Assessment of Dyslexia in the Urdu Language

Haidry, Sana

1.4 Overview of the Thesis

In the above sections we have established the need for the thorough assessment of reading difficulties in Urdu. We have also explained why and how we can adopt the dual-route model of reading to assess Urdu reading processes and impairments. **Chapter 2** covers the development and validation process of our Urdu reading/dyslexia test battery. The broad aim of this first study was to try and understand the nature of typical and impaired reading processes in the Urdu language. Our specific objectives were: (1) to develop a test battery to assess the ways in which reading impairments present in Urdu; (2) to understand the patterns of deficits of key reading processes in Urdu by comparing and profiling young struggling readers and typically developing age peers; and (3) to test the application of the dual-route reading model (DRM) to Urdu. The choice for and content of each individual test is described in detail as well as their validity and reliability, along with the results of our group comparisons (typical vs. struggling readers).

Apart from the fact that Urdu is spoken by more than 500 million people around the world, it still is an under-researched language. Certain features of Urdu, such as its dual orthography and the fact that alphabetic letters can change shape according to their position in a word (the letter-position effect), make it an interesting language to study. In **Chapter 3** we elaborate on the dual-orthography aspect by studying the effects of inconsistencies in the Urdu orthography due to the presence and absence of diacritics or short-vowel markers. We investigated the transparency and lexicality effect, and the application of the dual-route model by exploring which of certain presented words are read more accurately and which faster, namely (a) words with or without diacritics and (b) words or pseudowords. We again sought to gauge the difference between typical and struggling readers in terms of reading performance and reading speed.

Chapter 4 revolves around the letter-position effect. In the study presented we sought to understand the effect of the Urdu orthography on letter-position processing in beginning readers. As alluded to above, an intriguing aspect of the Urdu orthography is that many letters change their shape according to their position in a word. We examined reading accuracy with the hypothesis that scores would be higher for words with changed-shape migrated-letter cognates as compared to when the shapes of the letters remain the same. We expected this to be true for both the typical and the struggling readers but more so for the latter group. We additionally investigated word-frequency effects and differential effects between migrations of initial and final letters compared to migrations of middle letters.

The thesis concludes with a discussion of all the studies in **Chapter 5**, where also future directions in the field are proposed.

CHAPTER 2

Assessing Dyslexia in the Urdu Language

Abstract

Background: As in Pakistan tests to diagnose reading and spelling impairments are mostly based on English tests, there is a dire need for a dedicated test battery that assesses these basic skills in the Urdu language in the early stages of reading acquisition to thus allow timely identification and remediation of any deficits.

Aim and Method: To develop and validate a test battery to identify reading disability (dyslexia) in young children mastering the Urdu orthography. Based on dual-route model (DRM) of reading (in English) and dyslexia batteries in other alphabetic languages, tests for letter knowledge, word and pseudoword reading and spelling, and phonological abilities were constructed and administered to 167 typical readers (TR) and 128 struggling readers (SR) aged 7-11 years (grades 3-7; 150 boys, 145 girls) to establish the tests' reliability and validity, create profiles of the reading-related cognitive functions of proficient and struggling readers and test the applicability of dual-route model to Urdu.

Results: Test reliability was very high. Content validity was substantiated through the high correlation between two independent ratings. Overall, correlations of accuracy and speed measures confirmed the test battery to have high construct validity. All TR-SR differences were significant ($\alpha = .01$) in detriment of the SR group. Effect sizes (ES) were the highest for the spelling measures ($g > 2$), followed by the reading measures, where ES for accuracy ($g > 1.50$) were higher than those for speed ($g < 1.50$) and ES for pseudoword reading and spelling ($g > 2.5$) higher than those for word reading and spelling ($g = 1.59$-2.37). The medium ES for rapid automatized naming (RAN) and vocabulary were lower than those for reading and spelling.

Conclusion: The developed test battery based on DRM was reliable and valid, and differentiated well between TR and SR which can be interpreted as DRM being applicable to Urdu's dual orthography as well.

Key words: assessment, dyslexia, Urdu, dual-route model, dual orthography, reading, spelling

2.1 Introduction

Urdu, one of Pakistan's national languages, is the second most spoken language in the world (Grimes, 2000; Rahman, 2004; Ulrich, 2015). Despite having 588 million speakers including 70 million native speakers (Lewis et al., 2016; Ulrich, 2015), it still is a very much under-researched language (Farukh & Vulchanova, 2014). Hence, the nature of reading processes in Urdu and reading impairments, such as dyslexia, are not yet fully understood nor studied in-depth, which is an important reason why struggling readers are not properly identified, assessed or remediated in Pakistan. This paucity creates a vicious cycle of negative consequences such as academic underachievement, relatively low national education levels and un(der)employment, which potentially affects quality of life (Qin, 2016; Schulte-Körne & Bruder, 2010).

In Pakistan around six million children are deprived of quality education (Education for All, 2015). The most common barriers include high illiteracy rates, poverty, gender discrimination, a lack of training, low salaries and high workload for teachers (National Education Management Information System Pakistan-NEMIS, 2012-13). Spending less than 3% of the national budget on education, the Pakistani government appears negligent to these problems (Naeem, Mehmood, and Saleem, 2014). It is, accordingly, hard to determine whether children's reading problems in Urdu emanate from a lack of attention and available resources at school or a high prevalence of illiteracy in the family, or whether they can be attributed to dyslexia.

Note that in this study, unless indicated otherwise, we adopt the DSM-5 definition (American Psychiatric Association, 2013) for developmental dyslexia, being a specific learning disorder characterized by problems with accurate and/or fluent word recognition, decoding, and spelling abilities, and being distinct from reading and/or spelling difficulties caused by poor or inadequate instruction or an impoverished home environment (Snowling, 2000; Stanovich, 1988).

An exact dyslexia prevalence rate for Pakistani children is not yet available. Of all persons with disabilities in Pakistan, an estimated 25% are children aged 5-14 years (United Nation's Children Fund- UNICEF, 2003) of whom only 2% had access to special schools (Bureau of Statistics, 1998; Shahzadi, 2000). Screening 200 primary schoolgirls (grades 3, 4 and 5) for specific learning disorders (as based on DSM-5 criteria), Irshad (2005) concluded that 75 met the criteria, while the author also observed high rates of emotional problems such as depression, anxiety, lack of confidence and poor self-image. Ashraf and Majeed (2011) assessed 250 boys and 250 girls (aged 11-17 years) attending grades 6, 7 and 8 of Pakistani government schools and found 5% of the students to meet the DSM-5 diagnostic criteria for dyslexia.

It is widely accepted that early identification of dyslexia and timely remedial intervention are crucial for the academic outcome of individuals facing these learning difficulties in their childhood (Rothenberger, 2005; Stanovich, 1986; Webster-Stratton and Taylor, 2001). Previous studies have clearly shown that timely remediation positively affects children's reading skills, potentially preventing future problems such as un(der)employment (Bradley, 1988; Clay, 1991). Unfortunately, in the absence of a proper assessment battery for dyslexia in Urdu, tests developed for English are used without validation for Urdu. These tests do not assess phonological and orthographic specificities of Urdu. In a review of the literature, we did not find a single test for Urdu that can reliably distinguish between children with dyslexia and typically developing children.

