Ann. of Dyslexia (2012) 62:22–52
DOI 10.1007/s11881-011-0060-2

Is a cerebellar deficit the underlying cause of reading disabilities?

Shahrzad Irannejad · Robert Savage

Received: 7 February 2011 / Accepted: 22 September 2011 / Published online: 13 December 2011

Abstract This study investigated whether children with dyslexia differed in their performance on reading, phonological, rapid naming, motor, and cerebellar-related tasks and automaticity measures compared to reading age (RA)-matched and chronological age (CA)-matched control groups. Participants were 51 children attending mainstream English elementary schools in Quebec. All participants completed measures of IQ, word and nonword reading fluency, elision, nonword decoding, rapid naming, bead threading, peg moving, toe tapping, postural stability, and muscle tone. Results from both group contrasts and analyses at the individual case level did not provide support for claims of motor–cerebellar involvement in either typical or atypical reading acquisition. Results were more consistent with a phonological core process account of both typical reading and reading difficulty. Phonological deficits for children with dyslexia compared to RA-matched controls were, however, only evident in group contrasts. Findings thus also have important implications for identifying at-risk readers among their same-aged peers.

Keywords Balance automaticity · Cerebellar deficit · Dyslexia · Phonological deficit · Reading disability · Reading-level design

Reading disability or developmental dyslexia is one of the most common forms of learning disability (Shaywitz, Morris, & Shaywitz, 2008). Currently, the most widely accepted cognitive model of dyslexia is the phonological core deficit (PCD) theory. The PCD model has been supported by a large and convergent body of evidence from reading-level comparison studies, remedial intervention, and longitudinal research studies (e.g., Herrmann,

S. Irannejad (✉)
Office of Students with Disabilities (OSD), McGill University, 3600 McTavish Street, Room 3100, Brown Student Services Building, Montreal, Quebec H3A 1Y2, Canada
e-mail: shahrzad.irannejad@mcgill.ca

R. Savage
Faculty of Educational and Counseling Psychology, McGill University, 3700 McTavish Street, Montreal, Quebec H3A 1Y2, Canada
e-mail: robert.savage@mcgill.ca

Matyas, & Pratt, 2006; National Reading Panel, 2000; Rack, Snowling, & Olson, 1992; Savage & Pompey, 2008; Snowling, 1998). To summarize, many researchers agree that phonological awareness may be a causal factor in reading acquisition (Savage & Carless, 2005). As supported by substantial evidence, tasks tapping into phonological processing and phonological awareness seem to be concurrently related to reading ability (e.g., Badian, 1993; Savage, Carless, & Ferraro, 2007) and predict subsequent achievement and difficulties in reading after control for literacy (e.g., Cossu, Shankweiler, Liberman, Tola, & Katz, 1988; Pennington & Lefly, 2001; Wagner & Torgesen, 1987; Wesseling & Reitsma, 2001). The potentially causal role of deficits in phonological awareness and nonword processing in reading difficulties also seems to be well supported in studies that have used a reading-age match design in which individuals with dyslexia are matched to a group of younger readers who are reading at the same developmental level as those with dyslexia (e.g., Bradley & Bryant, 1983; Duncan & Johnston, 1999; Goswami & Bryant, 1989; Savage et al., 2005a; Thompson & Johnston, 2000). Further, a systematic review of findings from reading-level design studies that have investigated nonword processing seems to indicate that deficits were identified more clearly when children were asked to read bisyllabic and polysyllabic nonwords than when they read monosyllabic nonwords (Rack et al., 1992).

The cerebellar deficit

One recent extension of the PCD theory proposes that all of the underlying causes of dyslexia, including the phonological deficits, lie in more fundamental cerebellar deficits. The foundation of the cerebellar deficit theory (CDT) began within a cognitive-level framework of automatization deficit in which dyslexia was viewed as a symptom of a more general and pervasive deficit in skill acquisition (Nicolson & Fawcett, 1990). Nonetheless, as Fawcett, Nicolson, and Dean (1996) have noted, the cause of deficit in skill automatization was not explained within this cognitive-level structure. More recently, deficits in automatization were integrated within a neurological-level framework. At a neurological level, abnormal cerebellar–vestibular areas of the brain have been linked to automatization difficulties (Fawcett, Nicolson, & Maclagan, 2001; Nicolson, Daum, Schugens, Fawcett, & Schulz, 2002; Nicolson & Fawcett, 1999; Nicolson, Fawcett, & Dean, 2001; Reynolds, Nicolson, & Hambly, 2003). In its recent form, the CDT is, therefore, suggested to unify a neurological and a learning viewpoint and provide a parsimonious explanation for the existing deficit theories of dyslexia (Fawcett et al., 2001; Nicolson & Fawcett, 1999; Nicolson et al., 2001). Fawcett and colleagues 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 leading to deficits in writing, spelling, rapid naming, phonological processing, and finally, reading (Fawcett et al., 2001; Nicolson & Fawcett, 1999; Nicolson et al., 2001).

The cornerstone of the CDT appears to rest on the claim that children with dyslexia show impairment in their balance and motor skills. Evidence for this claim is derived from studies that used clinical cerebellar tests that assess postural stability or balance time, muscle tone, gait, or movement to investigate whether children with dyslexia display deficits in traditional signs of cerebellar impairment similar to patients with gross damage to their cerebellum (e.g., Fawcett & Nicolson, 1999; Fawcett et al., 1996). Ability to maintain balance was usually assessed using a blindfold balance task. Here, children had to maintain posture in standing position on one or two feet with arms

stretched out in front or on their sides upon a balance challenge (i.e., a gentle push) that was applied to their back without warning using a calibrated stability tester. Degree of wobble was then measured. Muscle tone in upper limbs was measured in standing or sitting position using tasks such as arm displacement or arm shake. Finally, ability to maintain or initiate complex voluntary movement was assessed using tasks such as repeated finger to finger pointing or toe-tapping speed. Using these tasks, Fawcett et al. (1996) compared three groups of children with dyslexia with mean ages of 10, 14, and 18 years with three groups of normal readers who were matched to those with dyslexia on chronological age (CA) and IQ. Fawcett et al. (1996) reported deficits for all cerebellar tasks in children with dyslexia. Higher incidence rates for impairments in children with dyslexia in this study, as opposed to those cited for cerebellar patients, led the authors to the conclusion 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). Fawcett and Nicolson (1999) attempted to replicate these findings in another study using a larger sample of 126 children with dyslexia and normal readers who were divided and matched into 4 different age groups (8–9, 10–11, 12–13, and 14–16 years). Significantly worse performance was reported in children with dyslexia, compared to CA-matched controls, in all three tests of cerebellar function (i.e., in postural stability, arm shake, and toe-tapping speed). Performance on postural stability and arm shake was reported to be exceptionally poor in all four dyslexia groups. In addition, children with dyslexia, compared to CA-matched controls, showed more difficulties on tests of nonword repetition and segmentation. Similar behavioral findings for cerebellar tasks and phonological processing skills were also reported in other studies by the authors (e.g., Fawcett & Nicolson, 1995a, Fawcett et al., 2001).

The CDT has gone through several, at times incongruent, revisions in the past few years. As indicated above, based on their findings, Fawcett et al. (1996) concluded that, as opposed to an impairment specific to a single region, a deficit across the entire cerebellum and hence across a range of cognitive processes is involved in dyslexia. Furthermore, according to Nicolson et al. (2001), the CDT predicted impairment in "cognitive, information processing, and motor skills." In a later discrepant claim, however, Nicolson and Fawcett (2006) disagreed with the assumption that a cerebellar deficit is linked only with motor and balance difficulties. In contrast to their earlier claims, the authors suggested that the "cerebellar deficit…claims only that the language-related regions of the cerebellum are affected in dyslexia" (p. 261). Nonetheless, the recent modifications made by the authors (e.g., Nicolson & Fawcett, 2006; Nicolson & Fawcett, 2007) to the CDT appear more congruent with a phonological deficit model or language-based deficits of dyslexia. Consequently, one may question the causal model of the CDT, the association between balance and motor difficulties and literacy impairment, and more importantly, the effectiveness of the motor and cerebellar tasks in diagnosing dyslexia.

Regardless of the noted discrepancies, examining balance and motor skills have been a major part of the original author's studies. The extended difficulties in static cerebellar tasks of balance and muscle tone have been claimed to have diagnostic power and to distinguish between the discrepancy-based dyslexia and garden-variety poor readers (Fawcett et al., 2001; Nicolson & Fawcett, 2006). Furthermore, some of the motor tasks as well as the balance task are currently part of the battery included in the Dyslexia Early Screening Test (Fawcett & Nicolson, 1995b) and the Dyslexia Screening Test (DST) (Fawcett & Nicolson, 1996). Despite the authors' recent revisions to the CDT in which deficits only in language-related regions of the cerebellum are proposed, impairments in postural stability and some motor skill deficits are addressed through

remedial techniques in the Dore Achievement Centers (Reynolds & Nicolson, 2007) through a range of exercises prescribed for the treatment of dyslexia. Findings for these cerebellar and balance intervention studies have been reviewed and critically appraised by Irannejad and Savage (2009). Consequently, given that (a) balance and motor skills have been an important part of the original authors' investigation of the CDT, (b) balance and motor tasks are a part of the author's dyslexia screening tests and dyslexia training, and (c) many other researchers who have investigated the CDT have included tasks related to postural stability and motor skills in their studies, it seems plausible to investigate an association (even if indirect) between a cerebellar deficit, balance impairment, possibly some motor difficulties, phonological awareness, rapid naming, and reading deficits.

Despite the growing body of research and attempts to replicate balance and motor impairments, findings remain inconsistent across studies. A critical appraisal of the behavioral evidence related to CDT points to three major issues that seem to explain the inconsistencies across evidence. These issues are explained below.

