WISC-III Cognitive Profiles in Children with Developmental Dyslexia: Specific Cognitive Disability and Diagnostic Utility

Abstract
This study analyzed the usefulness of Wechsler Intelligence Scale for Children (WISC-III) in identifying specific cognitive impairments that are linked to developmental dyslexia (DD) and the diagnostic utility of the most common profiles in a sample of 100 Portuguese children (50 dyslexic and 50 normal readers) between the ages of 8 and 12. Children with DD exhibited significantly lower scores in the Verbal Comprehension Index (except the Vocabulary subtest), Freedom from Distractibility Index (FDI) and Processing Speed Index subtests, with larger effect sizes than normal readers in Information, Arithmetic and Digit Span. The Verbal-Performance IQs discrepancies, Bannatyne pattern and the presence of FDI, ACID and SCAD profiles (full or partial) in the lowest subtests revealed a low diagnostic utility. However, the receiver operating characteristic (ROC) curve and the optimal cut-off score analyses of the composite ACID, FDI and SCAD profile scores showed moderate accuracy in correctly discriminating dyslexic readers from normal ones. These results suggested that in the context of a comprehensive assessment, the WISC-III provides some useful information about the presence of specific cognitive disabilities in DD.

Keyword: Developmental dyslexia, WISC-III, cognitive profiles, children.
Practitioner Points:

- Children with DD revealed significant deficits in the WISC-III subtests that rely on verbal abilities, processing speed and working memory.
- The composite ACID, FDI and SCAD profile scores showed moderate accuracy in correctly discriminating dyslexics from normal readers.
- WISC-III may provide some useful information about the presence of specific cognitive disabilities in DD.

Introduction

Developmental dyslexia (DD) is one of the most common learning disabilities (LD), affecting approximately 5% of school-age children (Ramus, 2003) and leading to substantially lower reading performance than expected according to the child’s chronological age, intelligence and school grade (American Psychiatric Association, 2000). DD can be conceptualized as a specific LD that is neurobiological in origin and is characterized by difficulties with accurate and/or fluent word recognition as well as poor spelling and decoding abilities (International Dyslexia Association, 2002; Lyon, Shaywitz, & Shaywitz, 2003). Deficits in the phonological domain have consistently been found to be the primary cause of this disorder (see for a review: Fletcher, 2009; Ramus, 2003; Snowling, 2000; Vellutino, Fletcher, Snowling, & Scanlon, 2004), although other cognitive deficits, such as in working memory (Berninger, Raskind, Richards, Abbott, & Stock, 2008; Swanson, 1999, 2011), executive functions (Altemeier, Abbott, & Berninger, 2008; Brosnan et al., 2002; Helland & Asbjørnsen, 2000; Reiter, Tucha, & Lange, 2005), processing speed (Shanahan et al., 2006; Thomson, 2003; Willcutt, Pennington, Olson, Chhabildas, & Hulslander, 2005) and attention (Marzocchi, Ornaghi, & Barboglio, 2009) have also been linked to DD.
According to the Fourth Edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV; American Psychiatric Association, 2000) criteria, intellectual assessment may play an important role in diagnosis because IQ has to be at least normal and there has to be a significant discrepancy between actual reading ability and intellectual ability. Although some authors have argued that intelligence tests are not necessary for the definition of LD (Siegel, 1989, 1992), others support the idea that the discrepancy between achievement and intelligence is important to the concept of LD (Meyen, 1989; Torgesen, 1989). In the context of a comprehensive psychological assessment, the Wechsler Intelligence Scale for Children (WISC, WISC-R, WISC-III and WISC-IV; Wechsler, 1949, 1974, 1991, 2003a) is the most frequently used, not only to exclude intellectual impairments that could explain reading difficulties, but also to analyze specific cognitive deficits that may be useful for diagnosis.

Because specific cognitive deficits are linked to DD, it can be expected that dyslexic children would show weakness in some subtests of the Wechsler scales. The phonological theory postulates that children with DD have a specific impairment in the representation, storage and/or retrieval of speech sounds, and this impairment plays a central and causal role in this disorder (Ramus et al., 2003). The temporary storage of material that has been read is dependent on working memory, and working memory impairments have been related to specific characteristics of children with DD (Beneventi, Tønnessen, Erslan, & Hugdahl, 2010; Fiorello, Hale, & Snyder, 2006; Kibby & Cohen, 2008; Swanson, 1999, 2011). The Digit Span and Arithmetic subtests require processes from the phonological loop and the central executive of Baddeley’s (1992, 2002, 2003) working memory model. The forward Digit Span is frequently used as a measure of the phonological loop, whereas the backward Digit Span measures the executive system (Rosenthal, Riccio, Gsanger, & Jarratt, 2006). Several studies have shown that children with DD or other LDs exhibit lower performance in the Digit Span and Arithmetic subtests (Daley & Nagle, 1996; Helland & Asbjørnsen, 2004;
Mayes, Calhoun, & Crowell, 1998; Rotsika et al., 2009; Thomson, 2003; Ward, Ward, Hatt, Young, & Mollner, 1995; or in the Working Memory Index from WISC-IV (Clercq-Quaegebeur et al., 2010). Coding and Symbol Search are also two subtests in which some children with DD showed impairment (Prifitera & Dersh, 1993; Shanahan et al., 2006; Thomson, 2003; Willcutt et al., 2005), as well the Information and Vocabulary subtests (Daley & Nagle, 1996) because of the relationship between intelligence and reading development known as the “Matthew effect” (Cain & Oakhill, 2011; Stanovich, 1986). As Clercq-Quaegebeur et al. (2010) stated, with less exposure to text, these children fail to build a large lexicon and enrich their general knowledge.