Although a reading and writing comprehension test for Urdu has been developed for secondary schoolchildren (Khan et al, 2011), no tests have been designed to assess reading-related processes, such as letter and sound knowledge, letter-to-sound association (decoding), sound blending, whole-word recognition, vocabulary, spelling (encoding) and rapid naming (fluency). Moreover, most research has focused on older children (Ashraf and Majeed, 2011; Irshad, 2005; Khan et al, 2011; Naeem, Mahmood & Saleem, 2014), while assessments conducted in the early stages of reading acquisition are vital to identify dyslexia and initiate remedial interventions early. Furthermore, although researchers strongly advocate the use of culturally fair tests to obtain a reliable diagnosis of dyslexia, the tests currently used in Pakistani schools are all developed in Western countries (Naeem, Mahmood & Saleem, 2014).

In the present study, we seek to fill this gap in knowledge through targeted tests. We present and test a tool to assess reading and spelling in young Urdu-speaking children that discriminates struggling readers from typically developing peers and helps identify children with dyslexia at an early stage. Assessing dyslexia in a language with a deep orthography calls for a different method than in a language with shallow orthography (Masterson, 2000). Urdu is exceptional in this regard as it has a dual orthography: both deep and shallow, depending on the absence and presence of diacritics[1] respectively. This specific characteristic makes Urdu an interesting language to study in the context of models of reading. Since Urdu and English are both alphabetic languages and there is no dedicated reading model for Urdu, we largely based our dyslexia battery on the dual-route model (DRM) of reading (Coltheart, 1978), which has been shown to be applicable in most alphabetic languages (Levy et al., 2009; Ziegler, Perry & Coltheart, 2000), and does not only account for normal reading processes but also impaired reading and dyslexia. We aim to test the applicability of DRM to Urdu with our newly developed assessment tool, to contribute to a better understanding of the nature of reading processes and reading impairments (dyslexia) in Urdu. Urdu is an Indo-Aryan language having 40 letters; it is orthographically similar to Arabic and Persian, cursive in nature, and written from right to left.

Orthographies of languages can be classified on a continuum from shallow (one-to-one letter-sound correspondence) to opaque (multiple letters corresponding to more than one sound). This transparency of letter-to-sound mapping is called orthographic depth (Ziegler & Goswami, 2005). As mentioned before, Urdu has fluctuating orthographic depth due to the inclusion and exclusion of diacritics. It has regular spelling-to-sound correspondences when fully written out with diacritics but when diacritics are missing successful word identification happens with contextual help (Mumtaz and Humphreys,

[1] Diacritics are three short vowel markers in the form of strokes placed above and/or below the letters. These are as follows: /ə/ symbolised by 'َ◌' placed above the consonant بَ, /ɪ/ symbolised by 'ِ◌' placed under the consonant بِ and /ʊ/ symbolised by 'ُ◌' placed above the consonant بُ.

2001). Usually, diacritics are omitted in reading materials for adults and from children's reading books from grade 2 onwards (Rao et al., 2011). Another interesting aspect of Urdu is that letters change their shape according to their position in a word. For instance, the isolated form 'غ' (/ɣ/) can be written as غـ, ـغـ and ـغ in initial, medial and final positions, respectively. Investigating reading acquisition in young children can provide more insight into how inclusion or exclusion of diacritics and letter positions affect normal and impaired reading processes in Urdu. We developed assessments of the strengths and weaknesses of beginner readers based on the functionality of particular components and/or pathways of the dual-route model (DRM - see Figure 2.1).

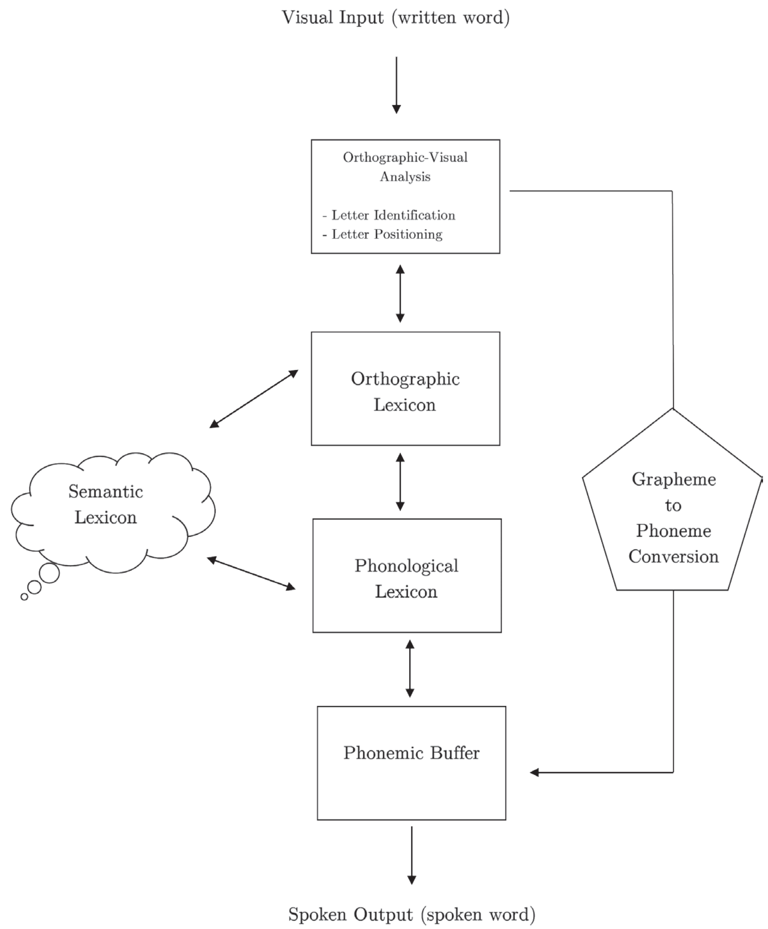

Figure 2.1: The components of the dual-route model of reading aloud (DRC Model - Coltheart, Rastle, Perry, Langdon & Ziegler, 2001, Figure 6, page 213).

2.1.1 Theoretical Perspective – Dual-Route Model (DRM)

According to the DRM, there are two separate routes for reading aloud: letter-to-sound conversion or indirect route, and visual word recognition or direct route. In the indirect route letters are converted into phonological representations. Pseudowords (e.g. 'lat', 'gat') and novel words are primarily read through this route. In the direct route, the reader recognizes the visual word form of the whole word without performing letter-to-sound conversions. Irregular words (e.g. 'enough') and familiar words are primarily read through this route. Successful reading then requires the two routes to be applied proficiently. Malfunction of any of these processes will lead to specific reading issues (Coltheart, 2005). If the development of the orthographic lexicon is impaired, for instance, this can lead to errors in reading irregular words while reading regular words and pseudowords would be unaffected, a disorder which is known as surface dyslexia (Castles and Coltheart, 2013; McDougall, Borowsky, MacKinnon & Hymel, 2005) If, conversely, letter-to-sound conversion is impaired, this will elicit errors in reading new words and pseudowords, a disorder which is commonly labelled as phonological dyslexia (Castles and Coltheart, 2013; McDougall, Borowsky, MacKinnon & Hymel, 2005). If the impairment occurs at an earlier processing stage, such as at the level of orthographic-visual analysis, this will produce errors in letter identification and letter-position coding (Kezilas et al, 2014). Having multiple letter forms for each letter, Urdu might seem more difficult in some respects than English. Then again, letters that have different shapes in initial, middle and final positions, facilitate in word recognition and thus potentially reduce the frequency of letter position errors (Friedmann and Haddad-Hanna, 2012).