(a) *Method by which motor and cerebellar tasks have been assessed*. Early supportive evidence for the CDT has been derived from studies using Likert scales to measure cerebellar tasks of postural stability and muscle tone (e.g., Fawcett & Nicolson, 1999; Fawcett & Nicolson, 1995a; Fawcett et al., 1996). These scales may be insensitive and also less reliable because they depend on the subjective estimation of observers (Shipley & Harley, 1971). A few studies (e.g., Brown et al., 1985; Moe-Nilssen, Helbostad, Talcott, & Toennessen, 2003; Needle, Fawcett, & Nicolson, 2006) that have attempted to explore deficits in postural stability using more objective measures that also yield continuous quantitative data seem to provide different results often reporting no group differences (e.g., Brown et al., 1985, Needle et al., 2006). For example, Needle et al. (2006) obtained postural stability data from 17 adults with dyslexia and 20 adult normal readers using a motion-tracking system to measure degree of wobble during a balancing task. For the balancing task, participants were not pushed in the back but were only required to maintain a heel-to-toe position for a minute at a time, moving as little as possible. To measure the degree of wobble, motion-tracking sensors attached to the forefinger of each outstretched hand were used that tracked movement in three dimensions at a rate of 120 Hz. The recorded data were then analyzed in a computer. Needle et al. (2006) reported no differences between the group with dyslexia and control group in the heel-to-toe balancing task.

A study by Moe-Nilssen et al. (2003) that used a triaxial accelerometer sensor to investigate both postural stability and gait did find some support for the CDT in a sample of 22 children with dyslexia and 18 normal readers from Norwegian schools (ages 10 to 12). Moe-Nilssen et al. (2003) reported group differences in mean walking speed and impaired balance compared to controls specifically on tests of undisturbed balance with eyes open and fixated on a visual target on the wall. However, the findings of Moe-Nilssen et al. (2003) may be hard to interpret due to some factors related to their methodology. For example, the children in their study were randomly selected from a teacher-selected sample which included the top and bottom 5% readers from 200 pupils. Upon sample selection, only the poor readers were administered different reading as well as intelligence measures, but this procedure 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. Given that group comparisons were not made between poor and average readers results will be affected since above-average readers will be better than average readers in processes underlying reading such as phonological and speed naming skills (Savage et al., 2005b). Another important missing element in the methodology used in the study of Moe-Nilssen et al. (2003), as well as other studies reviewed here, is the lack of a reading-level design.

(b) *Lack of a reading-level design*. The second issue that may underlie inconsistencies across the evidence base for the CDT appears to be related to a lack of a reading-level design in cerebellar deficit studies despite its importance for drawing causal interpretations. Generally, a reading-level design that includes both a reading age (RA)- and a CA-matched 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 RA-matched 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, 1995a; Fawcett & Nicolson, 1999; Moe-Nilssen et al., 2003; Nicolson & Fawcett, 1990; White et al., 2006). In the study by Fawcett et al. (1996) described earlier, the authors have acknowledged the importance of a reading-level design as noted by Bryant and Goswami (1986) and reported using two of the CA-matched control groups with mean ages of 14 and 10 as RA-matched controls for the two older dyslexia groups with mean ages of 18 and 14. However, one of the problems with the design was that Fawcett et al. (1996) did not perform statistical tests to investigate whether the groups were in fact matched on their RA, an issue when observed mean differences between groups was sometimes large. Given this shortcoming, the causal interpretation of results obtained in the study may be questioned.

Savage et al. (2005a) carefully matched children on standardized reading tests and then compared performances on phonological, rapid naming, and reading tasks as well as postural stability/balance in nine 10-year-old poor readers to nine CA-matched and nine younger RA-matched controls. Poor readers were reported to perform significantly poorer than their CA-matched peers on naming speed and some of the phonological tasks (e.g., nonword reading and spoonerisms). They were also found to perform significantly poorer than their RA-matched controls in nonword reading. Interestingly, however, they performed significantly *better* compared to their younger RA-matched peers in their postural stability, suggesting that maturational factors may explain motor task performance. Similar conclusions have been drawn from a more recent study by Stoodley, Fawcett, Nicolson, and Stein (2006).

(c) *Lack of homogeneity in samples involved in studies.* Finally, the third issue related to incongruent findings for the CDT seems to be a lack of homogeneity of samples involved in the studies. 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, any 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 (c.f., Campbell & Stanley, 1963). In studies on dyslexia, the sample of individuals with dyslexia is usually drawn

from an extreme group that may differ in many ways from typical reading comparison groups. For example, participants characterized as having dyslexia are usually drawn from clinics or schools for dyslexia and diagnosed based on empirically criticized discrepancy-based definitions (e.g., Fawcett & Nicolson, 1999; Moe-Nilssen et al., 2003; White et al., 2006). The origin of control groups is either unknown or the group is recruited from public schools. As a result, any group differences found are likely to be confounded by the fact that the groups may be inherently different (Jackson & Butterfield, 1989).

Individual differences in cerebellar tasks

Beyond methodological issues in group comparison studies, there has been relatively little research exploring individual differences in tasks that assess cerebellar processing (Brookes & Stirling, 2005; Savage et al., 2005b). Only one study to date has explored individual differences in children experiencing significant reading problems. This absence is perhaps surprising because, as Ramus, White, and Frith (2006) have noted, individual data are important as they assess the extent to which children who are diagnosed as having dyslexia have similar sensorimotor and cognitive profiles and whether a single deficit could underlie all or some reading problems and has clear clinical implications for casework.

White et al. (2006) report a multiple case of 23 children with dyslexia and 22 normal readers (ages 8 to 12) who were matched to those with dyslexia on age, nonverbal intelligence, and gender. All children were reported to have a nonverbal 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. To investigate individual differences on measures, White et al. first removed participants in their control group who had "abnormally low performance" (defined as performance more than 1.65 SD below the mean of the control group) on any of the measures used in their study. Individual differences were then explored on a series of summary factors which were created to represent the different tasks that underlie modalities such as literacy, phonology, and motor domains in their study. The investigation of White et al. of individual children's performances suggested that the majority of cases of reading difficulties seemed to be directly caused by deficits in phonological processing which could not be accounted for by motor or other sensory impairments. White et al. report the existence of a "sensorimotor syndrome" suggested by 14 cases that showed motor, visual, and/or auditory deficits in addition to phonology and literacy deficits. Sensorimotor task performance did not, however, directly explain reading difficulties, whereas phonological deficits did predict reading difficulties.

The work of White et al. (2006) was critiqued in a series of response articles. Tallal (2006) noted the fact that the group differences found in the study of White et al. (2006) might have been intensified because of the fact that outliers in the control group were detected and removed. Goswami (2006) pointed out the fact that the study of White et al. did not include a RA-matched control group. 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 RA-matched control group on sensorimotor tasks. Without a reading-level control group which relies on raw and not age-standardized scores as were used in the study of White et al. (2006), interpretation of findings becomes

difficult. There is also a need to extend the findings of White et al. (2006) in different contexts to ensure that results are not limited to a specific context or curriculum.

The present study

This study attempted to address the three aforementioned methodological problems identified as possible explanations for inconsistent findings across cerebellar deficit studies to test the CDT using both group- and individual-level analyses.

(a) *Method by which motor and cerebellar tasks have been assessed*. This study included a wide range of motor and cerebellar measures including the static cerebellar tests, which are suggested to have a unique diagnostic power by Nicolson and Fawcett (2006). As opposed to the 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 the "Method" section) also added to the reliability of findings in this study. Statistical data reduction techniques were also employed to obtain extra, more reliable clusters of motor and cerebellar measures.

(b) *Lack of a reading-level design*. As indicated earlier, there is a scarcity of reading-level design in studies that have investigated the CDT. Given the importance of this design in drawing causal interpretations (Goswami, 2006) which has also been acknowledged by Fawcett et al. in their 1996 study and in order to address the critique of the study of White et al. (2006) which did not include a reading-level control, the present study included both RA- and CA-matched control groups.

(c) *Lack of homogeneity in samples involved in studies*. In order to minimize the impact of methodological weaknesses associated with sample homogeneity, the sample for this study was a nonclinical sample drawn from the same school system as done by Savage et al. (2005a). The present study was the first study in Canada to attempt investigating cerebellar deficits and a range of reading measures in a nonclinical sample of mainstream elementary students. The reason for using a mainstream sample was 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. The attempt was also to draw a larger sample in order to ensure obtaining a wide range of reading abilities including poor, average, and good readers.

From the view that a cerebellar deficit explains both reading and motor difficulties in children with dyslexia, then, according to the CDT, children with dyslexia should exhibit deficits in motor and cerebellar-related tasks compared to the RA- and CA-matched controls. In addition, considering the alleged link between a cerebellar deficit, phonological processing, rapid naming, and reading (Fawcett et al., 2001; Nicolson & Fawcett, 1999; Nicolson et al., 2001), the motor and cerebellar measures should also be as successful as phonological and rapid naming measures in distinguishing between the participants in the dyslexia group and those in the RA-matched control group in this study. Hence, the two questions investigated in this study are (a) do a group of children with dyslexia differ in their performance on any of the motor and cerebellar tasks, as well as phonological and rapid naming related measures when compared to both RA- and CA-matched control groups? and (b) does the cerebellar deficit provide a good explanatory model at the individual level?

Method

Participants

In the city where the study took place, access to a monolingual English-speaking sample is extremely limited. To recruit an English-speaking population, we approached English school boards in the city to obtain what are referred to here as *English-dominant bilinguals*, where somewhat more exposure to English than French is found in their background. Altogether, 348 parents whose children attended grades 1 to 5 were eligible[1] to participate in this study. All the parents were contacted. Only 89 provided consent for their children to participate in the present study. Out of the 89, 2 children could never be enrolled for testing due to absence from school. From the remaining 87 participants, data for 2 could not be used. One was excluded because of the child's difficulties in understanding and following directions and the other child did not complete the entire battery of tests required for the study. The remaining 85 participants completed the entire battery of measures included in this study. Data for these 85 participants were then used to select the 3 groups (i.e., dyslexia, RA-matched, and CA-matched). The selection process is described in the next section. This final pool included 51 children, ages 6 to 13 (30 boys and 21 girls).