The discrepancy between Verbal IQ (VIQ) and Performance IQ (PIQ) in LD samples has been analyzed in a large number of studies (e.g., Daley & Nagle, 1996; Riccio & Hynd, 2000; Rotsika et al., 2009; Rourke, 1998; Slate, 1995). Although some studies have suggested that a significant VIQ-PIQ difference may be an important indicator of LD (Riccio & Hynd, 2000; Rourke, 1998), others did not find VIQ-PIQ differences to be useful in differentiating children with LDs from other groups of children (Humphries & Bone, 1993; Kavale & Forness, 1984).

As a result of the cognitive impairment observed in those with DD and LDs, many studies have tried to identify specific WISC profiles [e.g., Bannatyne pattern, Developmental Index, Learning Disabilities Index, Successive and Simultaneous Processing, Freedom from Distractibility Index (FDI), a pattern of low scores on the Arithmetic, Coding, Information and Digit Span subtests (ACID) or on the Symbol Search, Coding, Arithmetic and Digit Span subtests (SCAD)], although inconsistent results were obtained. A study about the usefulness of the WISC-III in the context of psychological assessment found that a total of 89% of school psychologists used profile analysis, and almost 70% listed it as among the most
beneficial features (Pfeiffer, Reddy, Kletzel, Schmelzer, & Boyer, 2000). In the present study, only the most common profiles were analyzed: Bannatyne pattern, FDI, ACID and SCAD.

Bannatyne (1968) suggested that WISC subtest scores could be re-categorized to identify children with LD. He argued that rather than relying on the traditional VIQ and PIQ, WISC subtest scores could be re-categorized into four composite scores: spatial abilities (Block Design, Object Assembly and Picture Completion), conceptual abilities (Vocabulary, Similarities and Comprehension), sequential abilities (Digit Span, Coding and Arithmetic) and acquired knowledge (Information, Arithmetic and Vocabulary). Bannatyne (1971) reported that disabled readers exhibited a specific pattern: spatial abilities > conceptual abilities > sequential abilities. Subsequent studies found support for Bannatyne’s classification system on WISC and WISC-R (Clarizio & Bernard, 1981; Rugel, 1974; M. D. Smith, Coleman, Dokecki, & Davis, 1977), whereas others demonstrated its limited diagnostic validity (D’Angiulli & Siegel, 2003; Henry & Wittman, 1981; Kavale & Forness, 1984; McKay, Neale, & Thompson, 1985; Vance & Singer, 1979). Kaufman (1981) stated that although some studies reported statistically significant mean differences in the composite scores between LD and controls, the proportions of individuals in the LD group displaying the Bannatyne pattern are quite small and their contribution to differential diagnosis is limited.

Some studies explored the utility of the Bannatyne pattern with WISC-III in dyslexic and LD samples. The first work, by Prifitera and Dersh (1993), compared the baseline rates of the Bannatyne WISC-III pattern in three groups of children: those with LD, those with Attention Deficit Hyperactivity Disorder (ADHD) and those without disabilities. They found baseline rates of 33% for children with LD, 47% for children with ADHD, and 14% for children without disabilities. Although Ho, Gilger, and Decker (1988) found that this pattern was reliable and specific to their dyslexic twin sample, Smith and Watkins (2004) suggested that the use of the Bannatyne WISC-III pattern is not recommended because they only found a
sensitivity of 22.4% and a specificity of 86.1% for the LD group, a sensitivity of 24% and a specificity of 86.1% for the dyslexic group, and 13.9% of false-positives were identified in the normative group.

Freedom from Distractibility (FD) was identified in a factor analysis of WISC-R and includes the Arithmetic, Coding, and Digit Span subtests (Kaufman, 1975; Reynolds & Kaufman, 1990). With the publication of WISC-III, four factors were included: Verbal Comprehension Index (VCI), Perceptual Organization Index (POI), Processing Speed Index (PSI), and FDI. The WISC-III FDI only consists of the Arithmetic and Digit Span subtests because the Coding subtest (which had previously been part of Kaufman’s FD) was included with Symbol Search in the PSI. Prifitera, Weiss, and Saklofske (1998) stated that FDI is a misleading name for this construct because it encourages naive interpretations and may be better conceptualized as an index of working memory. A considerable number of studies analyzed the significant mean score differences of FD and FDI between clinical samples and typically developing children, with inconsistent findings (Ackerman, Holloway, Youngdahl, & Dykman, 2001; Anastopoulos, Spisto, & Maher, 1994; Mayes & Calhoun, 2004; Mayes et al., 1998; Prifitera & Dersh, 1993; Slate, 1995; Snow & Sapp, 2000). Thomson (2003) showed that 80% of children with DD had significantly lower mean scores on the FDI and PSI compared to the VCI and POI, whereas Alm and Kaufman (2002) also found that POI > VCI > FD in a sample of dyslexic adults.

The addition of the Information subtest to the FD triad resulted in another WISC profile: the ACID profile. Using WISC-III standardization sample, Prifitera and Dersh (1993) found that the full ACID pattern was quite rare (only 1.1% of the children from the standardization sample showed this profile), although it was more common in the LD (5.1%) and ADHD (12.3%) samples. In a sample of children with LDs, the prevalence of the ACID profile was 4.7% (Ward et al., 1995). In a Greek dyslexic sample, the prevalence was 6.7%,
whereas the prevalence of the SCAD profile was 2.4% (Rotsika et al., 2009). Watkins, Kush, and Glutting (1997a) found sensitivities of 4% in the full profile (with a specificity of 99%) and 19% in the partial profile (with a specificity of 94%) in the dyslexic group, whereas the receiver operating characteristic (ROC) curve analysis resulted in an area-under-the-curve (AUC) value of .68. When analyzing group differences of the ACID and FDI profiles, the dyslexic group showed significantly lower scores than the normal reading group (Ackerman et al., 2001).