To enable us to successfully identify reading disabilities in children first learning to read Urdu, we took the DRM and existing dyslexia batteries for other alphabetic languages as our starting point. We constructed Urdu-specific tests to assess letter knowledge and word and pseudoword reading and spelling, as well as tests gauging phonological abilities, since these are the most prominent reading-related functions according to the definition of dyslexia adopted by the International Dyslexia Association (IDA; Lyon, Shaywitz and Shaywitz, 2003). To determine its reliability and validity, the test battery was administered to a large group of typical and struggling readers. Then we compared the reading and spelling performance as well as the profiles of reading-related cognitive functions of the typical and struggling readers. Since there was no theoretical model available for Urdu, we largely based our test battery on DRM. As Urdu is an alphabetical language similar to English, we expected to find comparable reading difficulties in Urdu-speaking children pertaining to the various DRM components, which can be identified using different targeted reading measures for each component.

2.2 Method

This research was ethically approved by the Macquarie University Human Research Ethics Committee (Australia, Reference No. 5201300826) and by the Research Ethics Committee (CETO) of the Faculty of Arts of the University of Groningen (The Netherlands). Informed consent was available for all the children.

2.2.1 School and Participant Selection

Mainstream schools in the city of Karachi were sent letters informing them about our study, while orientation meetings were arranged for staff from interested schools. From these, we selected three private middle-income schools (i.e. the Meezan Academy, SET School and Green Island School) that already had screening and remedial setups in place for their students.

It is important to mention here that Karachi is a multilingual society and there are hardly any monolingual children because most are exposed to two or more languages from an early age given that Urdu is the 'lingua franca' for people from different provinces. All the provinces have their main regional language, a national language (Urdu) and an official language of correspondence, which is English. For most families the regional and national languages are the two main languages spoken at home. The regional language is mostly referred to as the heritage language, i.e. the language of the geographic region the child's family originally stems from. Urdu enjoys the status of "second first language." In Pakistani schools, Urdu and English are both regular compulsory subjects from grade 1 onwards. Although English is the official language of instruction in most private Pakistani schools, in the classroom the children communicate in Urdu and English, while during breaks and playtime they predominantly speak Urdu.

The children we wished to recruit had to have been exposed to Urdu and English in terms of listening, speaking, reading and writing from an early age, both at school and at home. Accordingly, the main inclusion criterion for our study was that children needed to speak and understand Urdu and that they had received at least two years of formal reading instruction in Urdu. The principals and teachers of the three selected schools were invited to attend orientation sessions about the study and required sample, after which they prepared lists of typical (TR) and struggling (SR) readers per class. The children's teachers were requested to compose an SR group based on academic performance (grades), classroom reading and spelling performance, including children showing reluctance to read. The final selection was performed through stratified random sampling until the target sample size was reached. The descriptives of the recruited sample can be found in Table 2.1.

A total of 295 children (150 boys and 145 girls), aged 7-11 years attending grades 3-7 were recruited, of whom 167 were characterized as typical readers by their teachers and parents and 128 as struggling readers[2]. Of these 128 struggling readers, a subset of 32 students had been diagnosed as having dyslexia in English based on the ASTON index (Newton and Thomson, 2003) prior to our study. However, no details of the diagnostic reports were shared. Both TR and SR group belonged to the same socio-economic status. Although sample size is often problematic in between-subject designs, we believe the present sample size generates sufficient statistical power to pursue the aims of this study (Faul et al., 2007).

[2] In this study, the term *struggling readers* refers to children experiencing reading and spelling problems and it is not known exactly how far they lag behind in comparison to their typically developing peers nor what causes the arrears. If not otherwise specified, the term dyslexia is used when the DSM-5 criteria for a specific learning disorder in reading and/or spelling are met (APA, 2013).

Table 2.1: Descriptives of the typical and struggling readers included in the study

Variables	TR group	SR group	Total
Number	167	128	295
Gender			
Female	83	62	145
Male	84	66	150
Total	167	128	295
Age (in years)			
7	25	24	49
8	17	40	57
9	43	26	69
10	34	19	53
11	48	19	67
Total	167	128	295
Grade/Class			
2	0	11	11
3	37	54	91
4	33	31	64
5	38	15	53
6	29	11	40
7	30	6	36
Total	167	128	295

Note: TR = typical readers; SR = struggling readers

2.2.2 Selection of the Test Items

We selected 300 words from grade 3-to-7 Urdu text books, excluding compound words (mostly derived from Persian), words with heavy Arabic influence and commonly mispronounced words. On the basis of these, we constructed the same number of corresponding pseudowords matched for the number of syllables. Next, frequency and age of acquisition of the words were rated by grade 3-to-7 teachers, after which 10 female and 10 male volunteers (18-80 years) with a minimum of 12 years of formal education performed a second independent rating on the same variables. Frequency was rated on a 3-point scale, with 1 denoting low and 3 denoting high frequency. The selected words were all typically acquired before the age of 7 years. On the basis of high correlations (frequency, Spearman $\varrho = 0.81$ and age of acquisition, Spearman $\varrho = 0.83$) between the two ratings, final test items were selected fulfilling the requirements of each test specified in the descriptions below.

2.2.3 Procedure

We ran a pilot test with 20 typical (TR; 10 boys) and 20 struggling readers (SR; 10 boys) whom we randomly selected from all relevant grades, while controlling for socio-

economic status. The pilot sample was representative of the total sample. Testing was on an individual basis and performed during school hours at the child's school. After the pilot, a small number of items was replaced with more contextually relevant and culturally fair items.

Actual testing was also conducted at the child's school in individual 45-60 minute sessions. Breaks were offered, if needed. All tests were paper and pencil tasks. Children could stop and leave the session at any time. Results were shared with the child's parents and teachers but not directly with the children themselves. Individual feedback and recommendations were provided after the session.

2.3 Theoretical Framework of Dyslexia Assessment Test

As mentioned above, we largely based the design of our Urdu reading assessment battery on the DRM for reading (Coltheart, 1978), which distinguishes various components and interconnected processes, depicted in Figure 2.1. Each component has specific functions and any deficits in these functions will manifest as different types of reading impairments (Coltheart, 1978; 2005; 2006), whose causes fall outside the focus of this paper (for a detailed aetiological account, see Reid, 2016). Each test taps into a corresponding DRM component (see Table 2.2 below). Following the DSM-5 definition of dyslexia, we also added a word and pseudoword spelling test, as well as other tests gauging cognitive skills known to be predictive of dyslexia, such as phonological awareness (PA) and rapid automatized naming (RAN). Since dyslexia is considered a multi-factorial condition, these cognitive skills are often associated with dyslexia but by themselves not sufficient or necessary for dyslexia to be diagnosed.

Table 2.2: The functions and reading impairments per DRM component

Components and pathways	Function/Deficits	Manifestations if defective	Reading Impairment
Orthographic-visual analysis	Letter identity Letter-position coding	Letter substitutions and omissions in isolation and in context (word) Transposition of letters	Letter-identification-dyslexia Letter-position-dyslexia
Processing routes	Whole-word recognition Letter-to-sound conversion	Problems in reading accuracy and speed of irregular words Problems in reading accuracy and speed of new words and pseudowords	Surface dyslexia Phonological dyslexia
Conceptual and semantic lexicon	Comprehension of words	Reading without knowing the meaning of words	Semantic dyslexia

2.4 Description of the Tests

All the tests in our battery are at the letter and word level as previous research has shown that tests at the sentence level contribute little to the identification of dyslexia (Tops, Callens, Bijn, & Brysbaert, 2013). By definition, dyslexia is about reading words; reading comprehension is functionally important but not a core characteristic of the deficit.