Ethnicities reported in the sample included White Canadian (41%), White other (12%), Asian which included south/southeast and west Asian (25.5%), African American (6%), Hispanics (2.5%), Chinese (1%), and other (1%). The ethnic origin for 11% of participants was not reported. Results of a chi-square statistics comparing the proportion of first languages spoken (as reported by parents) to those available from Canada Statistics 2006 for the city where the study took place showed that the attempt to target the English-speaking population was successful. The chi-square was significant, $\chi^2(2, N=85)=240.29$, $p<0.000$, indicating that the sample included more Anglophones or *English-dominant bilinguals* than the reported proportion in this city (i.e., 62% vs. 12%) and fewer Francophones or *French-dominant bilinguals* (i.e., 8% vs. 65%).

Description of the dyslexia, RA-matched, and CA-matched groups

The first criterion for being included in the dyslexia group was participants' grade level. Among the original sample of 85 children, only those who attended grades 3, 4, or 5 were considered. The next and last criterion was participants' performance on the word identification subtest of the Wide Range Achievement Test—Third Edition (WRAT-3) (Wilkinson, 1993). Only those who obtained standard scores below 90 (i.e., below the 25th percentile) on the word identification subtest were included in the dyslexia group. This cutoff point has been used in other studies of dyslexia as well (Fletcher et al., 1994; Shaywitz et al., 2004; Siegel & Ryan, 1988). Using grade level and the aforementioned cutoff point, only 17 participants met the criteria for inclusion in the dyslexia group. No other criteria were applied to the selection of participants in this group.

As the next step, participants in the two control groups were selected to match to the dyslexia group. For the RA-matched group, only participants who were in grades 1 and 2 were considered. Following procedures used in other studies (e.g., Grainger, Bouttevin,

[1] The term "eligible" refers to children who attended schools in which English was the first language of instruction during elementary years. These were also the schools, and teachers, who provided permission for conducting the present study. It should also be noted that only in a limited number of schools in the city where the study took place English is the first language of instruction during elementary years.

Truc, Bastien, & Ziegler, 2003; Olson, Wise, Conners, Rack, & Fulker, 1989; Savage et al., 2005a; Snowling, 1981), these participants were matched to the dyslexia group based on their RA on the word identification subtest of the WRAT3. Similar to the procedure used by Savage et al. (2005a), the initial selection was to choose those children whose RA was exactly equal to their CA, then move outward to select children whose RA was as close as possible to their CA until 17 participants were included in the RA-matched group.

For the CA-matched group, only participants in grades 3, 4, and 5 were considered so that their CA matched the children in the dyslexia group. Following the procedure of Savage et al. (2005a), those whose RA was exactly equal to their CA were selected first. Next, selection moved outward to include children with the smallest difference between their reading and CA until 17 children were selected in the CA-matched group.

To confirm if the groups were matched appropriately, data on RA (in months) and CA were submitted to two one-way analyses of variance (ANOVA). Data on IQ (including verbal and nonverbal skills) were submitted to another three one-way analyses to investigate if the three groups matched on their level of intellectual functioning. As shown in Table 1, a significant main effect of group was found for RA [$F(2, 48)=31.22$, $p<0.001$] and CA [$F(2, 48)=20.57$, $p<0.001$]. Simple main effect comparisons confirmed a successful match on both RA and CA. The RA of participants in both dyslexia and RA-matched groups was lower than that of the CA-matched control. Additionally, participants in dyslexia and CA-matched groups were significantly older than those in the RA-matched group. Results also indicated that the three groups matched on the level of intellectual functioning (IQ: $F(2, 48)=2.34$, $p=0.11$; vocabulary: $F(2, 48)=0.93$, $p=0.40$; and matrix reasoning (MR): $F(2, 48)=2.51$, $p=0.09$).

Results of a chi-square test also indicated that the three groups matched on the first language (L1) spoken by participants, $\chi^2(2, N=47)=2.94$, $p=0.569$. For the majority of participants within all three groups, English was reported to be L1. For a smaller proportion of the participants within each group, French and a language other than English or French were reported as L1. Nonetheless, as previously explained, 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.

Procedure

Following ethical approval, measures were administered and scored blindly. The whole procedure took approximately 70 to 90 min per child. For most of the participants, testing

Table 1 Mean and standard deviations on RA, IQ and IQ subscales, and CA by group

Source	Dyslexia		RA-matched		CA-matched		*df*	*F*	*p*	Source of effect
	M	SD	M	SD	M	SD	Between subjects			
RA (in months)	83.53	8.97	86.41	5.81	103.59	8.83	2	31.22	0.00	(1=2)<3
IQ (standard scores)	92.59	13.68	101.24	12.26	101.18	14.24	2	2.34	0.11	1=2=3
Vocabulary (scaled scores)	8.59	2.69	9.47	2.07	9.76	3.01	2	0.93	0.40	1=2=3
MR (scaled scores)	8.53	3.28	10.82	3.38	10.65	3.32	2	2.51	0.09	1=2=3
CA (in months)	113.00	17.70	87.47	4.84	103.76	8.81	2	20.57	0.00	(1=3)>2

was completed in two sessions at school sites in rooms provided by the school. A brief measure of cognitive functioning and reading, phonological, and rapid naming tasks were administered in the first session that lasted 45 to 60 min. A series of simple motor and cerebellar tasks were completed in the second session that lasted approximately 30 min. For only a few participants, testing was completed in one session in the university's research laboratory. All participating children 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.

Cognitive and reading measures

Wechsler Abbreviated Scale of Intelligence (Psychological Corp., 1999) To obtain an estimate of verbal intelligence, the Vocabulary (Voc.) subtest of the Wechsler Abbreviated Scale of Intelligence was administered in which children were asked to define words pronounced by the examiner. To obtain an estimate of nonverbal intelligence, the MR subtest was used in which children were asked to select an option from five choices that best completed a visual pattern. For the sample in the present study, the reliability coefficient, using Cronbach's α, was 0.88 for the Voc. and 0.85 for the MR subtests, which are similarly high as the published reliability estimates of 0.89 and 0.92 (Psychological Corp., 1999) reported for the standardization sample, respectively.

Word identification of the WRAT-3 (Wilkinson, 1993) Depending on age, examinees aged 6 to 75 are asked to either begin with pronouncing the letters of the alphabet or the list of words ordered in decreasing fluency and increasing complexity. For the present sample, the reliability coefficient obtained for this subtest was 0.95, which is similar to the published median reliability estimates, ranging from 0.90 to 0.95, reported for the standardization sample used for WRAT-3 reading subtests.

Test of Word Reading Efficiency (Wagner, Torgesen, & Rashotte, 1999) The Test of Word Reading Efficiency (TOWRE), designed for ages 6 to 25, was administered to measure word reading accuracy and fluency which have been recommended as the best way to identify children who have fallen behind in these skills (Torgesen, 2002). The child is asked to read a list of real words (sight word reading) and a list of nonwords (phonemic decoding), ordered in increasing orthographic complexity and length, as quickly as possible. The number of real words and nonwords read accurately within 45 s is recorded. The published test–retest reliability reported for the TOWRE ranges from 0.83 to 0.93. A test–retest reliability estimate for TOWRE tasks could not be obtained for the present sample because participants were only tested once.

Elision subtest of the Comprehensive Test of Phonological Processing (Wagner et al., 1999) The elision subtest provides an estimate of phonological awareness skills. Examinees aged 7 to 24 are asked to repeat a word, minus a single sound, which can be in the initial or middle position (e.g., say *cup* without the *k* sound or say *powder* without the *d* sound). For the present sample, an internal consistency coefficient α of 0.95 was obtained, which was similarly high as the published reliability estimate of 0.93 reported for the Elision subtest.

Word attack subtest of the Woodcock Reading Mastery Tests—Revised (Woodcock, 1987; Woodcock & Johnson, 1989) The ability to apply phonetic skills to pronounce a list of nonwords ordered in increasing complexity is measured in this test designed for ages 5 to 75. For the present sample, a Cronbach's α of 0.97 was obtained, which is above 0.90 as

the published reliability estimate reported for the Woodcock Reading Mastery Tests—Revised subtests, including the word attack subtest.

Rapid digit, letter, and object naming subtests of the Comprehensive Test of Phonological Processing (Wagner et al., 1999) Examinees, ages 5 to 24, are presented with two pages (one page at a time) containing four rows and nine columns of either six randomly arranged digits, letters, or objects. They are asked to name the stimuli as quickly as possible. In each task, the total number of seconds taken to name all of the digits, letters, and objects on both pages presented is measured. The published test–retest reliability estimates for Comprehensive Test of Phonological Processing subtests are reported to range from 0.74 to 0.97.

Motor and cerebellar tasks

Peg moving (Fawcett & Nicolson, 1995a) The *peg moving* task is known to be sensitive to cerebellar damage (Haggard et al., 1995; Miall & Christensen, 2004). Similar to Fawcett and Nicolson's (1995a) study, children were asked to hold a pegboard steady with their nonpreferred hand and move a row of pegs in the top row of the board, one at a time, with their preferred (i.e., writing) hand as quickly as possible, jumping over one empty row into the next. Children repeated this procedure until five rows were completed. 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, 1995a). 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 average correlation obtained among mean peg moving times for the 5 trials was 0.69, which is the same as the published test–retest reliability reported by Annett, Hudson, and Turner (1974) for 53 adults. Using the last two trials of the peg moving test, an internal consistency α coefficient of r=0.81was also calculated for the sample.

Bead threading of the Dyslexia Screening Test (Fawcett & Nicolson, 1996) Children were asked to hold an 85-cm string (3 mm in diameter) vertically in their writing hand and thread 12 round wooden beads (each 4 cm in diameter with a hole of approximately 0.5 cm) on the thread as quickly as possible with their other hand. The dependent variable was the number of beads threaded in 30 s starting from the time they touched the first bead. Children with no cerebellar deficit were expected to thread more beads in 30 s than children with a mild cerebellar deficit (Fawcett & Nicolson, 1995a). The published test–retest reliability reported for the bead threading task is 0.76.