In analyzing the ACID profile results, Kaufman (1994) notes that the contribution of the Information subtest is minimal and that the differences between the clinical (ADHD and LD) and nonclinical (typically developing children) groups are largely attributable to the subtests comprising the FDI and PSI. He suggested the use of the SCAD profile because it is less vulnerable to contamination from school learning (Information subtest). Ward et al. (1995) also examined the frequency of SCAD profiles in their LD sample and obtained 19.6% true-positives and 16% false-positives. In the Daley and Nagle (1996) LD sample, the full SCAD profile was observed in 2% of the sample (partial profile was observed in 8%), the full ACID profile was observed in 1% of the sample (partial profile was observed in 12%), and the Bannatyne pattern was observed in 26% of the subjects. The mean SCAD and ACID scores were significantly different from the mean scores of the remaining subtests. Other studies showed more diagnostic utility. For instance, Thomson (2003) found that 40% of children with DD displayed a complete ACID profile, and 50% displayed a complete SCAD profile. Considering the subtest-level data, 68% presented the lowest scores on Digit Span and Coding, and 62% presented the lowest scores on Coding, Digit Span, and Symbol Search.

Thus, the results of empirical studies have shown a large cognitive variability and an inconsistency in identifying a specific profile. This variability might be related to the definition of DD used (e.g., discrepancy criterion, reading achievement criterion, response to
intervention criterion), sample characteristics (e.g., clinical, school-referred), selection criteria (e.g., cut-off scores, comorbidity), assessment measures (e.g., IQ, reading, spelling, phonological processing), and others. A large body of studies has been conducted on English-speaking samples, but it is also particularly important to analyze the presence of such profiles in samples with native languages other than English (some exceptions: Clercq-Quaegebeur et al., 2010; Filippatou & Livaniou, 2005; Rotsika et al., 2009). For example, it is known that phonological processing, reading fluency and accuracy, and the prevalence of DD subtypes are influenced by specific linguistic characteristics (Boets et al., 2010; Jiménez, Rodríguez, & Ramírez, 2009; Sprenger-Charolles, Colé, Lacert, & Serniclaes, 2000). The present study is an extension of previous studies analyzing the usefulness of WISC-III in identifying the specific cognitive impairments that are associated with DD and the diagnostic utility of the most common profiles. It makes a unique contribution by using a sample of Portuguese children (no similar studies in European Portuguese orthography were found) and performs an optimal cut-off score analysis (the few studies that previously used ROC curve did not compute this type of analysis). The study had the following goals: (i) to analyze the discrepancy between VIQ and PIQ; (ii) to identify characteristic patterns of subtest strengths and weakness in children with DD; and (iii) to analyze the discriminant power of the most common WISC-III profiles through sensitivity-specificity values, ROC curve analysis and optimal cut-off scores.

Method

Participants

Participants were 100 Portuguese children between the ages of 8 and 12 ($M = 9.81; SD = 1.34$) in the 3rd to 6th school grades. In the dyslexic group ($N = 50$), 74% were male and 26% were female, with a gender ratio of 2.8 (clinical based sample). A recent population
based study found a prevalence of DD in school age Portuguese children of 5.4%, with a gender ratio of 1.5 (Vale, Sucena, & Viana, 2011). This is consistent with the hypothesis that the prevalence of boys with DD is significantly higher in referred or clinical samples than in population samples (e.g., Hawke, Olson, Willcut, Wadsworth, & DeFries, 2009; Rutter et al., 2004; Shaywitz, Shaywitz, Fletcher, & Escobar, 1990; Wadsworth, DeFries, Stevenson, Gilger, & Pennington, 1992). The mean age of the dyslexic group was 9.80 years with a standard deviation of 1.38 years ($N_{8y} = 9, N_{9y} = 17, N_{10y} = 8, N_{11y} = 7, N_{12y} = 9$). Twenty-six percent of children with DD had school retention, and 36% were participants in special education systems (the Portuguese special education system establishes for children with DD the possibility of individual curriculum adjustment, adjustment in the assessment process, and personalized pedagogical support with a specialized teacher). Ninety-four percent had attended kindergarten, and 30% have relatives with reading difficulties. In the normal reader group ($N = 50$), 64% were male and 36% were female, with a mean age of 9.82 years and a standard deviation of 1.32 years ($N_{8y} = 7, N_{9y} = 19, N_{10y} = 8, N_{11y} = 8, N_{12y} = 8$). All normal readers attended kindergarten, only 2% had school retention, and 4% have relatives with reading difficulties. No statistically significant differences were found between groups with regard to gender $\chi^2(1) = 1.169, p = .387$, age $\chi^2(4) = 0.487, p = .975$ and school grade $\chi^2(3) = 1.776, p = .620$.

Criteria for Inclusion. For both groups, only children with the following criteria were included: (i) WISC-III Full Scale IQ (FSIQ) $\geq 90$; (ii) native speakers of European Portuguese; (iii) at least two years of school attendance; (iv) absence of a visual, hearing or motor impairment; and (v) exclusion of a language impairment, emotional disturbance, dyscalculia, disruptive behavior disorder (ADHD, oppositional defiant disorder and conduct disorder), neurological impairment or other psychiatric disorders. These children were not included in order to ensure that cognitive deficits were not associated with any of these
disorders. For the normal reader group, children with special educational needs were also excluded.

In the dyslexic group, only children who were previously diagnosed with DD by a psychologist, child psychiatrist, developmental pediatrician or child neurologist, and simultaneously having a score lower than or equal to the 15th percentile in a reading fluency and accuracy test («O Rei» Assessment Test of the Reading Fluency and Precision; Carvalho & Pereira, 2009) administered during the testing session were included. These cut-off score criteria (WISC-III FSIQ ≥ 90 and both reading fluency and accuracy measures ≤ 15th percentile) are similar to (and in some cases stricter than) the inclusion criteria used by several other authors (e.g., Ackerman et al., 2001; Frijters et al., 2011; Reiter et al., 2005; Siegel, 1992; Siegel & Ryan, 1989; Stanovich & Siegel, 1994; Swanson, 1999, 2011).

Measure

The WISC-III (Wechsler, 1991) is an individually administered intelligence test, including 13 subtests ($M = 10; SD = 3$), for children between the ages of 6 to 16 that measures different intellectual abilities and yields three composite IQs scores ($M = 100; SD = 15$): VIQ, PIQ and FSIQ; and four index scores: VCI, POI, PSI and FDI.