2.4.1 Letter Identification

This test corresponds to the first function of the DRM's orthographic-visual analysis component. It assesses letter-name and letter-sound knowledge, thus gauging a child's ability to name and sound-out Urdu letters. Deficiencies in both these skills distinguish struggling from typical readers in early reading development. Both are predictors of later word reading (de Jong & van der Leij, 1999; Lonigan, Burgess, Anthony, & Barker, 1998; Wimmer et al., 2000). As to letter naming, the child is shown a letter and asked to give the corresponding name (active naming), while in the letter-sound test, the administrator vocalizes the letter sound after which the child is to point out the matching letter (letter-sound matching). There are 20 items for each of the two skills (for sample items, see Figures 2.2 and 2.3) with accuracy being the only outcome measure (reading speed was not recorded).

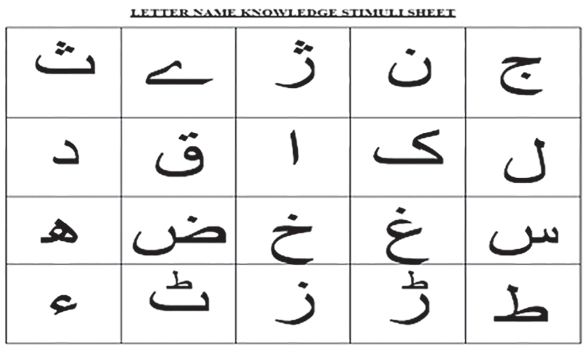

Figure 2.2: Sample items from the letter-naming test

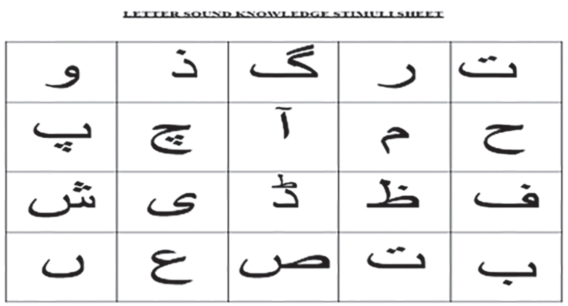

Figure 2.3: Sample items from the letter-sounding test

Two of the selected letters, آ and ں , are variations of their original forms ا and ن , respectively. There was proper representation of consonants and vowels in both test sheets.

2.4.2 Letter-position Processing

Coinciding with the second function of the DRM's orthographic-visual analysis component, this test evaluates a child's ability to code the order of letters in words. Errors in letter position coding or letter migrations while reading are a hallmark symptom of letter position dyslexia (Kohnen et al., 2012). The test consists of 50 words presented on two separate lists, each including 25 words of a cognate pair. List 1 includes 25 high-frequency words and list 2, 25 low-frequency cognates. Each of the words in list 1 has a cognate in list 2 in which the letters has changed position, of which half are visually similar cognates (letters keeping the same shape) and half visually less similar cognates (letters changing shape). Children read aloud the 25 words of list one first, followed by those on list two. Outcome measures are accuracy and reading speed (for the two lists). Table 2.3 depicts some examples.

Table 2.3: Sample items from the letter position test

Words	Roman	IPA	English Meaning	Same-Shape Cognate	Roman	IPA	English Meaning
جواب	Jawaab	/j ə v aː b/	Answer	واجب	Waajib	/v aː j ɪ b /	Obligatory
ماتھا	Maatha	/m ɑː t̪ʰ ə /	Forehead	تھاما	Thama	/ t̪ʰ ɑː m ə /	Hold
اور	Aur	/ ɔ r/	And	وار	Waar	/v aː r/	Attack
Words	**Roman**	**IPA**	**English Meaning**	**Change-Shape Cognate**	**Roman**	**IPA**	**English Meaning**
لات	Laat	/l aː t̪ /	Kick	تال	Taal	/ t̪ aː l/	Rhythm
سڑک	Sarak	/ s ə ɽ ə k /	Road	سکڑ	Sukar/d	/ s ʊ k ə ɽ /	Shrink
بادل	Badal	/b aː d ə l/	Cloud	بدلا	Badla	/b ə d l ə /	Revenge

2.4.3 Pseudoword Reading (Letter-Sound Correspondence)

Corresponding to the letter-to-sound conversion route of the DRM, this test gauges a child's ability to decode words based on his/her knowledge of the relationship between Urdu letters and sounds. If letter-to-sound conversion is impaired, errors will be made in reading new words and pseudowords, a reading difficulty known as phonological dyslexia (Castles and Coltheart, 1993; McDougall, Borowsky, MacKinnon & Hymel, 2005). The test consists of 30 pseudowords of increasing complexity that cannot be read by whole-word recognition. They are presented with diacritics (vowel markers) that enable the child to read it aloud correctly as long as s(he) knows the corresponding rules. Outcome measures are accuracy and reading speed. Examples are given in Table 2.4.

Table 2.4: Sample items from the pseudoword reading test

Words	Roman
ذِخ	Zikh
بَتُش	Batush
زِلچابُن	Zilchabun

2.4.4 Word Reading (Whole-Word Recognition)

Covering the visual-word-recognition pathway and orthographic-input lexicon component of the DRM, this test assesses a child's whole-word recognition skills. If this direct route is impaired, it will lead to errors in reading sight words or irregular words while the child may still be able to correctly articulate regular words and pseudowords, which would be indicative of surface dyslexia (Castles & Coltheart, 1993). The test comprises 30 existing words (without vowel diacritics) of increasing difficulty (i.e. one, two or three syllables) that can only be read through the direct route as the readers would not know the pronunciation unless they were familiar with them. Each word has one correct pronunciation only. Of the 10 words in each syllable category half are high- and half low-frequency words. Outcome measures are accuracy and reading speed. Examples are provided in Table 2.5.

Table 2.5: Sample items from the word reading test

Words	Roman	English meaning
پیار	piyar	love
مچھر	machchar	mosquito
کھلونے	khilonay	toys

2.4.5 Word Reading With and Without Diacritics

Consistent with the two routes of DRM, this test gauges the processing of words with and without full phonological information. It consists of two different lists each comprising the same 30 words either with diacritics commanding one pronunciation only or without diacritics where each word has two or three possible pronunciations. It differs from the pseudoword reading test in that in the present test real words are presented, while it differs from the word reading test in that the words included have not one but two or three possible pronunciations (and meanings). Outcome measures are accuracy and reading time. Examples are provided in Table 2.6 below.