Postural stability of the DST (Fawcett & Nicolson, 1996) To assess balance automaticity, the procedure used in Fawcett and Nicolson's (1996) DST and Fawcett and Nicolson's (1999) study was followed. The balance challenge was repeated in six trials. In three, participants' arms were at sides, and in the next three, the arms were straight out in front. Degree of sway was assessed and recorded using a PASCO three-axis acceleration sensor that was attached to a Velcro belt and secured to the hip of the participant. The data were recorded by the three-axis acceleration sensor and then analyzed directly through a data collection and analysis interface installed on a computer. During each of the six trials, accelerations were recorded over a period of approximately 7.023 s, as shown on the x-axis

in Fig. 1. The summary data reported by the interface were the mean of all accelerations recorded for the entire recording period as well as the minimum and maximum acceleration points.

The most appropriate way to measure the degree of sway in this task was to calculate a deviation score for each participant (Dr. P. Stapley, personal communication, October 6, 2008). 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. For each participant, the time period between 0 to approximately 1.5 s included the recordings prior to the administration of the balance challenge. To control for any anticipatory sway once the task was learned, the recording area between 0 and 0.4 s was chosen as the area prior to administration of the balance challenge which was used to calculate the deviation score. As shown in Fig. 1, this area was calculated by multiplying the length of the rectangle, which was the time period of 0.4 s, by the mean of accelerations (i.e., 0.684) recorded over this time period which included 40 acceleration recordings (i.e., $0.684 \times 0.4 = 0.274$). 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 s (i.e., 0.274) was then subtracted from the maximum degree of acceleration rate (i.e., 12.158) recorded upon administering the balance challenge. For each participant, the final scores were two average deviation scores (arms at sides and arms in front). In accordance with Fawcett and Nicolson's (1999) predictions, children with no cerebellar deficit were expected 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 was 0.61 for the three trials with arms at sides and 0.18 for the three trials with arms straight out in front. For postural stability task with arms at sides, the internal consistency coefficient α was 0.79 for all three trials and 0.70 for the last two trials. For postural stability task with arms straight out in front, the internal consistency coefficient α was 0.35 for all three trials and 0.24 for the last two trials. Due to the small correlations and internal reliability obtained for the postural stability task with arms straight out in front, 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

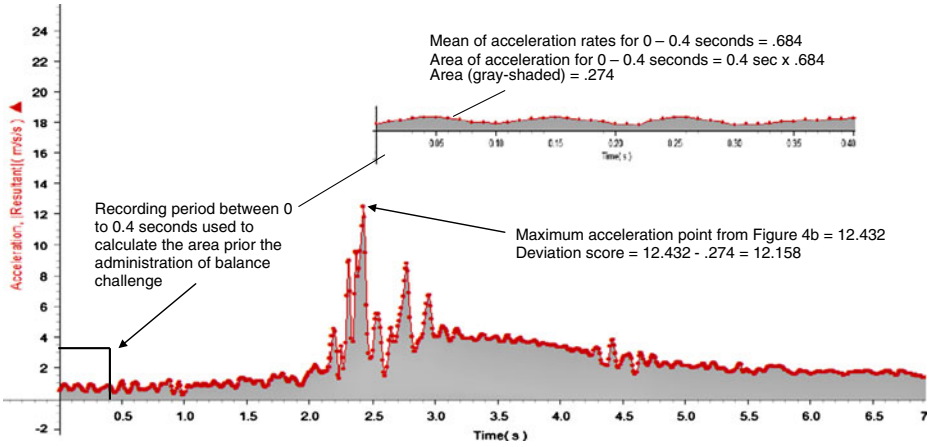

Fig. 1 Graphic representation and data recorded for participant with arms straight out in front with illustration of recording period prior to administration of challenge (i.e., 0–0.4 s)

period. The average correlation obtained among the mean acceleration areas calculated for the three postural stability trials with arms straight out in front was 0.43. The internal consistency coefficient α was 0.69 for all three trials and 0.57 for the last two trials.

Muscle tone (Fawcett & Nicolson, 1999) The Wartenberg Pendulum Test was used to assess muscle tone. The pendulum test has been used in studies of adults, such as healthy elderly clients 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 (Nordmark & Anderson, 2002), as an objective measure of muscle tone. Here, an electronic goniometer and an angle sensor were used to measure and record the angle at the knee joint. The angle sensor was connected to the computer through a USB link, analogous to the one used in the postural stability task, and data were sent digitally to interface data collection and analysis software. Following Nordmark and Anderson (2002), participants wore the electronic goniometer on their right knee. They sat on the edge of a table high enough that their feet did not touch the floor. In order to decrease measurement error related to voluntary muscle activities (i.e., assisting or resisting pendular movement) that have also been reported in some previous studies (Brown et al., 1988), time was devoted for practice prior to actual recordings until participants felt comfortable with the procedure. In a few cases, relaxation was achieved by closing the eyes during the procedure.

Once the task was fully understood, the experimenter lifted the participant's right foot so that the leg was extended at 180°. The foot was then dropped to swing like a pendulum. This procedure was repeated three times for each participant and each time the angle at the knee joint was recorded by the angle sensor. The most commonly measured parameter that is 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 et al., 1988; Brown et al., 1988; Nordmark & Anderson, 2002). As shown in Fig. 2, to calculate this ratio, the measurements taken from each recording were the angle of the knee joint at its maximum bend (i.e., 103.295°) and at its resting position (i.e., 84.242°). 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). For the present sample, the average correlation between the muscle tone ratios obtained for all three trials was 0.76. The internal consistency coefficient α computed was 0.85 for ratios obtained for the last two trials and 0.90 for all three trials, which were similar to published test–retest reliability of 0.96 (Bohannon, 1987) or 0.84 (Stillman, Phty, McMeeken, & Phty, 1995) for the pendulum test that has been measured in some studies for spasticity.

Toe-tapping speed (Fawcett & Nicolson, 1999) Following the procedure used in Fawcett and Nicolson's (1999) study, 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 a laptop computer to record the tapping sound. Sounds were recorded on the computer and the speed of tapping was assessed accurately using a digital audio editor, namely, the Audacity sound wave analysis software. 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. 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). However, as Fawcett and Nicolson (1999) have suggested, many other factors could also affect speed of

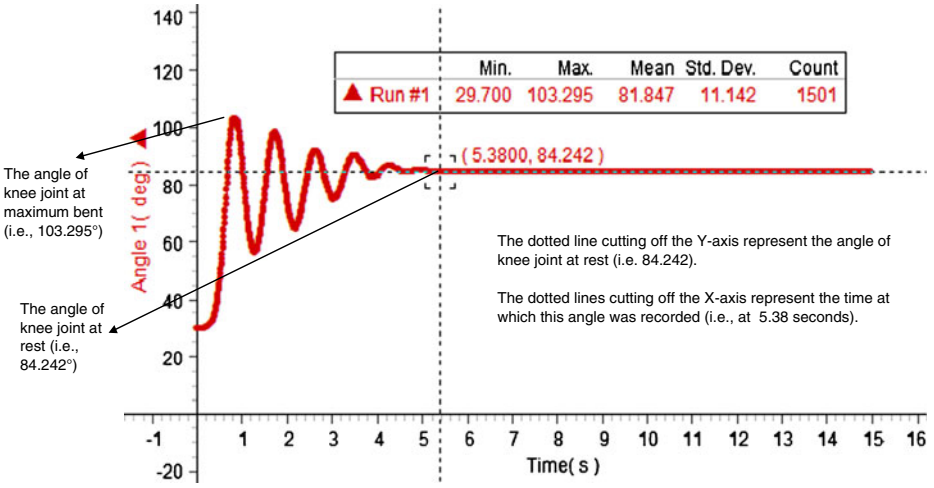

Fig. 2 Angle recording traces using electronic goniometer and angle sensor for a participant in the sample

tapping. The Spearman–Brown split-half reliability as well as the internal consistency coefficient α obtained for the tapping time for the left and right foot was 0.89 in both cases.

Results

Preliminary analyses

Normality Due to the presence of outliers, the assumption of normality was not met for measures of elision, rapid letter, digit, and object naming, as well as word attack (all items) and word attack (polysyllables).[2] Most outliers were within the dyslexia group (e.g., two on all items of word attack, three on polysyllable items of word attack, three on elision, or one on rapid object naming). Some of the outliers were also within the RA-matched control group (e.g., two on each rapid digit and letter naming). While none of the children within the RA-matched group were outliers on word attack, six were outliers on elision with only two that were more than 1 SD below the mean and four that were more than 1 SD above the mean. Finally, there were only two outliers within the CA-matched control group with one on all items of the word attack and the other on rapid object naming. Several recommended methods were used to correct the outlying cases, including (a) regression in which measures with large correlations were used to predict the measure with outliers, (b) rows and columns were used to calculate the expected values for measures with outliers, and (c) score alteration in which the score(s) on the outlying cases were changed so that they were deviant but not as deviant as they were previously (Tabachnick & Fidell, 2007). Subsequently, the main analyses were performed several times, namely, (a) with uncorrected measures, (b) with the outliers excluded, and (c) with corrected measures.

[2] Given that performance on more complex nonwords have been found to be more likely in identifying a deficit in individuals with dyslexia (Rack et al., 1992), group and individual differences on polysyllable nonwords in word attack task were also investigated.

Given that the pattern of results obtained for word attack (all items), TOWRE word and nonword reading, and the three rapid naming measures yielded similar findings in all analyses, the results presented here are those performed on the uncorrected measures. For elision, the pattern of results for the analyses using the corrected measures was the same as the analyses performed excluding the outlying cases. Performing nonparametric tests on elision yielded similar findings. Hence, considering the large proportion of outlying cases (i.e., three in the dyslexia group and six in the RA-matched control) that needed to be transformed to normalize elision, the results of the nonparametric tests performed on the uncorrected elision are presented here. Nonparametric tests were also used for word attack that represented the polysyllables because the distribution of this measure for the dyslexia group was markedly different than the other two groups.