All participants were tested with the Portuguese version of WISC-III (Wechsler, 2003b), which was normed on a representative sample of 1354 children. The factor structure of the Portuguese version of WISC-III, analyzed through an exploratory and confirmatory factor analysis, yielded a three-factor model (VCI, POI and PSI). Thus, in this study, the FDI was analyzed as a profile (sum of the scaled scores of Arithmetic and Digit Span) rather than as an index score. The Mazes subtest was not administered.

Procedures
WISC-III administration was included as part of a broad neuropsychological protocol that also comprised a neuropsychological battery as well as reading and spelling measures. The testing was conducted in two sessions (with an interval of 10 to 15 days), lasting approximately 90 minutes per session, in a clinic or school setting during a weekday. The WISC-III was administered during the first session. All measures were administered by the first author in a standard order. No incentives were offered in exchange for participation.

**Data Analysis**

Data were analyzed using the Statistical Package for Social Sciences (SPSS 19.0). Independent, paired and one-sample t-tests, repeated measures and multi-factor ANOVA were calculated to investigate the significance of differences in WISC-III IQs, index scores, subscales and profiles between groups. Cohen’s $d$ or eta squared ($\eta^2$) were additionally calculated to determine the effect sizes of these differences. According to Cohen’s (1988) criteria, $d$ effect sizes are considered to be large if exceeding 0.80, moderate if at 0.50, and small if less than 0.20; whereas for $\eta^2$, .01 constitutes a small effect, .06 a medium effect and .14 a large effect.

ROC curve analysis was performed to examine the differential discriminatory power of WISC-III profiles for the diagnosis of DD. ROC curve analysis systematically sweeps across all possible true positive (sensitivity) and false positive (1-specificity) values of a diagnostic test, graphically illustrates the test’s full range of diagnostic utility and calculates the AUC, which provides an accuracy index for the test (Fawcett, 2006; McFall & Treat, 1999; Metz, 1978; Watkins et al., 1997a). The more accurately a test is able to discriminate between groups (children with DD vs. normal readers), the more its ROC curve will deviate toward the upper left corner of the graph. The AUC is the average of the true positive rate, taken uniformly over all possible false positive rates (Krzanowski & Hand, 2009) that range
between .5 and 1.0. An AUC value of 1.0 is perfectly accurate because the sensitivity is 1.0 when the false positive rate is .0, whereas an AUC value of .5 reflects a completely random classifier. An AUC of .5 to .7 indicates low test accuracy, .7 to .9 indicates moderate accuracy, and .9 to 1.0 indicates high accuracy (Swets, 1988).

**Results**

**IQs and Index Scores**

A paired sample t-test showed a statistically significant difference between VIQ and PIQ for normal readers $t(49) = 3.542, p = .001, d = 0.46$ with VIQ $>$ PIQ; but a non-significant difference for children with DD $t(49) = -1.651, p = .105, d = 0.32$. The absolute mean Verbal–Performance discrepancy for the dyslexic group was 3.55 ($SD = 15.05$; range = 0–37) and 5.30 ($SD = 10.57$; range = 0–32) for normal readers.

Table 1 shows the percentage of the discrepancies between VIQ and PIQ, compared with the findings of Daley and Nagle (1996) and Rotsika et al. (2009). Thirty-eight percent of normal readers and 44.9% of children with DD had a difference $\geq$ 11 points. Only 8% (VIQ $>$ PIQ) of normal readers showed a difference $\geq$ 21 points, in contrast with 20.4% of the dyslexic group (16.3% showed VIQ $<$ PIQ). Thirty percent of normal readers had a PIQ $>$ VIQ discrepancy (and 66% a VIQ $>$ PIQ), whereas among the dyslexic group, the percentage was 55.1% (and 42.9% a VIQ $>$ PIQ).

--- (Insert Table 1 about here) ---

Statistically significant differences were found between children with DD and normal readers for FSIQ and VIQ (see Table 2). The VIQ scores of normal readers were 12.86 points greater, and for FSIQ they were 9.71 points greater. For PIQ, no significant difference was found. Relative to the WISC-III index scores, children with DD showed significantly lower VCI ($p < .001$) and PSI ($p < .01$) scores.
A two-factor analysis of variance was conducted to explore the impact of group x gender, group x age, and group x school grade on the three IQs and on the three index scores, but no significant differences were found.

--- (Insert Table 2 about here) ---

**Subtests Scores**

As shown in Table 2, Vocabulary was the only subtest with no statistically significant difference from the six VIQ subtests. The largest effect sizes were observed in the Information, Arithmetic and Digit Span subtests. The lowest scores for both groups were in the Digit Span subtest, with significant differences between groups in forward and backward span. In the PIQ subtests, significant differences with moderate effect sizes were found in Object Assembly, Coding and Symbol Search. That is, children with DD showed significantly lower scores than normal readers in the subtests included in the WISC profiles (Bannatyne sequential abilities, FDI, ACID and SCAD): the Information, Arithmetic, Digit Span, Coding and Symbol Search subtests (and also in the Similarities, Comprehension and Object Assembly subtests).

A two-factor ANOVA was also conducted to explore the impact of group x gender, group x age, and group x school grade on the 12 WISC-III subtests. No significant differences were found.

**Profiles Scores**

Statistically significant differences between the dyslexic and normal reader groups were found in six of the seven analyzed WISC-III profiles (see Table 2). Very large effect sizes were observed in the composite score of Bannatyne’s sequential abilities, FDI, ACID and SCAD, in which the mean scores of the dyslexic group were 1.43SD, 1.54SD, 1.59SD and
1.25SD, respectively, below the mean scores of the group of normal readers. Once again, a two-factor analysis of variance found no significant differences for group x gender, group x age, and group x school grade on these profiles.