Table 2.6: Sample items from the word reading test with and without diacritics

Words with diacritics	Roman	English meaning	Same words without diacritics	Roman	English meaning
سُونا	soona	deserted	سونا	sona/soona	gold/deserted
گُھٹنا	ghutna	knee	گھٹنا	ghatna/ghutna	decrease/knee
تُلا	tula	persistent	تلا	tala/tula	fry/persistent

2.4.6 Word and Pseudoword Spelling

In congruence with the DSM-5 definition of dyslexia, we also included a word and pseudoword spelling test to assess the children's ability to encode words (sound-to-letter conversion) based on their knowledge of the relationship between Urdu letters and sounds (pseudoword spelling), and whole-word memorization (word spelling). There is high co-morbidity and interdependency between reading and spelling impairments (Ehri & Snowling, 2004). Spelling by sound-letter conversion requires letters to be matched to corresponding speech sounds (Berninger et al., 1998; Ehri, 1998; Moats, 2004). Skilled spelling through whole-word memorization is supported by underlying knowledge of a word's print form, speech sounds and meaning (Ehri & Snowling, 2004). Once the spelling of a word has been mastered, its representation is committed to memory and easily accessed for fluent reading (Snow et al. 2005). Of the 30 items in this test, 15 are existing words arranged in the order of difficulty (five single syllable, five two and the remaining five three syllables). The other 15 items are matching pseudowords containing the same number of syllables. In this test the investigator says the word and child writes it down on a blank sheet of paper. The outcome measure is accuracy. Examples are shown in Table 2.7.

Table 2.7: Sample items from the spelling test of words and pseudowords

Words	Roman	English meaning	Pseudowords	Roman
درد	dard	Pain	لرد	lard
ملک	mulk	Country	چُلک	chulk
پہاڑ	pahaard	Mountain	فِگاڑ	figaard

2.4.7 Vocabulary

To measure functioning of the semantic lexicon and conceptual system components of the DRM, this test assesses a child's knowledge of the meaning(s) of words. The literature shows relationships between vocabulary deficits and at-risk readers (Pierce et al., 2007). Reading comprehension is the product of adequately decoding words and attaching the correct meaning to them (Snow, Burns & Griffin, 1998), with vocabulary size being directly connected to reading comprehension (Steven, 1999). A child may suffer from semantic dyslexia if they read words without attaching meaning to them (Pierangelo & Giuliani, 2006). Our vocabulary test has 62 items, each consisting of four pictures with each set being depicted on a separate page. One picture is the target picture and the others are: a phonological distracter (having the same initial phoneme as the target word), a semantic distracter (having a similar meaning to the target word) and an unrelated picture. The test administrator says the word aloud after which the child is expected to point to the correct picture. The outcome measure is accuracy. A sample item is provided in Figure 2.4.

Figure 2.4: Sample item from the vocabulary test

In Figure 2.4, the vegetable (lady finger or okra) in the upper right picture depicts the target word '**bh**indi', while the picture below it shows a 'gobhi' (cauliflower), a semantically related word. The upper left picture shows a '**bh**aalu' (a teddy bear), a phonetically related word, while the bottom left picture is of a 'kawwa' (crow), an unrelated word.

2.4.8 Phoneme and Syllable Deletion

Not having a corresponding DRM component, this test gauges phonological awareness (PA) and was included because PA has consistently been shown to be the single most crucial predictor of reading impairment (Park & Uno, 2015; Snowling, 1995; Snowling 2000; Torgesen, 1996; Wolf & Bowers, 1999). Phoneme blending and segmentation (addition and deletion) are typical tests to quantify PA in impaired readers (Lee, 2008). In our test, the investigator asks the child to leave out either the initial, middle or final phoneme or syllable of an orally presented word and then vocalize this truncated word. An example: the examiner orally presents the word تب ('/t ə b/' meaning 'then') and instructs the child to leave out the '/t/' sound (which is in initial position), after which the child is expected to say '/ə b/'. This is the only oral test in the battery. It consists of 30 items, 10 items each for initial, medial and final phoneme/syllable deletion, while in all three sets one-, two- and three-syllable words are equally represented. The outcome measure again is accuracy. See Table 2.8 for sample items.

Table 2.8: Sample items from the phoneme- and syllable-deletion test (phonological awareness)

Words	Roman	Meaning	Phoneme/ syllable deleted	Roman	Desired response	Roman
ماتم	matam	mourning	تم	tam	ما	ma
پانی	pani	water	پا	pa	نی	ni

2.4.9 Rapid Automatized Naming (RAN)

This test has no corresponding DRM component but was included because there is evidence of the predictive role of RAN in reading accuracy and fluency for Urdu (Farukh & Vulchanova, 2014). The test comprises four stimuli sheets showing colours, objects and digits (one for English, one for Urdu).The digits range from 1-9 and are presented in different order on the two sheets. The child needs to name the stimuli as accurately and as quickly as possible, with reading speed being the sole outcome measure.

Table 2.9 provides an overview of all the factors and corresponding tests and skills evaluated.

Table 2.9: Overview of the factors and corresponding tests

Factors	Tests	Measures
Letter Level		
Letter Knowledge	Active letter naming Recognition of letter sounds	Letter-name knowledge Letter-sound knowledge
	Phoneme deletion	Phonological awareness Phonological sequencing Phonological memory
Word Level		
Reading	Word Pairs (Letter Position)	Letter-position effect
	Pseudowords	Phonological decoding (letter-sound mapping)
	Words without Diacritics (1 reading possibility) (2 or 3 reading possibilities)	Visual-word recognition
	Words with Diacritics	Phonological Decoding Visual-word recognition
Spelling	Pseudowords Words	Phonological encoding (sound-to-letter mapping) Orthographic knowledge
Vocabulary	Vocabulary	Lexical retrieval Word-meaning knowledge
Rapid automatized naming (RAN)	Colours Objects Digits-English Digits-Urdu	Fast lexical retrieval

2.5 Results

2.5.1 Reliability

First, we calculated the reliability measures of the whole test battery (see Table 2.10). The overall test reliability was very high, Cronbach's $a = 0.94$; Guttman split-half y= 0.97, where we considered reliability coefficients to be high from .70 onwards (Fraenkel & Wallen, 2006).

Table 2.10: Descriptives and reliability coefficients per test

Tests	**Total Items**	**Mean**	**Median**	**(SD)**	**Cronbach's α**	**Guttman split-half y**
Letter-name knowledge	20	18.64	19	*(2.09)*	.82	.80
Letter-sound knowledge	20	18.94	20	*(2.29)*	.87	.80
Phoneme and syllable deletion	30	25.42	29	*(6.72)*	.95	.90
Reading letter position HF	25	20.57	23	*(6.11)*	.95	.94
Reading letter position LF	25	18.95	22	*(6.90)*	.95	.95
Reading pseudowords	30	20.12	26	*(10.32)*	.97	.94
Reading words	30	22.89	27	*(8.62)*	.97	.94
Reading without diacritics	30	23.87	28	*(8.42)*	.97	.96
Reading with diacritics	30	20.44	24	*(9.16)*	.96	.95
Spelling words	15	9.81	12	*(5.22)*	.94	.94
Spelling pseudowords	15	9.23	12	*(5.75)*	.96	.95
Vocabulary	62	61.47	62	*(1.09)*	.68	.47

Note: Reliability coefficients = Cronbach's α and the Guttman split-half y, HF = high-frequency, LF = low-frequency

Next, Cronbach's a and the Guttman split-half y of each test were determined separately based on the accuracy scores for all individual tests, except for RAN because it has speed as its sole outcome measure. For each test descriptives and reliability coefficients are given in Table 2.10. For all tests both reliability measures were higher than 0.80, except for Vocabulary where reliability was in the medium range, $a < .70$ and $y \leq 0.60$.

2.5.2 Validity

Content and construct validity, sensitivity and specificity were calculated for the overall test.