Reduction of motor and cerebellar variables Arguably, it might be best to investigate each of the motor and cerebellar measures individually. Nonetheless, in order to maintain economy of reporting and more importantly to increase the reliability and robustness of results, Steven's (1996) recommendation was followed and the motor (i.e., peg moving and bead threading) and cerebellar measures (i.e., postural stability, muscle tone, and toe tapping) were reduced into separate clusters using a principal component analysis (PCA). It should be noted that both toe tapping and postural stability tasks were significantly positively skewed. 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). Following Tabachnick and Fidell's (2007) recommendations, transformations were applied to correct these variables and the transformed variables were entered in the PCA. Altogether, three PCAs were performed to investigate how these measures loaded together. In the first analysis, postural stability and toe tapping loaded strongly on separate components, explaining 57.51% of the variance. Loadings for bead threading and peg moving were not well defined as they loaded on both components. Muscle tone not only loaded poorly on the two extracted components, it was also not correlated significantly with any of the other motor and cerebellar measures. In a second PCA that excluded muscle tone, the variance explained improved slightly from 57.61% to 65.64%, but the pattern of loadings remained similar to the first analysis. An inspection of the scree plots for this analysis indicated the presence of a third eigenvalue with a value of 0.91. Subsequently, a third PCA was conducted that included the third eigenvalue. As shown in Table 2, this led to extraction of three clear components which together explained a total of 80.51% of the variance. The solutions extracted from this PCA remained stable when the analysis was repeated using an oblique instead of a varimax rotation as well as a principal axis factoring. The three components which were used in the main analyses included (a) motor (i.e., bead threading and peg moving), (b) postural stability, and (c) toe tapping components (i.e., the two postural stability and toe tapping tasks, respectively). Muscle tone was used as a separate measure in subsequent analyses.

Main analyses

Question 1 Do a group of children with dyslexia differ in their performance on any of the motor and cerebellar tasks, as well as phonological and rapid naming related measures when compared to both RA- and CA-matched control groups?

Table 2 Component loadings and communalities (h^2) for PCA and varimax rotation on motor and cerebellar measures with eigenvalues above 0.90

Measure	Component 1	Component 2	Component 3	Communality h^2
Log10 postural stability (hands at sides)	−0.01	0.83	−0.24	0.75
Log10 postural stability (hands stretched out in front)	−0.14	0.87	−0.07	0.77
Inverse log toe tapping (left foot) (right foot)	0.91	−0.12	0.25	0.90
Inverse log toe tapping (right foot) (left foot)	0.95	−0.05	0.07	0.92
Bead threading	0.06	−0.09	0.90	0.83
Peg moving	0.28	−0.27	0.72	0.67
% of variance explained by component	43.56	22.06	14.89	

Ten one-way ANOVA were performed with word attack (all items), TOWRE word and nonword reading, rapid digit, letter and object naming, postural stability, toe tapping, motor components, and muscle tone ratio as the dependent measures. The adjusted α level using the Bonferroni method was 0.005 (α of 0.05 divided by 10). In each analysis, there was one between-subjects factor with three levels (group: dyslexia vs. RA-matched vs. CA-matched). As indicated earlier, the results reported are based on the analyses performed on the uncorrected measures. For elision and word attack (polysyllables) measures, nonparametric tests were used that are described in the text.

As shown in Table 3, these analyses revealed a main effect of group for word attack (all items) ($F(2, 48)=23.24$, $p<0.001$), word attack (polysyllables) ($\chi^2(2, N=51)=27.29$, $p<0.001$), elision ($\chi^2(2, N=51)=16.733$, $p<0.001$), word reading and nonword reading efficiency ($F(2, 48)=16.46$, $p<0.001$ and $F(2, 48)=14.53$, $p<0.001$ respectively), but not for rapid digit ($F(2, 48)=1.94$, $p=0.16$), letter ($F(2, 48)=0.29$, $p=0.75$), and object naming ($F(2, 48)=0.43$, $p=0.65$), postural stability ($F(2, 48)=0.88$, $p=0.42$), toe tapping components ($F(2, 48)=1.49$, $p=0.24$), and muscle tone ratio ($F(2, 48)=0.52$, $p=0.60$). For significant main effects of ANOVA, simple comparisons were explored using the adjusted α level noted earlier. For pairwise nonparametric Mann–Whitney U tests which normally follow a significant omnibus test, family-wise error rate is already controlled for up to three pairwise comparisons (Howell, 2010). At an $\alpha=0.05$ for word attack and elision and $\alpha=0.005$ for word and nonword reading efficiency, these comparisons revealed that the performance of the dyslexia group was significantly poorer than the CA-matched group on word attack (all items), word attack (polysyllables) ($U=10.50$, $p<0.001$), elision ($U=30.50$, $p<0.001$), as well as word and nonword reading efficiency. Additionally, for polysyllabic nonwords, the dyslexia group was also significantly poorer than the RA-matched group ($U=92.50$, $p=0.03$). The difference between these two groups was also marginally significant for elision ($U=87.50$, $p=0.05$). Of note were the simple comparisons in the motor component. While the main effect of group was not significant for this component at the corrected α level of 0.005, it was significant at the α level of 0.05, $F(2, 48)=4.30$, $p<0.05$. Main effect

comparisons indicated that participants in the dyslexia group performed better on motor tasks compared to those in the RA-matched control group.

As shown in Table 4, the magnitude of effects was also explored in terms of Cohen's *d* effect size ratios (i.e., difference between standardized means) (Cohen, 1988; Tabachnick & Fidell, 2007), which have been used in some of the previous studies (e.g., Fawcett & Nicolson, 1999; Savage et al., 2005a). Throughout the table, a negative score indicates poorer performance. Moderate positive effect sizes were found for nonword reading and elision for contrasts made between the dyslexia group and the RA-matched control. This was combined with large positive effect sizes for contrasts of the dyslexia group and the CA-matched control as well as the RA- and CA-matched controls for the same measures. In comparison to the moderately large effect size for nonword reading that represented all test items, the effect size for polysyllable nonwords for the contrast between the dyslexia group and the RA-matched control was large. For word and nonword reading efficiency, only the effect sizes for contrasts between the dyslexia group and the CA-matched control and the RA- and CA-matched controls were largely positive. Among the rapid naming measures, a moderate positive effect size was found for rapid digit naming when the dyslexia group's performance was contrasted against the CA-matched control and when the performance of the RA-matched group was contrasted against the CA-matched group.

Among motor and cerebellar measures, small positive effect sizes were found when the dyslexia group's performance on postural stability was contrasted against the performance of the RA- and CA-matched controls. Conversely, for contrast of the dyslexia group with the RA-matched 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 group with the RA-matched control combined with a small negative effect size for contrast of the dyslexia group and the CA-matched control and a moderate positive effect size for contrast of the RA- and CA-matched controls.

Question 2 Does the cerebellar deficit provide a good explanatory model at the individual level?

Similar to the study of White et al. (2006), individual differences were investigated by displaying the individual data on the outcome measures for this study in figures. The measures included (a) nonword decoding (all items), (b) nonword decoding (polysyllables), (c) elision, (d) alphanumeric rapid naming summary score (i.e., letters and digit naming), (e) word reading efficiency, (f) nonword reading efficiency, (g) postural stability summary score, (h) toe tapping summary score, (i) muscle tone, and (j) motor summary score. Following the procedure of White et al., summary scores for alphanumeric rapid naming, postural stability, toe tapping, and motor summary scores were calculated in that standardized *z* scores were averaged (calculated once in relation to the performance of the RA-matched group and then in relation to the performance of the CA-matched group, respectively) for each participant on each group of tasks. To create the rest of the scores (i.e., "a," "b," "c," "e," "f," and "j"), scores on RA, word attack (all items), word attack (polysyllables), elision, TOWRE word and nonword reading, and muscle tone ratio were each saved as standardized z scores (in relation to RA- and CA-matched controls, respectively).

Unlike White et al. (2006) who used a cutoff point of −1.65, the cutoff point for identifying extreme poor cases on the outcome measures in the present study in the group comparison reported previously was the same cutoff point used to identify individuals with dyslexia (i.e., standard scores below 90 or −0.67 SD), shown by a broken line in figures. On all figures, values on the *y*-axis are *z* scores. For clarity, all lines that corresponded to the 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. In all figures, positive scores indicated good performance and negative scores indicated poor performance. For economy of reporting, only some of the key figures are displayed.

Literacy outcome measures Among reading measures, nonword decoding (all items) (Fig. 3a), alphanumeric rapid naming, as well as word and nonword efficiency did not successfully differentiate between participants in the dyslexia group and those in the RA-matched control. A similar pattern was found for the alphanumeric rapid naming when participants in the dyslexia group were compared to those in the CA-matched control. However, the pattern was

Table 3 Mean and standard deviations on word attack, elision, TOWRE word and nonword 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) (WA)	10.65	5.42	14.76	8.50	25.88	5.89	2	23.24	0.00	(1=2)<3
WA (polysyllables)	0.47	1.13	1.53	1.77	4.71	2.29	2	27.29[a]	0.00	1<2<3
Elision	6.82	2.92	8.94	4.52	12.47	4.38	2	16.73[b]	0.00	1<2<3
TOWRE word reading (TWR)	35.65	15.31	33.35	14.20	57.82	11.39	2	16.46	.00	(1=2)<3
TOWRE nonword reading (TNWR)	13.00	9.23	12.76	7.85	27.65	10.42	2	14.53	0.00	(1=2)<3
Rapid digit naming (RDN)	46.71	12.64	46.59	12.53	40.00	8.44	2	1.94	0.16	1=2=3
Rapid letter naming (RLN)	49.35	12.77	46.76	11.62	45.88	16.34	2	0.29	0.75	1=2=3
Rapid object naming (RON)	70.71	14.45	68.00	9.51	66.82	13.15	2	0.43	0.65	1=2=3
Postural stability (PS)	0.04	0.88	−0.29	1.05	−0.36	0.85	2	0.88	0.42	1=2=3
Toe tapping (TT)	0.16	0.70	−0.38	1.02	0.007	1.03	2	1.49	0.24	1=2=3
Motor	0.50	0.91	−0.42	1.00	0.15	0.86	2	4.30	0.02	1>2=3
Muscle tone ratio (MTR)	0.02	1.04	0.06	0.81	−0.27	1.19	2	0.52	0.60	1=2=3

For word attack (polysyllables) and elision variables, nonparametric tests were used