Additionally, we were interested in investigating how these profiles operate only in the dyslexic group. A repeated measures analysis of variance revealed significant differences between the Bannatyne’s composite scores, $F(3, 47) = 6.358, p < .001$, $\eta^2 = .293$. A comparison of main effects (Bonferroni $p < .05$) showed that spatial and conceptual abilities > sequential abilities, and conceptual abilities > acquired knowledge. The means of the FDI ($M = 8.54; SD = 1.44$), ACID ($M = 8.71; SD = 1.22$) and SCAD ($M = 9.00; SD = 1.37$) scaled scores calculated for the dyslexic group were significantly different from the mean of 10.30 ($SD = 1.37$) for the remaining subtests: FDI $t(49) = -8.514, p < .001$, $d = 1.25$; ACID $t(49) = -9.006, p < .001$, $d = 1.22$ and SCAD $t(49) = -6.614, p < .001$, $d = 0.94$.

**Discriminant Power of WISC-III Profiles**

Following the criteria of Prifitera and Dersh (1993), children were considered to be positive for the full profile when their scores on the four ACID (excluding Symbol Search and Mazes) and SCAD subtests or on the two FDI subtests were less than or equal to the scores on the remaining subtests. For the partial profile, scores on any three of the four ACID and SCAD subtests had to be less than or equal to the scores on the remaining subtests. We were also interested in investigating the presence of FDI in the three and four lowest-scoring subtests, as well of ACID and SCAD in the five and six lowest-scoring subtests.

Table 3 shows the diagnostic accuracy of WISC-III profiles. The Bannatyne pattern showed a sensitivity of .22 (i.e., 22% of the dyslexic children were correctly diagnosed) and a specificity of .90 (i.e., 90% of normal readers were classified by the Bannatyne pattern as not having a disability). The number of children displaying the Bannatyne pattern did not differ
between the dyslexic and normal reader groups, $\chi^2(1) = 1.986, p = .159$, kappa = .124. The full FDI, ACID and SCAD profiles misclassified the children with DD, only 0% to 8% of whom were properly diagnosed (true positive). The presence of ACID in the lowest six subtests and FDI in the lowest four subtests showed a greater diagnostic utility. For ACID, a sensitivity of .45, a specificity of .94, a positive predictive power of .88 (i.e., 88% of children classified as positive are children with DD) and a negative predictive power of .64 (i.e., 64% of children classified as negative are normal readers) were obtained, whereas FDI revealed a sensitivity of .39, a specificity of .90, a positive predictive power of .79 and a negative predictive power of .60. The number of children displaying ACID in the six lowest-scoring subtests differed between the dyslexic and normal reader groups $\chi^2(1) = 17.830, p < .001$, kappa = .391; as did the presence of FDI in the four lowest-scoring subtests $\chi^2(1) = 9.646, p < .01$, kappa = .289.

--- (Insert Table 3 about here) ---

These results suggested that the presence of the full Bannatyne, FDI, ACID and SCAD profiles did not efficiently distinguish between children with and without DD. However, the composite scores of these profiles showed statistically significant differences with large effect sizes between groups (see Table 2). Thus, a ROC curve analysis was performed because it is independent of prevalence rates and cut-off values (McFall & Treat, 1999; C. B. Smith & Watkins, 2004). This analysis was conducted for FDI, ACID and SCAD composite scores; the Bannatyne pattern was excluded because is a dichotomous variable (presence vs. absence).

As shown in Figure 1, the ACID and FDI ROC curves are elevated over the reference line. The AUC value for ACID was .875 ($p < .001, SE = .033, 95\%CI = .810–.941$), i.e., a randomly selected child with DD will have a lower ACID score than a randomly selected child without DD approximately 87.5% of the time. The AUC values for FDI and SCAD were .862 ($p < .001, SE = .036, 95\%CI = .792–.933$) and .809 ($p < .001, SE = .042, 95\%CI = .727–
.891), respectively. As Swets (1988) noted, these AUC values are indicative of a moderate accuracy in discriminating between dyslexic and non-dyslexic children.

--- (Insert Figure 1 about here) ---

Additionally, the Youden index (Youden, 1950) was calculated ($J = \text{sensitivity} + \text{specificity} - 1$) to analyze the optimal cut-off scores for FDI, ACID and SCAD (note that all children in the sample have a FSIQ $\geq 90$ with a mean of 103.43). For FDI, the optimal cut-off score was 17.50 ($J = .552$), yielding a sensitivity of .61 and a specificity of .94; for ACID, the optimal cut-off score was 37.50 ($J = .573$), which yielded a sensitivity of .67 and a specificity of .90; and for SCAD, a cut-off of 41.50 ($J = .437$) yielded a sensitivity of .84 and a specificity of .60.

**Discussion**

In this study, approximately 55% of children with DD had scores such that PIQ > VIQ (30% in normal readers), and 20.4% revealed a VIQ-PIQ discrepancy equal to or above 21 points (compared to 8% in normal readers and 18.1% in the WISC-III Portuguese standardization sample). Non-significant differences were found in the Verbal-Performance discrepancy, and the mean difference was smaller than for normal readers. This finding was also reported by other studies with LD (Pereira & Simões, 2005) and children with DD (Rotsika et al., 2009), and demonstrated the minimal practical value of the VIQ–PIQ discrepancy for differential diagnosis (Kaufman, 1981). The analyses of IQ mean scores between groups revealed that dyslexics had significantly lower scores in FSIQ and VIQ, which is consistent with a large number of studies (e.g., Ackerman et al., 2001; Laasonen, Leppämäki, Tani, & Hokkanen, 2009; Rotsika et al., 2009; Swartz, Gfeller, Hughes, & Searight, 1998). The VIQ comprises subtests that are more strongly associated with school learning, verbal abilities and working memory, and it was therefore expected that these tests
would be more difficult for these children. Results from the WISC-III index scores showed that, in addition to VCI, processing speed can also be an additional risk factor, whereas a non-significant difference was found in POI, which may indicate that nonverbal reasoning ability is not compromised in these children with DD. These findings were also reported by previous studies (Ackerman et al., 2001; Prifitera & Dersh, 1993; Thomson, 2003).

