

Rochelle, K. S., & Talcott, J. B. (2006). Impaired balance in developmental dyslexia? A meta-analysis of the contending evidence. *Journal of Child Psychology and Psychiatry, 47, 1159-1166.*

Roessner, V., Banaschewski, T., Fillmer-Otte, A., Becker, A., Albrecht, B., Uebel, H., et al. (2008). Color perception deficits in co-existing attention-deficit/hyperactivity disorder and chronic tic disorders. *Journal of Neural Transmission, 115*, 235-239.

Rucklidge, J. J., & Tannock, R. (2002). Neuropsychological profiles of adolescents with ADHD: Effects of reading difficulties and gender. *Journal of Child Psychology & Psychiatry, 43*, 988-1003.

Rudel, R. G., Denckla, M. B., & Broman, M. (1978). Rapid silent response to repeated target symbols by dyslexic and nondyslexic children. *Brain & Language, 6*, 52-62.

Saklofske, D. H., Caravan, G., & Schwartz, C. (2000). Concurrent validity of the Wechsler Abbreviated Scale of Intelligence (WASI) with a sample of Canadian children. *Canadian Journal of School Psychology, 16*, 87-94.

Savage, R. (2004). Motor skills, automaticity and developmental dyslexia: A review of the research literature. *Reading & Writing, 17*, 301-324.

Savage, R. (2007). Cerebellar tasks do not distinguish between children with developmental dyslexia and children with intellectual disability. *Child Neuropsychology, 13*, 389-407.

Savage, R., & Carless, S. (2005). Phoneme manipulation not onset-rime manipulation ability is a unique predictor of early reading. *Journal of Child Psychology and Psychiatry, 46*, 1297-1308.

Savage, R., Carless, S., & Ferraro, V. (2007). Predicting curriculum and test performance at age 11 from pupil background, baseline skills and phonological awareness at age 5. *Journal of Child Psychology and Psychiatry, 48*, 732-739.

Savage, R., & Frederickson, N. (2006). Beyond Phonology: What else is needed to describe the problems of below-average readers and spellers? *Journal of Learning Disabilities, 39*, 399-413.

Savage, R., Frederickson, N., Goodwin, R., Patni, U., Smith, N., & Tuersley, L. (2005a). Evaluating current deficit theories of poor reading: Role of phonological processing, naming speed, balance automaticity, rapid verbal perception and working memory. *Perceptual & Motor Skills, 101*, 345-361.

Savage, R., Frederickson, N., Goodwin, R., Patni, U., Smith, N., & Tuersley, L. (2005b). Relationships among rapid digit naming, phonological processing, motor automaticity, and speech perception in poor, average, and good readers and spellers. *Journal of Learning Disabilities, 38*, 12-28.

Savage, R., Lavers, N., & Pillay, V. (2007). Working memory and reading difficulties: What we know and what we don't know about the relationshiop. *Educational Psycholgy Review, 19*, 185-221.

Savage, R., Pillay, V., & Melidona, S. (2007). Deconstructing rapid automatized naming: Component processes and the prediction of reading difficulties. *Learning and Individual Differences, 17*, 129-146.

Savage, R., & Pompey, Y. (2008). What does the evidence say about effective literacy teaching? *Educational and Child Psychology, 25*, 17-26.

Savage, R. S., Abrami, P., Hipps, G., & Deault, L. (2009). A randomized controlled trial study of the ABRACADABRA reading intervention program in grade 1. *Journal of Educational Psychology, 101*, 590-604.

Scanlon, D. M., Vellutino, F. R., Small, S. G., Fanuele, D. P., & Sweeny, J. M. (2005). Severe reading difficulties - can they be prevented? A comparison of prevention and intervention approaches. *Exceptionality, 13*, 209-227.

Schmahmann, J. D., & Sherman, J. C. (1998). The cerebellar cognitive affective syndrome. *Brain, 121*, 561-579.

Schulz, G. M., Dingwall, W. O., & Ludlow, C. L. (1999). Speech and oral motor learning in individuals with cerebellar atrophy. *Journal of Speech, Language, & Hearing Research, 42*, 1157-1175.

Semrud-Clikeman, M., Guy, K., & Griffin, J. (2000). Rapid naming deficits in children and adolescents with reading disabilities and attention deficit hyperactivity disorder. *Brain & Language, 74*, 70-83.

Share, D. L. (1995). Phonological recoding and self-teaching: Sine qua non of reading acquisition. *Cognition, 55, 151-218.*

Share, D. L., Jorm, A. F., Maclean, R., & Matthews, R. (1984). Sources of individual differences in reading acquisition. *Journal of Educational Psychology, 76*, 1309-1324.