WA number of correct nonwords read, *Elision* number of correctly completed items, *TWR* number of correct words read in 45 s, *TNWR* number of correct nonwords read in 45 s, *RDN*, *RLN*, and *RON* total number of seconds taken to name digits, letters, and objects, *PS* PS component based on average deviation scores over three trials for each PS task, *TT* TT component based on total number of seconds taken to tap 10 times with each right and left foot, *Motor* motor component based on peg moving (i.e., mean time in seconds to complete the five peg moving trials) and bead threading (i.e., number of beads threaded in 30 s), *MTR* average ratio (for three trials) of angle of knee joint at its maximum bend and at its resting position

[a] χ^2 value for Kruskal–Wallis *H* omnibus test for word attack (polysyllables)

[b] χ^2 value for Kruskal–Wallis *H* omnibus test for elision

Table 4 Mean effect sizes measured in Cohen's *d* for the dyslexia group versus the RA- and CA-matched controls and for the RA-matched control versus the CA-matched control

Variable	Dyslexia vs.		RA-matched vs.
	RA-matched	CA-matched	CA-matched
Word attack (all items)	0.59	2.77	1.57
Word attack (polysyllables)	0.74	2.49	1.39
Elision	0.57	1.29[a]	0.82
Word reading efficiency	−0.16	1.69	1.96
Nonword reading efficiency	−0.03	1.53	1.66
Rapid digit naming	0.01	0.64	0.64
Rapid letter naming	0.22	0.24	0.06
Rapid object naming	0.23	0.29	0.11
Postural stability	0.35	0.48	0.08
Toe tapping	−0.64	−0.18	0.39
Motor	−0.99	−0.41	0.63
Muscle tone ratio	−0.04	0.27	0.33

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

different for nonword decoding (all items) (Fig. 3b). As illustrated in Fig. 3b, more participants in the dyslexia group fell below the cutoff point than those in the CA-matched control (i.e., nine vs. zero for all nonword items). A similar pattern was found for word and nonword reading efficiency (i.e., eight in dyslexia vs. zero in the CA-matched group for word reading efficiency and eight in dyslexia vs. one in the CA-matched group for nonword reading efficiency). Among the literacy measures that have been the focus of PCD theory, the patterns observed for nonword decoding polysyllables and elision in comparisons against the RA-matched control (Figs. 4a and 5a, respectively) suggested that *fewer* participants in the dyslexia group compared to the RA-matched control had scores falling above the mean of 0 (i.e., 3 vs. 9 for polysyllable nonwords and 3 vs. 10 for elision). On the other hand, in comparisons against the CA-matched control (Figs. 4b and 5b, respectively), the differentiation between participants was more clear below the cutoff point for polysyllable nonwords and elision with more participants below the broken line in the dyslexia compared to the CA-matched control group (i.e., 14 vs. 0 for polysyllable nonwords and 9 vs. 0 for elision).

Motor and cerebellar measures Among the motor and cerebellar measures, only muscle tone seemed to illustrate a trend in comparisons against the RA-matched control group. This means that somewhat more extreme poor cases were identified in the dyslexia group (i.e., four) compared to the RA-matched control (i.e., two), as seen in Fig. 6c. Nonetheless, none of the cases in the dyslexia group that were identified as outliers on the muscle tone measure seemed to have difficulties in any of the associated literacy measures, except for one case that was identified as an extreme poor case on alphanumeric rapid naming, as well as word and nonword reading efficiency measures. For motor measures, the trend indicated that in fact more individuals in the RA-matched group fell below the cutoff line compared to those in the dyslexia group as seen in Fig. 7a (i.e., nine vs. two). A similar pattern was found for toe tapping (nine in the RA-matched group vs. one in dyslexia). This trend also seemed to be present for toe tapping in comparisons against the CA-matched control (i.e., eight in the CA-matched group vs. three in dyslexia).

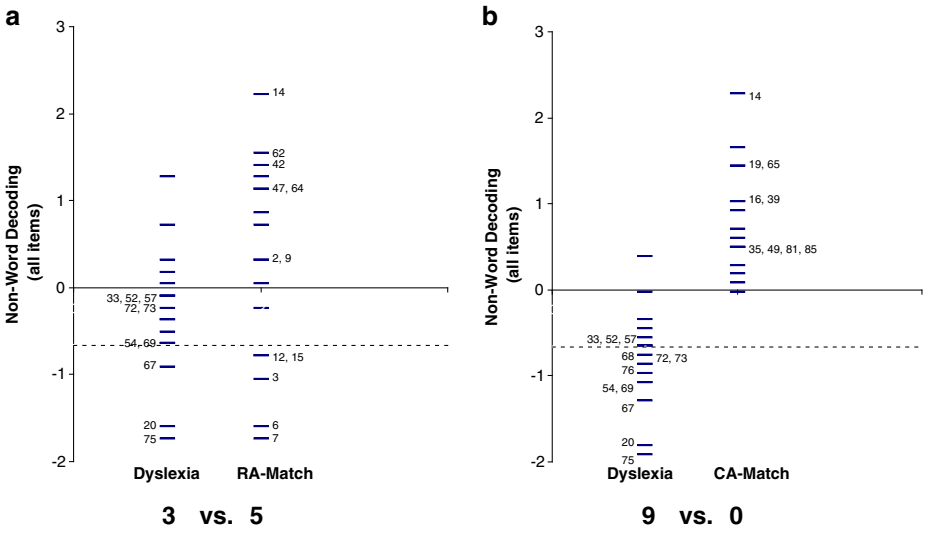

Fig. 3 Graph of individual performance for nonword reading (all items)

Discussion

In this study, two questions were investigated to test the CDT. A database of school-sampled children was analyzed using (a) principled group contrasts and (b) analysis of the deficits at an individual level of a dyslexia group versus the RA- and the CA-matched controls ($n=17$ in each case). 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).

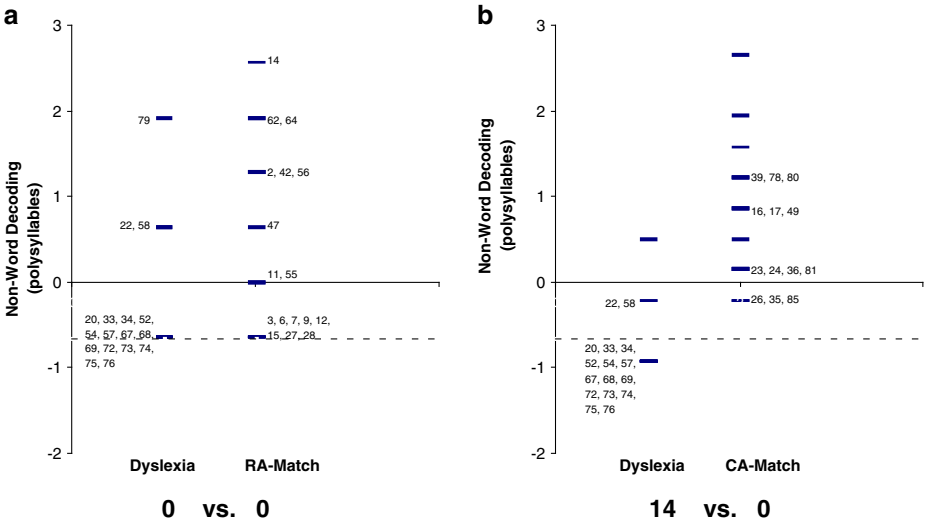

Fig. 4 Graph of individual performance for polysyllable nonwords

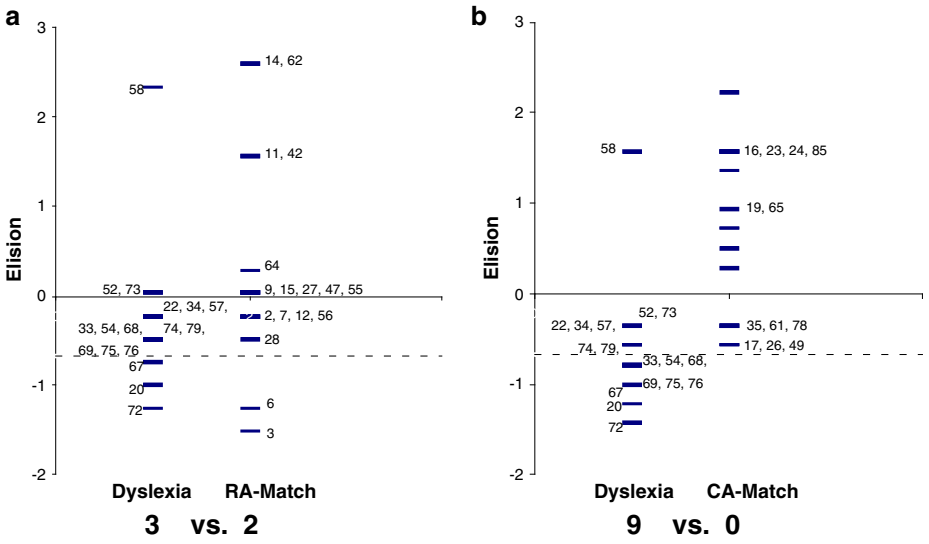

Fig. 5 Graph of individual performance for elision

Question 1 Do a group of children with dyslexia differ in their performance on any of the motor and cerebellar tasks, as well as phonological and rapid naming related measures when compared to both RA- and CA-matched control groups?

Motor and cerebellar measures Contrary to what would be expected on the basis of the CDT hypothesis, motor and cerebellar tasks did not successfully differentiate between participants in the dyslexia group and those in the RA- and CA-matched controls. For postural stability, this study's finding from group contrasts, in terms of effect size, against the CA-matched control was similar to the study of Savage et al. (2005a) indicating moderately poorer performance in participants in the dyslexia group compared to their same-aged peers only. Although this may provide some support for possible motor problems in the dyslexia group, findings of equal performances found in the dyslexia group and the RA-matched control were not consistent with a causal role for cerebellar processing. As indicated by Goswami and Bryant (1989), positive results in a CA-matched 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. In addition, the positive finding for postural stability may be explained by factors other than a cerebellar deficit. As a systematic review by Rochelle and Talcott (2006) has indicated, the relationship between balance deficits and dyslexia seem to be indirect and most strongly influenced by variables other than reading skills such as attention deficit hyperactivity disorder symptoms or developmental coordination disorder. Indeed, evidence has shown that differences between poor and same-aged good readers were solely explained by deficits in naming speed and phonological awareness once the effect of attention was controlled statistically (e.g., Raberger & Wimmer, 1999, 2003; Ramus, Pidgeon, & Frith, 2003).