At the WISC-III subtest level, inferential analysis showed that the six subtests with the largest effect sizes were Information, Arithmetic, Digit Span, Similarities, Coding and Symbol Search, with the dyslexic children exhibiting significantly lower scores. Significant differences were also found in Comprehension and Object Assembly (the only subtest from POI). Surprisingly, Vocabulary was the subtest with the second-highest scores among children with DD, and no statistical differences were found in comparison with normal readers (although the $p$-value was closer to statistical significance), most likely because some children with DD received a direct intervention from the special education system, may have had adequate cultural opportunities at home and may have revealed some intellectual curiosity for particular topics, thereby minimizing the impact of the low vocabulary knowledge that is common in these children. A similar finding was also reported in samples of dyslexic (Thomson, 2003) and LD children (Mayes et al., 1998). Contrary to the results reported by other studies (Clercq-Quaegebeur et al., 2010; Rotsika et al., 2009; Thomson, 2003), Similarities was one of the lowest subtests among children with DD. Similarities subtest requires greater demands on verbal abstract reasoning skills, that seems to be diminished in our dyslexic group. This unexpected finding was also observed in two studies with Portuguese learning disabled children, which included a subgroup of children with DD (Cardoso, 2007; Pereira & Simões, 2005). It would be particularly relevant clarify in subsequent studies if Similarities is (or it is not) also a “problematic” subtest for Portuguese children with DD.
The six subtests with the lowest scores for the dyslexic group (in order from lowest to highest) were Digit Span, Information, Similarities, Symbol Search, Picture Completion and Arithmetic. As reported by a large number of studies, dyslexics typically exhibit lower performance in Arithmetic, Digit Span, Information, Coding and Symbol Search (Mayes et al., 1998; Rourke, 1998; Thomson, 2003; Ward et al., 1995). These results seem to suggest that Portuguese children with DD demonstrated difficulties in the same subtests linked to specific cognitive deficits that other international studies have shown are impaired in DD.

Although inferential analysis showed that normal readers outperformed dyslexics in a large number of WISC-III subtests, the scaled scores of the dyslexic group were within norm (the exception was Digit Span), and therefore at risk of being unobserved in clinical assessment. Thus, the analysis of the most common WISC-III profiles linked to DD may provide additional diagnostic information beyond the subtest-level analyses. Relative to the WISC-III re-categorizations, Prifitera and Dersh (1993) stated that the Bannatyne pattern is useful for diagnostic purposes, but our results did not support their conclusion. The Bannatyne pattern was present only in 22% of dyslexic children and in 10% of the normal reader group. Smith and Watkins (2004) also reported similar percentages in their sample of children with DD and LD. Large, significant differences in FDI, ACID and SCAD profiles were also found, with children with DD scoring at least 1.25 SD below normal readers.

Elwood (1993) stated that the presence of a significant difference alone does not imply that the test can discriminate among subjects with sufficient accuracy for clinical use. We therefore additionally performed an analysis of the discriminant power (sensitivity-specificity values, ROC curve and optimal cut-off scores) of these three profiles.

Although the presence of full or partial FDI, ACID and SCAD profiles was more prevalent among dyslexics than among normal readers, the sensitivity and specificity values revealed a low diagnostic accuracy. However, when we analyzed the mean of the composite
scores, moderate accuracy was obtained. A randomly selected child with DD will have a lower FDI, ACID and SCAD score than a randomly select normal reader approximately 86.2%, 87.5% and 80.9% of the time, respectively. These results from the ROC curve analysis revealed a higher diagnostic accuracy than the findings reported by Watkins et al. (1997a; 1997b) in LD samples. One of the particularities of this study was the analysis of optimal cut-off scores for FDI, ACID and SCAD, because previous studies did not perform such analysis. For FDI, a score less than or equal to 17.50 correctly identified 61% of dyslexic children (6% false positives); for ACID, a score less than or equal to 37.50 correctly identified 67% of dyslexic children (10% false positives); and for SCAD, a score less than or equal to 41.50 correctly identified 84% of dyslexic children (40% false positives). The results from the optimal cut-off scores analysis showed greater diagnostic utility than the presence of full or partial profiles. New studies are needed to explore and compare the diagnostic accuracy of these and others cut-off scores in dyslexic samples. Compared to the other profiles, ACID showed a higher discriminant power.

In sum, our findings from a sample of Portuguese children were also consistent with previous studies that found that VIQ-PIQ discrepancies, the Bannatyne pattern and the presence of the FDI, ACID and SCAD profiles in the lowest-scoring subtests do not efficiently distinguish children with DD from those without DD. However, the composite scores of FDI, ACID and SCAD profiles showed greater diagnostic utility and subtest-level analyses may provide useful information beyond the global scores about the presence of specific cognitive impairments in children with DD. In clinical practice, weakness on a specific profile is not a sufficient diagnostic criterion for dyslexia; conversely, the lack of this profile should not exclude the possibility of dyslexia (Clercq-Quaegebeur et al., 2010; Thomson, 2003). As Mayes and Calhoun (2004, p. 566) asserted, “the presence or absence of profile types certainly should not be the basis for making a diagnosis. Profiles are clinically
useful because they may alert a clinician to certain diagnostic possibilities and they provide knowledge about the pattern of strengths and weaknesses that characterize certain disorders”. IQ tests yield information that is only a component of the DD diagnosis and decision-making process (Prifitera et al., 1998) and need to be viewed in the context of a more comprehensive assessment that must include other tests, such as phonological awareness, rapid naming, working memory, reading and spelling measures, executive functions, and others.