Shaywitz, B. A., Shaywitz, S. E., Blachman, B. A., Pugh, K. R., Fulbright, R. K., Skudlarski, P., et al. (2004). Development of left occipitotemporal systems for skilled reading in Children after a phonologically-based intervention. *Biological Psychiatry, 55*, 926-933.

Shaywitz, S. E., Morris, R., & Shaywitz, B. A. (2008). The education of dyslexic children from childhood to young adulthood. *Annual Review of Psychology, 59*, 451-475.

Shipley, R. E., & Harley, R. J. (1971). A device for estimating stability of stance in human subjects. *Psychophysiology, 7*, 287-292.

Shumway-Cook, A., & Woollacott, M. H. (2001). *Motor control: Theory and practical applications* (2nd ed.). Baltimore, Maryland: Lippincott Williams & Wilkins.

Siegel, L. S. (1988). Evidence that IQ scores are irrelevant to the definition and analysis of reading disability. *Canadian Journal of Psychology, 42*, 201-215.

Siegel, L. S. (1989). Why we do not need intelligence and analyses of learning disabilities. *Journal of Learning Disabilities, 8*, 514-518.

Siegel, L. S. (1992). An evaluation of the discrepancy definition of dyslexia. *Journal of Learning Disabilities, 25*, 618-629.

Siegel, L. S. (1993). Phonological processing deficits as the basis of a reading disability. *Developmental Review, 13*, 246-257.

Siegel, L. S. (2005). Reflections on research on reading disability with special attention to gender issues. *Journal of Learning Disabilities, 38*, 473-477.

Siegel, L. S., & Ryan, E. B. (1988). Development of grammatical-sensitivity, phonological, and short-term memory skills in normally achieving and learning disabled children. *Developmental Psychology, 24*, 28-37.

Siegel, L. S., & Ryan, E. B. (1989). Subtypes of developmental dyslexia: The influence of definitional variables. *Reading & Writing, 1*, 257-287.

Silveri, M. C., Leggio, M. G., & Molinari, M. (1994). The cerebellum contributes to linguistic production: A case of agrammatic speech following a right cerebellar lesion. *Neurology, 44*, 2047-2050.

Simmons, D. C., Coyne, M. D., Kwok, O., McDonagh, S., Harn, B. A., & Kame'enui, E. J. (2008). A longitudinal study of reading risk from kindergarten through third grade. *Journal of Learning Disabilities, 41*, 158-173.

Simos, P. G., Fletcher, J. M., Bergman, E., Breier, J. I., Foorman, B. R., Castillo, E. M., et al. (2002). Dyslexia-specific brain activation profile becomes normal following successful remedial training. *Neurology, 58*, 1203-1213.

Snowling, M. (1998). Dyslexia as a phonological deficit: Evidence and implications. *Child Psychology & Psychiatry Review, 3*, 4-11.

Snowling, M., & Hulme, C. (1994). The development of phonological skills. *Philosophical Transactions of the Royal Society of London, 346*, 21-27.

Snowling, M. J. (1981). Phonemic deficits in developmental dyslexia. *Psychological Research, 43*, 219-234.

Snowling, M. J. (1995). Phonological processing and developmental dyslexia. *Journal of Research in Reading, 18*, 132-138.

Snowling, M. J. (2006). Language skills and learning to read: The dyslexia spectrum. In M. J. Snowling & J. Stackhouse (Eds.), *Dyslexia speech and language* (2nd ed., pp. 1-14). England: Whurr Publishers.

Snyder, L. S., & Downey, D. M. (1995). Serial rapid naming skills in children with reading disabilities. *Annals of Dyslexia, 45, 31-49.*

Stanovich, K. E. (1993). A model for studies of reading disability. *Developmental Review, 13*, 225-245.

Stanovich, K. E. (1996). Towards a more inclusive definition of dyslexia. *Dyslexia, 2*, 154-166.

Stanovich, K. E. (1998). Refining the phonological core deficit model. *Child Psychology & Psychiatry Review, 3*, 17-21.

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.

Stillman, B., Phty, D., McMeeken, J., & Phty, D. (1995). A video-based version of the pendulum test: Technique and normal response. *Archives of Physical Medicine and Rehabilitation, 76*, 166-176.

Stoodley, C. J., Fawcett, A. J., Nicolson, R. I., & Stein, J. F. (2006). Balancing and pointing tasks in dyslexic and control adults. *Dyslexia, 12*, 276-288.

Stoodley, C. J., Harrison, E. P., & Stein, J. F. (2006). Implicit motor learning deficits in dyslexic adults. *Neuropsychologia Vol 44(5) 2006, 795-798.*

Strauss, E., Sherman, E. M. S., & Spreen, O. (2006). *A compendium of neuropsychological tests: Administration, norms, and commentary* (3rd ed.). New York: Oxford Press.