2.5.2.1 Content validity. Content validity was established during item selection (see method section above).

2.5.2.2 Construct validity. To judge the tests' construct validity, we calculated inter-test correlations for accuracy scores (Table 2.11), and speed measures (Table 2.12).

Table 2.11: Inter-test correlation matrix for accuracy

	LNK	**LSK**	**RLPHF**	**RLPLF**	**RPW**	**RW**	**SW**	**SPW**	**V**	**PSD**	**RWND**	**RWD**
LNK	1.00	.48	.59	.53	.43	.54	.46	.50	.28	.43	.53	.48
LSK		1.00	.59	.56	.44	.57	.49	.49	.29	.52	.57	.49
RLPHF			1.00	.90	.79	.89	.80	.74	.33	.60	.89	.83
RLPLF				1.00	.83	.90	.85	.78	.37	.60	.91	.87
RPW					1.00	.80	.79	.79	.37	.56	.80	.83
RW						1.00	.84	.76	.36	.62	.91	.88
SW							1.00	.89	.37	.62	.84	.83
SPW								1.00	.34	.64	.79	.75
V									1.00	.27	.30	.32
PSD										1.00	.67	.60
RWND											1.00	.90
RWD												1.00

Note: Spearman ρ was used for correlations. Interpretation: high correlation = .5 to 1.0 or -0.5 to -1.0; medium correlation = .3 to .5 or -0.3 to -0.5; low correlation: .1 to .3 or -0.1 to -0.3 (Cohen, 1988). All correlations were significant at the 0.01 level. Abbreviations: LNK=Letter Name Knowledge, LSK=Letter Sound Knowledge, RLPHF=Reading Letter Position High Frequency, RLPLF=Reading Letter Position Low Frequency, RPW=Reading Pseudowords, RW=Reading Words, SW=Spelling Words, SPW = Spellings Pseudowords, V=Vocabulary, PSD=Phoneme and Syllable Deletion, RWND=Reading without Diacritics, RWD=Reading with Diacritics

All the tests correlated positively. We computed significant ($a = .01$) and strong correlations between all reading and spelling tests. Similar high correlations were found in previous studies (Snow et al., 2005), indicating word reading, pseudoword reading and spelling to be interrelated and measuring similar domains. Correlations with vocabulary were moderate-to-low.

All speed measures correlated significantly ($a = .01$), with strong correlations for all reading tests. Correlations between reading and RAN were moderate, confirming previous findings where RAN was found to be a predictor for reading (Farukh & Vulchanova, 2014).

Taken together, the correlations of the accuracy and speed measures convincingly show the test battery to have high construct validity.

Table 2.12: Inter-test correlation matrix for speed measures (time in seconds)

	RLPHF	**RLPLF**	**RPW**	**RW**	**RANC**	**RANO**	**RANDE**	**RANDU**	**RWND**	**RWD**
RLPHF	1.00	.91	.62	.85	.51	.51	.59	.60	.83	.75
RLPLF		1.00	.66	.87	.53	.49	.53	.54	.84	.80
RSW			1.00	.70	.34	.31	.45	.40	.63	.66
RW				1.00	.51	.50	.54	.56	.85	.80
RANC					1.00	.60	.47	.58	.52	.42
RANO						1.00	.58	.62	.51	.43
RANDE							1.00	.66	.58	.50
RANDU								1.00	.56	.47
RWND									1.00	.87
RWD										1.00

Note: Spearman ρ was used for correlations. Interpretation: high correlation = .5 to 1.0 or -0.5 to 1.0; medium correlation = .3 to .5 or -0.3 to -0.5; low correlation: .1 to .3 or -0.1 to -0.3 (Cohen, 1988). RLPHF=Reading Letter Position High Frequency, RLPLF=Reading Letter Position Low Frequency, RPW=Reading Pseudowords, RW=Reading Words, RANC=Rapid Automatized Naming Colours, RANO=Rapid Automatized Naming Objects, RANDE=Rapid Automatized Naming Digits English, RANDU=Rapid Automatized Naming Digits Urdu, RWND=Reading without Diacritics, RWD=Reading with Diacritics

2.5.3 Group Comparisons

After having established the reliability and validity of our test battery, we ran group comparisons to see whether it would indeed distinguish between typical and struggling readers.

We compared both accuracy and speed scores for all the tests (except for RAN since speed was the sole measure) using Mann-Whitney U ($\alpha = .01$) and applying Bonferroni corrections because of the large number of tests. Group differences in accuracy and speed were expressed as effect sizes (ES; using Hedge's g to correct for any differences in sample size of the two groups), which are shown in Tables 2.13 and 2.14. We considered ES large when > 0.80, medium when between 0.50 and 0.80, and small when < 0.50. Positive values indicate a better performance for the typical readers as compared to their struggling peers.

As with the accuracy measures, large ES were found for all tests except for Letter-name knowledge and Vocabulary (both medium ES), with the SR group always performing poorer than the TR group. It should be noted that mean scores on the Vocabulary test were close to ceiling in both groups, indicating that this particular test was easy for all the children we tested.

Table 2.13: Accuracy scores for the typical (TR) and struggling readers (SR)

Tests	**Total Items**	**TR (n=167) M**	***(SD)***	**SR (n=128) M**	***(SD)***	**Effect Size Hedges' *g***
Letter-name knowledge	20	19.29	*(1.08)*	18.09	*(2.30)*	0.69*
Letter-sound knowledge	20	19.78	*(0.68)*	18.27	*(2.44)*	0.89*
Phoneme and syllable deletion	30	28.33	*(3.02)*	22.36	*(7.49)*	1.10*
Letter Position-high frequency	25	23.93	*(2.23)*	17.16	*(5.91)*	1.59*
Letter position-low frequency	25	23.15	*(2.38)*	14.36	*(6.63)*	1.86*
Pseudoword reading	30	27.41	*(3.06)*	11.55	*(8.75)*	2.55*
Word Reading (single pronunciation option)	30	28.02	*(2.95)*	17.27	*(8.53)*	1.77*
Words without diacritics (2-3 pronunciation options)	30	28.75	*(2.89)*	18.61	*(8.40)*	1.70*
Words with diacritics	30	26.21	*(4.04)*	13.87	*(8.35)*	1.96*
Spelling words	15	13.41	*(1.93)*	5.58	*(4.49)*	2.37*
Spelling pseudowords	15	13.24	*(2.41)*	4.28	*(4.34)*	2.64*
Vocabulary	62	61.77	*(0.55)*	61.16	*(1.27)*	0.65*

Note: All the differences are statistically significant (P<0.01)

Table 2.14: Speed measures (in seconds) for the typical (TR) and struggling readers (SR)

Tests	**Total Items**	**TR (n=167) M**	***(SD)***	**SR (n=128) M**	***(SD)***	**Effect Size Hedges' *g***
Letter position High Frequency	25	36.76	*(26.46)*	122.05	*(90.10)*	1.34*
Letter position Low Frequency	25	44.77	*(31.41)*	133.52	*(87.34)*	1.42*
Pseudoword reading	30	107.74	*(65.41)*	196.42	*(101.09)*	1.07*
Word reading (1 pronunciation option)	30	52.76	*(43.81)*	159.91	*(101.06)*	1.44*
Words without diacritics (2-3 pronunciation options)	30	44.03	*(30.28)*	123.77	*(76.01)*	1.45*
Words with diacritics	30	62.18	*(40.80)*	137.02	*(80.54)*	1.22*
RAN Colours	35	37.50	*(13.61)*	54.27	*(26.96)*	0.82*
RAN Objects	35	30.31	*(9.34)*	38.24	*(12.81)*	0.64*
RAN Digits English	35	17.80	*(9.45)*	30.57	*(28.41)*	0.72*
RAN Digits Urdu	35	26.11	*(16.36)*	41.70	*(24.14)*	0.77*