Furthermore, the present findings for motor and cerebellar measures differed from those reported in Fawcett and Nicolson's studies (e.g., Fawcett et al., 1996; Fawcett & Nicolson, 1999). For example, the large effect sizes reported by Fawcett et al. (1996) (i.e., −1 SD or

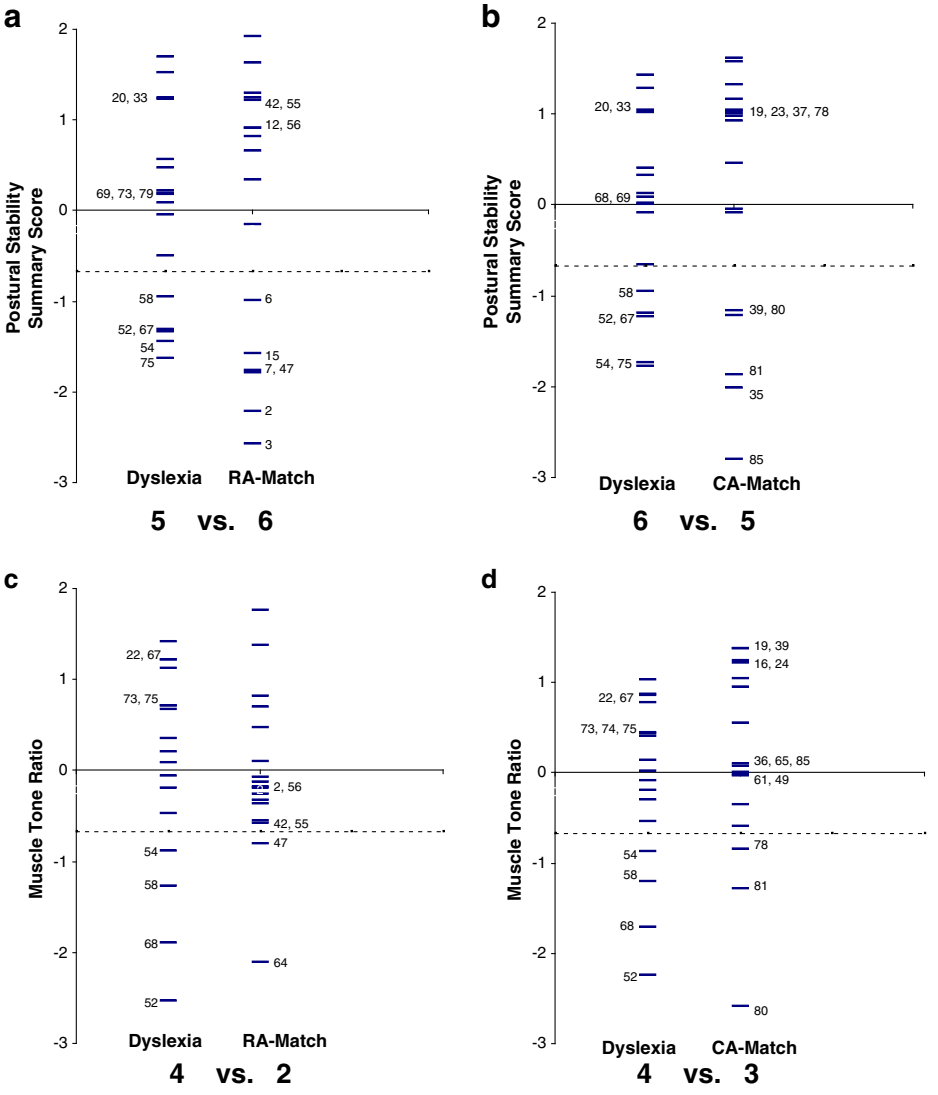

Fig. 6 Graph of individual performance for postural stability summary score and muscle tone

greater) were not found here. It may be argued that one reason for the small effect sizes found in the present study can be attributed to the dyslexia sample in this study. This is to say that given the somewhat lower mean score for verbal IQ for the dyslexia sample compared to the two controls, the individuals in the dyslexia sample may represent garden-variety poor readers as opposed to the discrepancy-based dyslexia. According to Nicolson and Fawcett (2006), the static cerebellar tasks of postural stability and muscle tone have the diagnostic power to distinguish between discrepancy-based dyslexia versus garden-variety poor readers. Nicolson and Fawcett's (2006) claim was based merely on one 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). Nonetheless,

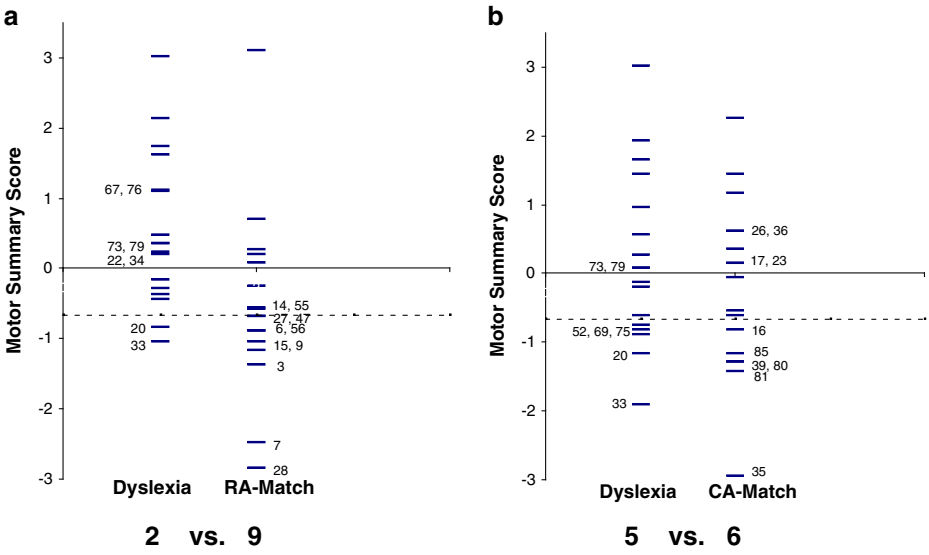

Fig. 7 Graph of individual performance for motor summary score

the findings of Fawcett et al. (2001) have not been replicated. Savage (2007) tested the claim of Fawcett et al. (2001) in a study comparing 25 children with a formal diagnosis of developmental dyslexia and 18 children with a formal diagnosis of intellectual disabilities (ages in sample ranging from 11 to 14 years). The two groups only differed on verbal and nonverbal cognitive disabilities. Savage (2007) found no group differences in spelling, word reading, phonological processing, or basic verbal response speed as well as in postural stability as measured by the DST. In fact, based on the DST criterions, both groups were identified as being at "no risk" on this measure.

Given that Nicolson and Fawcett's (2006) claim regarding the static cerebellar tests does not seem to be well supported, another explanation for the null findings in the present study may be related to the differing method of measurement for cerebellar tasks in this study as opposed to the earlier studies. It is possible that using more sensitive measurements such as accelerometer sensor or goniometer used in this study as opposed to observer ratings and Likert scales used in Fawcett and Nicolson's early body of research (e.g., Fawcett & Nicolson, 1999; Fawcett & Nicolson, 1995a; Fawcett et al., 1996) show that there may be no real cerebellar deficit. Indeed, findings from some other studies (e.g., Brown et al., 1985; Stoodley et al., 2006) that have used sensitive measures to assess postural stability are consistent with the present results. The present findings differed, however, from those obtained by Moe-Nilssen et al. (2003) who also used an accelerometer to assess postural stability in their study. Nonetheless, the impaired balance found in their study was exclusive to tests of undisturbed balance with eyes open and not when eyes were closed. In addition, as discussed earlier, the study of Moe-Nilssen et al. (2003) is potentially limited in not involving a reading-age match design.

Of additional interest were the findings for motor and toe tapping components. For both measures, neither of the group differences found was in the direction predicted by the CDT. Main analysis as well as comparisons using Cohen's *d* effect size metric indicated that in fact participants in the dyslexia group were significantly better than the younger normal readers. A similar trend was also found in the studies of Savage et al. for postural stability

and bead threading measures (Savage, 2005a, 2007). By using a reading-level design that included both RA- and CA-matched control groups, findings for the motor and toe tapping components were more consistent with the notion of developmental maturity than the CDT, as was also suggested by Savage et al. (2005a) for their postural stability results.

Beyond the data on the cerebellar and motor skills most directly related to the CDT hypothesis, the present study replicated and extended well-reported findings on the role of phonological processing in dyslexia. Phonologically based measures of elision and nonword reading (especially polysyllable nonwords) successfully differentiated between the reader groups in the sample. Both the dyslexia and the RA-matched groups performed significantly poorer than the CA-matched group in nonword decoding and phoneme deletion. The modest (0.59) and large (0.74) effect sizes for nonword reading obtained in this study were also comparable to overall combined effect sizes reported in some of the systemic reviews for nonword reading deficit (e.g., Herrmann et al., 2006; van IJzendoorn & Bus, 1994). Overall, findings for nonword reading and phoneme deletion in the present research were similar to many other studies that have also used reading-level designs (e.g., Ackerman & Dykman, 1993, Bowey, Cain, & Ryan, 1992; Bruck, 1992; Bowey & Hansen, 1994; Gillon & Dodd, 1994; Savage et al., 2005a) and are consistent with a PCD account of dyslexia.

Question 2 Does the cerebellar deficit provide a good explanatory model at the individual level?