Stuebing, K. K., Fletcher, J. M., LeDoux, J. M., Lyon, G., Shaywitz, S. E., & Shaywitz, B. A. (2002). Validity of IQ-discrepancy classifications of reading disabilities: A meta-analysis. *American Educational Research Journal, 39*, 469-518.

Sunseth, K., & Bowers, P. G. (2002). Rapid naming and phonemic awareness: Contributions to reading, spelling, and orthographic knowledge. *Scientific Studies in Reading, 6*, 401-429.

Swanson, H. L., Trainin, G., Necoechea, D. M., & Hammill, D. L. (2003). Rapid naming, phonological awareness, and reading: A meta-analysis of the correlation evidence. *Review of Educational Research, 73*, 407-440.

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5th ed.). New York: Pearson Education.

Tallal, P. (2006). What happens when dyslexic subjects do not meet the criteria for dyslexia and sensorimotor tasks are too difficult even for the controls? *Developmental Science, 9*, 262-264.

Tannock, R., Banaschewski, T., & Gold, D. (2006). Color naming deficits and attention-deficit/hyperactivity disorder: A retinal dopaminergic hypothesis. *Behavioral and Brain Functions, 2*, Retrieved December 18, 2008, from http://www.behaviorandbrainfunction.com/content/2/1/4

Thalheimer, W., & Cook, S. (2002). How to calculate effect sizes from published research articles: A simplified methodology. Retrieved February 16, 2010 from http://work-learning.com/effect_sizes.htm

Thompson, B. (2002). Statistical, practical, and clinical: How many kinds of significance do counselors need to consider? *Journal of Counseling & Development, 80*, 64-71.

Thompson, G. B., & Johnston, R. S. (2000). Are nonword and other phonological deficit indicative of a failed reading process? *Reading and Writing: An Interdisciplinary Journal, 12*, 63-97.

Torgesen, J. K. (2002). The prevention of reading difficulties. *Journal of School Psychology, 40*, 7-26.

Torgesen, J. K., & Wagner, R. K. (1998). Alternative diagnostic approaches for specific developmental reading disabilities. *Learning Disabilities Research & Practice, 13*, 220-232.

Torgesen, J. K., Wagner, R. K., Rashotte, C. A., Rose, E., Lindamood, P., Conway, T., et al. (1999). Preventing reading failure in young children with phonological processing disabilities: Group and individual responses to instruction. *Journal of Educational Psychology, 91*, 1-15.

Valle, M. S., Casabona, A., Sgarlata, R., Garozzo, R., Vinci, M., & Cioni, M. (2006). *BMC Musculoskeletal Disorders, 7*, Retrieved October 12, 2008, from http://www.biomedcentral.com/1471-2474/89

van IJzendoorn, M. H., & Bus, A. G. (1994). Meta-analytic confirmation of the nonword reading deficit in developmental dyslexia. *Reading Research Quarterly, 29*, 267-275.

Vellutino, F. R., Fletcher, J. M., Snowling, M. J., & Scanlon, D. M. (2004). Specific reading disability (dyslexia): what have we learned in the past four decades? *Journal of Child Psychology and Psychiatry, 45, 2-40.*

Vellutino, F. R., Scanlon, D. M., & Lyon, G. (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.

Vellutino, F. R., Scanlon, D. M., Sipay, E. R., Small, S. G., Pratt, A., Chen, R. S., et al. (1996). Cognitive profiles of difficult to remediate and readily remediated poor readers: Early intervention as a vehicle for distinguishing between cognitive and experiential deficits as basic causes of specfic reading disability. *Journal of Educational Psychology, 88*, 601-638.

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

Vellutino, F. R., Scanlon, D. M., & Tanzman, M. S. (1998). The case for early intervention in diagnosing specific reading disability. *Journal of Learning Disabilities, 36*, 367-397.

Vellutino, F. R., Scanlon, D. M., Zhang, H., & Schatschneider, C. (2008). Using response to kindergarten and first grade intervention to identify children at-risk for long-term reading difficulties. *Journal of Reading and Writing, 21*, 437-480.

Vicari, S., Marotta, L., Menghini, D., Molinari, M., & Petrosini, L. (2003). Implicit learning deficit in children with developmental dyslexia. *Neuropsychologia, 41*, 108-114.

Vukovic, R. K., & Siegel, L. S. (2006). The double-deficit hypothesis: A comprehensive analysis of the evidence. *Journal of Learning Disabilities, 39*, 25-47.

Waber, D. P., Marcus, D. J., Forbes, P. W., Bellinger, D. C., Weiler, M. D., Sorensen, L. G., et al. (2003). Motor sequence learning and reading ability: Is poor reading associated with sequencing deficits? *Journal of Experimental Child Psychology, 84*, 338-354.