The present study revealed some limitations that needed be addressed in future research. First, the pattern of WISC-III subtests scores found in children with DD was only compared to a control group and did not include other clinical samples (e.g., ADHD). Studies have found that such profiles also have some diagnostic utility in ADHD children (Mayes & Calhoun, 2004; Mayes et al., 1998; Prifitera & Dersh, 1993; Swartz et al., 1998). Second, the two groups were not matched for WISC-III FSIQ. This additional inclusion criterion would have been a better baseline to compare cognitive profiles differences between groups. Third, we established a strict cut-off score criterion for WISC-III FSIQ (≥ 90) in order to decrease Type I error (false positive). Obviously, such strict criterion increased Type II error (false negative), excluding from the sample some children with DD that had a WISC-III FSIQ lower than 90. Another limitation was the fact that this study did not analyze the effects of socioeconomic status or parental educational attainment.

References


Table 1

Percentages of Discrepancies between VIQ and PIQ in Normal Readers and Children with DD

<table>
<thead>
<tr>
<th>Number of Points</th>
<th>Normal Readers</th>
<th></th>
<th></th>
<th>Dyslexic</th>
<th></th>
<th></th>
<th>Daley &amp; Nagle</th>
<th>Rotsika et al.</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>4.0</td>
<td>4.0</td>
<td>2.0</td>
<td>2.0</td>
<td>---</td>
<td>4.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 – 5</td>
<td>14.0</td>
<td>18.0</td>
<td>32.0</td>
<td>12.2</td>
<td>14.3</td>
<td>26.5</td>
<td>29.5*</td>
<td>33.9</td>
</tr>
<tr>
<td>6 – 10</td>
<td>10.0</td>
<td>16.0</td>
<td>26.0</td>
<td>14.3</td>
<td>12.2</td>
<td>26.5</td>
<td>23.1</td>
<td>23.3</td>
</tr>
<tr>
<td>11 – 15</td>
<td>4.0</td>
<td>14.0</td>
<td>18.0</td>
<td>6.1</td>
<td>8.2</td>
<td>14.3</td>
<td>20.1</td>
<td>18.3</td>
</tr>
<tr>
<td>16 – 20</td>
<td>2.0</td>
<td>10.0</td>
<td>12.0</td>
<td>6.1</td>
<td>4.1</td>
<td>10.2</td>
<td>18.9</td>
<td>13.3</td>
</tr>
<tr>
<td>≥ 21</td>
<td>0</td>
<td>8.0</td>
<td>8.0</td>
<td>16.3</td>
<td>4.1</td>
<td>20.4</td>
<td>11.4</td>
<td>6.7</td>
</tr>
</tbody>
</table>

Note. * This percentage value relates to a discrepancy ranging from 0 to 5 points.
### Table 2

*Mean WISC-III Scores and Standard Deviations for Normal Readers and Children with DD*

<table>
<thead>
<tr>
<th></th>
<th>Normal Readers</th>
<th>Dyslexic</th>
<th>t(98)</th>
<th>p</th>
<th>d</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>IQs</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FSIQ</td>
<td>108.24 ± 11.64</td>
<td>98.53 ± 8.55</td>
<td>4.721</td>
<td>&lt;.001</td>
<td>0.95</td>
</tr>
<tr>
<td>VIQ</td>
<td>109.98 ± 11.20</td>
<td>97.12 ± 10.72</td>
<td>5.831</td>
<td>&lt;.001</td>
<td>1.17</td>
</tr>
<tr>
<td>PIQ</td>
<td>104.68 ± 11.91</td>
<td>100.67 ± 11.42</td>
<td>1.707</td>
<td>.091</td>
<td>0.34</td>
</tr>
<tr>
<td><strong>Index Scores</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VCI</td>
<td>108.90 ± 11.36</td>
<td>97.71 ± 11.14</td>
<td>4.942</td>
<td>&lt;.001</td>
<td>0.99</td>
</tr>
<tr>
<td>POI</td>
<td>104.04 ± 12.48</td>
<td>102.12 ± 11.95</td>
<td>0.780</td>
<td>.437</td>
<td>0.15</td>
</tr>
<tr>
<td>PSI</td>
<td>105.98 ± 14.75</td>
<td>97.22 ± 12.07</td>
<td>3.228</td>
<td>&lt;.01</td>
<td>0.65</td>
</tr>
<tr>
<td><strong>Subtests</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Information</td>
<td>10.74 ± 2.38</td>
<td>8.18 ± 1.93</td>
<td>5.859</td>
<td>&lt;.001</td>
<td>1.18</td>
</tr>
<tr>
<td>Similarities</td>
<td>11.70 ± 2.73</td>
<td>9.22 ± 2.45</td>
<td>4.739</td>
<td>&lt;.001</td>
<td>0.95</td>
</tr>
<tr>
<td>Arithmetic</td>
<td>12.10 ± 2.04</td>
<td>9.45 ± 2.11</td>
<td>6.348</td>
<td>&lt;.001</td>
<td>1.17</td>
</tr>
<tr>
<td>Vocabulary</td>
<td>11.72 ± 2.14</td>
<td>10.84 ± 2.59</td>
<td>1.847</td>
<td>.068</td>
<td>0.37</td>
</tr>
<tr>
<td>Comprehension</td>
<td>11.92 ± 2.53</td>
<td>10.61 ± 2.37</td>
<td>2.648</td>
<td>&lt;.1</td>
<td>0.53</td>
</tr>
<tr>
<td>Digit Span (DS)</td>
<td>9.76 ± 2.26</td>
<td>7.63 ± 1.75</td>
<td>5.221</td>
<td>&lt;.001</td>
<td>1.05</td>
</tr>
<tr>
<td>DS Forward*</td>
<td>7.36 ± 1.45</td>
<td>6.20 ± 1.13</td>
<td>4.403</td>
<td>&lt;.001</td>
<td>0.89</td>
</tr>
<tr>
<td>DS Backward*</td>
<td>4.56 ± 1.34</td>
<td>3.63 ± 0.97</td>
<td>3.929</td>
<td>&lt;.001</td>
<td>0.80</td>
</tr>
<tr>
<td>Picture Completion</td>
<td>10.16 ± 2.51</td>
<td>9.43 ± 2.70</td>
<td>1.394</td>
<td>.166</td>
<td>0.28</td>
</tr>
<tr>
<td>Picture Arrangement</td>
<td>11.18 ± 2.37</td>
<td>11.78 ± 2.57</td>
<td>-1.197</td>
<td>.234</td>
<td>0.24</td>
</tr>
<tr>
<td>Block Design</td>
<td>10.44 ± 2.50</td>
<td>10.47 ± 2.57</td>
<td>-0.058</td>
<td>.954</td>
<td>0.01</td>
</tr>
<tr>
<td>Object Assembly</td>
<td>10.90 ± 2.96</td>
<td>9.78 ± 2.32</td>
<td>2.099</td>
<td>&lt;.05</td>
<td>0.42</td>
</tr>
<tr>
<td>Coding</td>
<td>11.12 ± 2.70</td>
<td>9.61 ± 2.45</td>
<td>2.900</td>
<td>&lt;.01</td>
<td>0.58</td>
</tr>
<tr>
<td>Symbol Search</td>
<td>10.96 ± 3.12</td>
<td>9.31 ± 2.64</td>
<td>2.840</td>
<td>&lt;.01</td>
<td>0.57</td>
</tr>
</tbody>
</table>