The CDT theory predicts that motor and cerebellar-related measures should also be successful in distinguishing between participants in the dyslexia group and those in both the CA- and RA-matched controls at an individual level. This was, nonetheless, not supported by the findings of the present research. At first glance, only muscle tone seemed to be moderately successful in differentiating between participants in the dyslexia group and those in the RA- and CA-matched controls. Further exploration of the extreme poor cases in the dyslexia group indicated that poor performances on these measures were not necessarily accompanied by poor performance on literacy measures. The pattern of findings obtained for motor and toe tapping measures also seemed to follow those obtained from group contrasts and support the developmental maturation notion. In fact, for the toe tapping measure, individuals in the dyslexia group seemed not only to be better than those in RA-matched control but also those in the CA-matched control. Overall, deficits in motor and cerebellar measures did not seem to be common among participants in the dyslexia group in the present sample. 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 included an RA-matched control group in addition to a CA-matched control group that was used by White et al. (2006) and the findings were still replicated.

Similar to the findings of White et al. (2006), comparisons against the CA-matched control in the present study indicated that phonologically based measures successfully differentiated between participants in the dyslexia group and their same-aged peers at an individual level. A similar pattern was also found for word reading efficiency. 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 group and those in the RA-matched control though they were reliable differentiators of good and poor readers of the same age.

Limitations

At least two classes of limitation should be noted in the present study: those related to sample and those related to measures used.

Sample-related limitations Information on method of instruction used in schools was not available in the present study and could, therefore, not be controlled in the analyses. Usually, many studies on dyslexia, including those reviewed in this study, share this limitation. 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 (Savage & Deault, 2010).

Additionally, given that bilingualism/multilingualism is a common phenomenon in the city where the study took place, participant recruitment was also affected by factors that played an important role in limiting the number of potential Anglophone or *English-dominant bilingual* participants available to this study. Caution should also be advised in terms of generalizability of results given that the sample is not entirely monolingual. Nonetheless, it is important to highlight the fact that bilingualism/multilingualism is different within the context of this city, from where the present sample were drawn to that, for example, in the US where a child grows up learning his/her mother tongue as L1 (e.g., Spanish) and then learns English later as L2. Here, many children (most in fact) grow up in bilingual homes with exposure to English and French from birth. Hence, most children in this city can probably be considered equal status bilinguals at birth because they grow up equally strong in both languages. More importantly, every attempt was made to target the English-speaking population through the available English school boards and limiting school selection to those in which English was the leading language of instruction during elementary school years in order to obtain what are referred to in this city as *English-dominant bilinguals* where more exposure to English than French is found in their background. Consequently, the targeted sample in this project included mostly Anglophones with typical general language skills in English.

Another limitation may be concerning the age of participants in the RA-matched control. One could argue that the children in the present sample may be too young for exposure to literacy tasks. As Rack et al. (1992) have noted in their systematic review, for example, it may be too early to expose young 7-year-old readers to nonword decoding tests since these readers may experience developmentally normal difficulty with reading nonwords. This in turn can result in obtaining a smaller effect size for nonword reading deficit. Nonetheless, it is noteworthy that the present study did not seem to be affected by this factor. This was indicated by the magnitude of effect sizes for nonword reading, including all items (0.59) and polysyllables only (0.74), which were comparable to the overall combined effect size of 0.65 reported in the systemic review for nonword reading deficit by Herman et al. (2006). More importantly, inspections for univariate normality for variables within each group indicated that nonword reading variables were normally distributed within the RA-matched control group. Furthermore, as was noted in inspection for normality, most of the outliers found within the RA-matched control on the elision subtest had raw scores lying more than 1 SD above the mean.

Limitations related to measures Due to time restriction and the limited time allotted to testing in school settings, 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" section, 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 are reported to be high (i.e., 0.82 for verbal IQ, 0.76 for performance IQ, and 0.82 for the two-subtest full-scale IQ) (Saklofske, Caravan, & Schwartz, 2000).

Another limitation that needs to be addressed is related to motor and cerebellar measures. The lower reliabilities obtained for postural stability compared to other measures in this study despite using a sensitive device to assess this task may 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. 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, Nwigwe, & Wong Ho, 2000).

As Ramus et al. (2006) have advised, the onus is on proponents of sensorimotor theories to improve their methods. Given that 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. Evidently, the possible inadequacy of current methods to measure cerebellar skills may also be a potential reason for the null findings often reported (Ramus et al., 2006). The inadequacy and lower reliability of these measures may have also led to our overall null findings, especially considering the restrictions regarding sample recruitment that affected our final sample size and in turn possibly the power to detect meaningful differences. Nonetheless, within the same study and with the same constraints on factors such as size or statistical power, meaningful differences were detected for phonological processes which replicate previously well-established findings. Additionally, small differences were in fact detected when effect size ratios were explored. These revealed an effect sizes of 0.48 for postural stability in contrast to the CA-matched control and 0.35 in relation to the RA-matched group, providing some support for the CDT hypothesis, though effects were half the size of those for phonologically based nonword decoding (effect size=0.74). In addition, some evidence of maturation was also found in motor measures (e.g., effect sizes of 0.64 for toe tapping measure and 0.99 for motor tasks for the RA-matched control versus dyslexia group), a finding that clearly runs against the CDT irrespective of statistical power.

Implications of findings

Clinical implications Overall, the present findings provide support for a PCD rather than a cerebellar model of dyslexia that emphasizes phonological processing as a potential cause of poor reading and of variation in typical reading. The findings of this study may have important clinical implications as to what 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 group and the CA-matched control and the clear pattern of findings obtained from case-level analyses for elision, nonword 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.

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. This is to say that training focusing on improving phonologically based processes may be crucial for prevention of future reading difficulties and for improving reading practices in poor readers, as has been shown both at behavioral (e.g., Vellutino, Fletcher, Snowling, & Scanlon, 2004) and neurological levels (e.g., Shaywitz et al., 2004; Simos et al., 2002). 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 causes underlying reading difficulties.

Conclusion

To conclude, the present findings did not provide support for a cerebellar deficit account of dyslexia both in group- and case-level analyses. The findings 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 CDT. Specifically, a cerebellar deficit account of dyslexia was not supported despite (a) 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, (b) including a reading-level design, and (c) use of a more homogeneous sample that included nonclinical, mainstream children. Motor and cerebellar tasks did not successfully differentiate between participants in the dyslexia group and those in the RA- and CA-matched controls both at a group level and at an individual level. The positive findings for the dyslexia group and the RA-matched control suggest that a phonological deficit, independent from a cerebellar deficit, may provide the best-supported description for reading difficulties in the dyslexia group. 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.

Acknowledgments We extend our gratitude to school boards, schools, and teachers who showed interest in the present study and provided their permission for conducting this research. We thank all parents and children who participated in this study. We would like to acknowledge Dr. Ronald Stringer for his contribution to the methodology of assessing motor and cerebellar measures and Dr. Paul Stapley for providing his expertise and assistance in measuring the postural stability task. Portions of this research were supported by a grant from the Social Sciences and Humanities Research Council of Canada.

References

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

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

Badian, N. A. (1993). Predicting reading progress in children receiving special help. *Annals of Dyslexia, 43*, 90–109.

Bohannon, R. W. (1987). Variability and reliability of the pendulum test for spasticity using a Cybex II isokinetic dynamometer. *Physical Therapy, 67*, 659–661.
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 healthy, 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, 100*, 101–103.
Campbell, D. T., & Stanley, J. C. (1963). *Experimental and quasi-experimental designs for field settings*. Chicago: Rand McNally College.
Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). New York: Academic.
Corp, P. (1999). *Wechsler abbreviate scale of intelligence*. San Antonio: Psychological Corporation.
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.
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.
Fawcett, A. J., & Nicolson, R. I. (1995a). Persistent deficits in motor skill of children with dyslexia. *Journal of Motor Behavior, 27*, 235–240.
Fawcett, A. J., & Nicolson, R. I. (1995b). The dyslexia early screening test. *Irish Journal of Psychology, 16*, 248–259.
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., & 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.
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, 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.
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.
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, 21*, 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.
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.

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.
Howell, D. C. (2010). *Statistical methods for psychology* (7th ed.). Belmont: Cengage Wadsworth.
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.
Jackson, N. E., & Butterfield, E. C. (1989). Reading-level match designs: Myths and realities. *Journal of Reading Behavior, 21*, 387–412.
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.
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.
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.
National Reading Panel. (2000). Teaching children to read: Reports for the subgroups. http://www.nichd.nih.gov/publications/nrp/report.htm. Accessed 19 December 2008.
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.
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. (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. (2006). Do cerebellar deficits underlie phonological problems in dyslexia? *Developmental Science, 9*, 259–262.
Nicolson, R. I., & Fawcett, A. J. (2007). Procedural learning difficulties: Reuniting the developmental disorders? *Trends in Neurosciences, 30*, 135–141.
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.
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.
Pennington, B. F., & Lefly, D. L. (2001). Early reading development in children at family risk for dyslexia. *Child Development, 72*, 816–833.
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. 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., White, S., & Frith, U. (2006). Weighing the evidence between competing theories of dyslexia. *Developmental Science, 9*, 265–269.
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.
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.
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. (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. S., & Deault, L. (2010). 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.), *International handbook of psychology in education* (pp. 569–608). Bingley: Emerald.

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., & Pompey, Y. (2008). What does the evidence say about effective literacy teaching? *Educational and Child Psychology, 25*, 17–26.

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.

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.

Shipley, R. E., & Harley, R. J. (1971). A device for estimating stability of stance in human subjects. *Psychophysiology, 7*, 287–292.

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.

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. J. (1981). Phonemic deficits in developmental dyslexia. *Psychological Research, 43*, 219–234.

Snowling, M. (1998). Dyslexia as a phonological deficit: Evidence and implications. *Child Psychology & Psychiatry Review, 3*, 4–11.

Steven, J. (1996). *Applied multivariate statistics for social sciences* (3rd ed.). New Jersey: Lawrence Erlbaum.

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.

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.

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.

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.

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. (1999). *Comprehensive test of phonological processing*. Austin: PRO-ED.

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: Wide Range, Inc.

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

Woodcock, R. W., & Johnson, M. B. (1989). *Woodcock–Johnson psycho-educational battery—Revised (WJ-R)*. Allen: Developmental Learning Materials.