Wagner, R. K., & Torgesen, J. K. (1987). The nature of phonological processing and its causal role in the acquisition of reading skills. *Psychological Bulletin, 101*, 192-212.

Wagner, R. K., Torgesen, J. K., & Rashotte, C. A. (1994). Development of reading-related phonological processing abilities: New evidence of bi-directional causality from a latent variable longitudinal study. *Developmental Psychology, 30,* 73-87.

Wagner, R. K., Torgesen, J. K., & Rashotte, C. A. (1999). *Comprehensive test of phonological processing*. Austin, TX: PRO-ED.

Walker, H. K. (1990). The cerebellum. In H. K. Walker, W. D. Hall & J. W. Hurst (Eds.), *Clinical methods: The history, physical, and laboratory examinations.* Retrieved November 19, 2008 from http://www.ncbi.nlm.nih.gov/bookshelf/br.fcgi?book=cm

Walker, J. E., & Norman, C. A. (2006). The neurophysiology of dyslexia: A selective review with implications for neurofeedback remediation and results of treatment in twelve consecutive patients. *Journal of Neurotherapy Vol, 1,* 45-55.

Wartenberg, R. (1951). Pendulousness of the legs as a diagnostic test. *Neurology, 1*, 18-24.

Weber, J. M., Marx, P., & Schneider, W. (2002). Profitieren Legastheniker und allgemein lese-rechtschreibschwache Kinder in unterschiedlichem Ausmass von einem Rechtschreibtraining? [Are there different remedial potentials for dyslexics and garden-variety poor readers?]. *Psychologie in Erziehung und Unterricht, 49*, 56-70.

Wesseling, R., & Reitsma, P. (2001). Preschool phonological representations and development of reading skills. *Annals of Dyslexia, 51*, 203-229.

White, S., Milne, E., Rosen, S., Hansen, P., Swettenham, J., Frith, U., et al. (2006). The role of sensorimotor impairments in dyslexia: A multiple case study of dyslexic children. *Developmental Science, 9*, 237-255.

Wilkinson, G. S. (1993). *Wide range achievement test 3*. Wilmington, DE: Wide Range, Inc.

Willcutt, E. G., & Pennington, B. F. (2008). Comorbidity of reading disability and attention deficit hyperactivity disorder: Difference by gender and subtype. *Journal of Learning Disabilities, 33*, 179-191.

Wimmer, H., Mayringer, H., & Raberger, T. (1999). Reading and dual-task balancing: Evidence against the automatization deficit explanation of developmental dyslexia. *Journal of Learning Disabilities, 32*, 473-478.

Wolf, M. (1982). The word-retrieval process and reading in children and aphasics. In K. Nelson (Ed.), *Children's language* (pp. 437-493). Hillsdale, NJ: Erlbaum.

Wolf, M. (1991). Naming speed and reading: The contribution of the cognitive neurosciences. *Reading Research Quarterly, 26*, 123-141.

Wolf, M. (1997). A provisional, integrative account of phonological and naming-speed deficits in dyslexia: Implications for diagnosis and intervention. In B. A. Blachman (Ed.), *Foundations of reading acquisition and dyslexia: Implications for early intervention* (pp. 67-92). Mahwah, NJ: Lawrence Erlbaum.

Wolf, M. (1999). What Time May Tell: Towards a new conceptualization of developmental dyslexia. *Annals of Dyslexia*, 3-28.

Wolf, M., Bally, H., & Morris, R. (1986). Automaticity, retrieval processes, and reading: A longitudinal study in average and impaired readers. *Child Development, 57*, 988-1000.

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

Wolf, M., Bowers, P. G., & Biddle, K. (2000). Naming-speed processes, timing, and reading: A conceptual review. *Journal of Learning Disabilities, 33*, 387-407.

Wolf, M., & O'Brien, B. (2006). From the sumerians to images of the reading brain: Insights for reading theory and intervention. In G. D. Rosen (Ed.), *The dyslexic brain* (pp. 5-19). Mahwaw, NJ: Lawrence Elrbaum Associates.

Wolf, M., O'Rourke, A. G., Gidney, C., Lovett, M., Cirino, P., & Morris, R. (2002). The second deficit: An investigation of the independence of phonological and naming-speed deficits in developmental dyslexia. *Reading and Writing: An Interdisciplinary Journal, 15*, 43-72.

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

Woodcock, R. W., & Johnson, M. B. (1989). *Woodcock-Johnson psycho-educational battery - revised (WJ-R)*. Allen, Texas: Developmental Learning Materials.

APPENDIX A

Certificate of Ethical Approval