Profiles
<table>
<thead>
<tr>
<th>Ability</th>
<th>Mean ± SD</th>
<th>Mean ± SD</th>
<th>t-value</th>
<th>p-value</th>
<th>Cohen's d</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spatial Abilities</td>
<td>31.50 ± 6.10</td>
<td>29.67 ± 5.79</td>
<td>1.526</td>
<td>.130</td>
<td>0.30</td>
</tr>
<tr>
<td>Conceptual Abilities</td>
<td>35.34 ± 5.99</td>
<td>30.67 ± 6.04</td>
<td>3.854</td>
<td>&lt;.001</td>
<td>0.77</td>
</tr>
<tr>
<td>Sequential Abilities</td>
<td>32.98 ± 4.76</td>
<td>26.69 ± 4.00</td>
<td>7.101</td>
<td>&lt;.001</td>
<td>1.43</td>
</tr>
<tr>
<td>Acquired Knowledge</td>
<td>34.56 ± 5.23</td>
<td>28.46 ± 5.20</td>
<td>5.807</td>
<td>&lt;.001</td>
<td>1.16</td>
</tr>
<tr>
<td>FDI</td>
<td>21.86 ± 3.30</td>
<td>17.08 ± 2.89</td>
<td>7.654</td>
<td>&lt;.001</td>
<td>1.54</td>
</tr>
<tr>
<td>ACID</td>
<td>43.72 ± 6.10</td>
<td>34.87 ± 4.91</td>
<td>7.929</td>
<td>&lt;.001</td>
<td>1.59</td>
</tr>
<tr>
<td>SCAD</td>
<td>43.94 ± 7.02</td>
<td>36.00 ± 5.50</td>
<td>6.254</td>
<td>&lt;.001</td>
<td>1.25</td>
</tr>
</tbody>
</table>

*Note.* Raw scores are presented for Forward and Backward Digit Span.
Table 3

Diagnostic Accuracy of Specific WISC-III Profiles in Children with DD versus Normal Readers

<table>
<thead>
<tr>
<th></th>
<th>Sensitivity (True Positive)</th>
<th>Specificity (True Negative)</th>
<th>False Positive</th>
<th>False Negative</th>
<th>Positive Predictive Power</th>
<th>Negative Predictive Power</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bannatyne pattern</td>
<td>.22</td>
<td>.90</td>
<td>.10</td>
<td>.78</td>
<td>.69</td>
<td>.54</td>
</tr>
<tr>
<td>FDI</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>full profile</td>
<td>.02</td>
<td>.98</td>
<td>.02</td>
<td>.98</td>
<td>.50</td>
<td>.51</td>
</tr>
<tr>
<td>in 3 lowest subtests</td>
<td>.18</td>
<td>.94</td>
<td>.06</td>
<td>.82</td>
<td>.75</td>
<td>.54</td>
</tr>
<tr>
<td>in 4 lowest subtests</td>
<td>.39</td>
<td>.90</td>
<td>.10</td>
<td>.61</td>
<td>.79</td>
<td>.60</td>
</tr>
<tr>
<td>ACID</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>full profile</td>
<td>.08</td>
<td>.98</td>
<td>.02</td>
<td>.92</td>
<td>.80</td>
<td>.52</td>
</tr>
<tr>
<td>partial profile</td>
<td>.20</td>
<td>.96</td>
<td>.04</td>
<td>.80</td>
<td>.83</td>
<td>.55</td>
</tr>
<tr>
<td>in 5 lowest subtests</td>
<td>.22</td>
<td>.98</td>
<td>.02</td>
<td>.78</td>
<td>.92</td>
<td>.56</td>
</tr>
<tr>
<td>in 6 lowest subtests</td>
<td>.45</td>
<td>.94</td>
<td>.06</td>
<td>.55</td>
<td>.88</td>
<td>.64</td>
</tr>
<tr>
<td>SCAD</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>full profile</td>
<td>.00</td>
<td>.98</td>
<td>.02</td>
<td>1.00</td>
<td>.00</td>
<td>.50</td>
</tr>
<tr>
<td>partial profile</td>
<td>.08</td>
<td>.92</td>
<td>.08</td>
<td>.92</td>
<td>.50</td>
<td>.51</td>
</tr>
<tr>
<td>in 5 lowest subtests</td>
<td>.10</td>
<td>.96</td>
<td>.04</td>
<td>.90</td>
<td>.71</td>
<td>.52</td>
</tr>
<tr>
<td>in 6 lowest subtests</td>
<td>.18</td>
<td>.92</td>
<td>.08</td>
<td>.82</td>
<td>.69</td>
<td>.53</td>
</tr>
</tbody>
</table>
Figure 1. ROC curve comparing true- and false-positive rates among children with DD and normal readers in the FDI, ACID and SCAD profiles.