Note: All the differences are statistically significant (P<0.01), higher score = more time taken to read

All tests had large ES, except for RAN (medium ES for RAN-objects, and Urdu and English digits) in detriment of the SR group. We can summarize our main findings as follows:

- All the differences were significant ($\alpha = .01$), with poorer SR outcomes;

- Effect sizes for spelling measures were largest (g > 2), followed by reading measures;
- Effect sizes for reading accuracy (g > 1.50) were larger than those for reading speed (g < 1.50).
- Effect sizes (g > 2.5) for pseudoword reading and spelling were larger than those for word reading and spelling;
- RAN and Vocabulary both had medium ES, showing that tests outside the reading and spelling domain had clearly smaller ES, which could mean that they are less clinically relevant for the TR-SR distinction but exist as a separate construct independent of reading and spelling factors.

2.5.3.1 Sensitivity and specificity. We calculated sensitivity and specificity in different ways manipulating cut-off scores (Pc 16 or Pc 10) and population (whole group, children diagnosed with dyslexia in English or typical readers only). First, out of our sample of 295 children, 128 were characterized as struggling readers by their teachers and parents. With TR versus SR as the diagnostic categories, we plotted receiver operating characteristic (ROC) curves for word and pseudoword reading and, word and pseudoword spelling. The areas under the ROC curve were very high for all the tests (> 0.9). We also calculated the specificity and sensitivity of our tests based on the DSM-5 clinical diagnostic criterion that a child scoring less than the 16th percentile is to be considered having dyslexia (APA, 2013). The tests' specificity was very high (> 0.9) but their sensitivity low (< 0.6).

Second, since the 16th percentile threshold is based on the assumption of a normal distribution but our sample was not normally distributed, we plotted the reading (words and pseudowords; Figure 2.5) and spelling (words and pseudowords; Figure 2.6) scores of the TR and SR groups and recalculated the sensitivity and specificity based on the intersection points of the sample distribution (cut-off score 23 for reading and 12 for spelling – Figures 2.5 and 2.6). With these criteria both the specificity and sensitivity were very high (> 0.9).

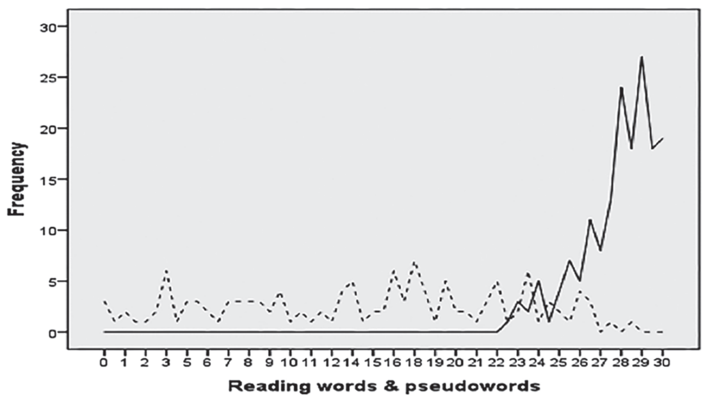

Figure 2.5: Distribution of typical and struggling readers on word and pseudoword reading where ____= typical and ------ = struggling readers

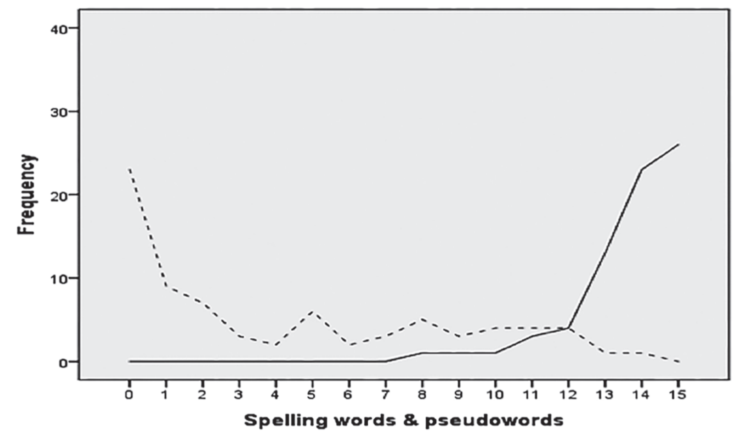

Figure 2.6: Distribution of typical and struggling readers on word and pseudoword spelling where ___ = typical and ------ = struggling readers

Third, we validated our results with another external criterion. As mentioned in the method section, of the 128 SRs, 32 had been diagnosed with dyslexia prior to our study based on the ASTON index for English. When we ran the sensitivity and specificity analysis on this diagnostic category, the area under the ROC curve (AUC), sensitivity and specificity were all very high (>0.9). All 32 children with dyslexia scored below the 16th percentile (DSM-5; APA, 2013).

Fourth, we conducted the different categorical analyses according to the *counting deficit method* (Pennington, 2012, p. 216) and determined 10% and 16% cut-off points based on the scores of typical readers only (excluding struggling readers). We calculated the new parameters with the value '0' if the score is above the 10th percentile and 16th percentile (condition is not present) and '1' if it is below (condition is present), for word reading, pseudoword reading, spelling of words and spelling of pseudowords. We then cross-tabulated the poor performers in the group of struggling and typical readers. Our results showed that only 3/167 typical readers fell in the lowest 10% in *reading words and pseudowords*, as compared to 97/128 of the struggling readers (98% specificity, 76% sensitivity), and 9/167 typical readers fell in the lowest 16% in reading words and pseudowords, compared to 108/128 of the struggling readers (95% specificity, 84% sensitivity). In case of *spelling words and pseudowords*, our results showed that 25/167 typical readers fell in the lowest 10% compared to 104/128 of the struggling readers (85% specificity, 81% sensitivity), and 42/167 typical readers fall in the lowest 16% in spelling words and pseudowords, compared to 114/128 of the struggling readers (75% specificity, 89% sensitivity). These outcomes are indicative of the robustness of our battery and further support that our test battery has high sensitivity and specificity based on the cut-off scores derived from the typical readers.

Finally, we took our analysis to a step further and explored the variability within the struggling reader group in order to observe whether our test provides the opportunity to assess specific patterns of reading and/or spelling deficits. We found that it does, as illustrated in the Tables 2.15 (16th percentile) and 2.16 (10th percentile).

Table 2.15: Profiling of struggling readers based on 16th percentile

Profiles	Deficit: Reading Words	Deficit: Reading Pseudowords	Deficit: Spelling Words	Deficit: Spelling Pseudowords	No. of struggling readers
Multiple deficit	√	√	√	√	96
Isolated reading deficit	√	√	x	x	2
Isolated spelling deficit	x	x	√	√	2
Word reading and spelling deficit	√	x	√	x	0
Pseudoword reading and spelling deficit	x	√	x	√	2
Multiple deficit except for word reading	x	√	√	√	8
Multiple deficit except for pseudoword reading	√	x	√	√	8
Multiple deficit except for word spelling	√	√	x	√	4
Multiple deficit except for pseudoword spelling	√	√	√	x	6
No deficit	x	x	x	x	0
Total struggling readers					128

Note: Profiles of struggling readers based on the 16th percentile criterion applied on the scores of typical readers, √ = presence of a deficit, x = absence of a deficit

When we interpreted both the profile tables, it was clear that the majority of struggling readers exhibited reading as well as spelling deficits, supporting the co-occurrence of these problems. However, we did observe specific patterns of deficits, such as some children had problems with reading only but not with spelling and vice versa. Moreover, there were few children with pseudoword reading and spelling issues but their word reading and spelling was not impaired, and there was only one child with the inverse condition. This indicates that phonological processing is probably more problematic for children struggling to read Urdu than the word recognition. We also found children who had problems with only one of the four skills while other three being unimpaired. Overall, our results support the argument that there are various profiles of dyslexia comprising of combinations of single and multiple deficits.

Table 2.16: Profiling of struggling readers based on 10th percentile

Profiles	Deficit: Reading Words	Deficit: Reading Pseudowords	Deficit: Spelling Words	Deficit: Spelling Pseudowords	No. of struggling readers
Multiple deficit	√	√	√	√	84
Isolated reading deficit	√	√	x	x	6
Isolated spelling deficit	x	x	√	√	7
Word reading and spelling deficit	√	x	√	x	1
Pseudoword reading and spelling deficit	x	√	x	√	4
Multiple deficit except for word reading	x	√	√	√	10
Multiple deficit except for pseudoword reading	√	x	√	√	3
Multiple deficit except for word spelling	√	√	x	√	1
Multiple deficit except for pseudoword spelling	√	√	√	x	6
Other combinations of deficit					6
No deficit	x	x	x	x	0
Total struggling readers					128

Note: Profiles of struggling readers based on the 10th percentile criterion applied on the scores of typical readers, √ = presence of a deficit, x = absence of a deficit

2.6 Discussion

The present study had three main objectives: (1) to develop and validate an assessment battery for young children to identify reading disability in Urdu, (2) to determine the reading and spelling profiles of typical readers (aged 7-11 years) and their struggling peers and (3) to test the applicability of the DRM to Urdu, which has a dual orthography. Our analyses of the data of 295 children showed the dyslexia battery to be reliable and valid as it clearly distinguished struggling from typical readers.

2.6.1 What Distinguishes Struggling from Typical Readers?

The scores on all tests were significantly lower for the struggling readers. Our study based on Urdu, representing the dual orthography, showed that the difference between typical readers and struggling readers was larger for reading pseudowords than for reading words. All children had more difficulties reading and spelling pseudowords than words. This *lexicality effect* was also reported in earlier studies (Landerl et al., 1997a; Rack, Snowling & Olson, 1992; Ziegler, Perry & Coltheart, 2003). Moreover, our finding is consistent with the literature on pseudoword reading deficits in consistent orthographies like German, where the effect manifests itself more in speed than in accuracy, and in less transparent languages, such as English, where accuracy is more impaired than speed (Landerl and Wimmer 2008; Ziegler et al., 2003). In DRM terms, this would suggest that, on average, our struggling readers showed a greater deficit in letter-to-sound conversion process than in direct word recognition.

Effect sizes of the group differences for letter-sound knowledge and, phoneme and syllable deletion were high, suggesting that poor phonological awareness (PA) plays an important role in the delay of reading acquisition in the struggling readers (Ehri & Snowling, 2004; Lyon et al., 2003; Ramus et al., 2003). Supporters of the phonological deficit hypothesis claim that dyslexia arises from a cognitive deficit in PA, making grapheme-to-phoneme conversions difficult, hampering reading acquisition (Bishop & Snowling, 2004; Lyon et al., 2003; Snowling, 2001; Vellutino et al., 2004; Ziegler & Goswami, 2005). Comparing our results for Urdu with those reported for Arabic, which also has a dual orthography and, like Urdu, is more (or less) consistent depending on the presence (or absence) of diacritics, we again see that PA is highly predictive of reading performance in typical and poor grade 1-5 readers (Abu-Rabia, Share, & Mansour, 2003; Al-Mannai & Everatt, 2005). Also in other languages PA is reported as one of the most important predictors for later literacy (Ehri, 2004; Pennington et al., 2012; Rath, 2001; Snowling & Melby-Lervag, 2016; Troia, 2004).

Effect sizes for spelling were the highest followed by those for the reading measures. Spelling and reading are highly connected and share the use of letter-sound conversions (Ehri, 2000). For reading, effect sizes were higher for accuracy than for speed, which is congruent with previous studies (e.g., Callens, Tops, & Brysbaert, 2012). High correlations (Tables 2.11 and 2.12) between reading and spelling tests confirm that reading and spelling impairments often co-occur (Katzir et al., 2006), as is also mentioned in many definitions of dyslexia (e.g. DSM-5; APA, 2013).

Finally, RAN and vocabulary had the lowest effect sizes of all tests, although they still were medium to high, signifying that the two factors do explain reading variability to some extent, which is consistent with previous findings (Furnes & Samuelsson, 2011; Kail, Hall, & Caskey, 1999; Kirby, Parrila, & Pfeiffer, 2003). The RAN-digit subtest (Urdu ES = 0.77, English ES = 0.72) was a stronger predictor of reading than the RAN-object subtest (ES = 0.64), again consistent with the literature (Neuhaus, Foorman, Francis, & Carlson, 2001).

The specific role of vocabulary in reading and spelling acquisition is still under debate and the current evidence is mixed. It was suggested that its contribution to reading and spelling varies across age groups and/or reading stages, where it seems a strong predictor in early reading acquisition but becomes less predictive with increasing proficiency (Vaessen & Blomert, 2010). Arguably, the medium effect sizes for vocabulary may be explained by the fact that our participants had received at least two years of formal reading instruction before participating in this study. Highly likely they acquired a substantial level of vocabulary that is reflected in the close-to-ceiling scores of both groups.

2.6.2 Assessing Dyslexia in Urdu

Our test battery was found to have high reliability and validity, and it also discriminated well between typical and struggling readers as a group. We ran multiple analyses based on different criteria, showing that the battery's sensitivity and specificity are overall very high (> 0.9). Using the internationally recognized 16^{th} percentile criterion (DSM-5), we found that all the children with a confirmed diagnosis of dyslexia in English were screened out, demonstrating that this external criterion is applicable to our assessment battery and supports its clinical relevance. Still, when we considered all the 128 struggling readers using the 16^{th} percentile cut-off score, the specificity rate remained very high (0.9) but the sensitivity rate dropped substantially (< 0.6), implying that not all the struggling readers were screened out. There can be two possible reasons for that. First, the SR group probably did not only consist of children with dyslexia but included in the SR group might be readers struggling due to delayed reading and/or delayed language development, having psychological problems and/or less favourable backgrounds (e.g. low parental socioeconomic status). Second, our findings point towards the heterogeneity of dyslexia. Reading and spelling impairments are manifested so differently that every child with dyslexia possesses a unique combination of symptoms. The large variance that is exhibited in our data of struggling readers could be explained by this heterogeneous nature of dyslexia.

In addition to group average comparisons, we explored the variability within the struggling reader group because our test allowed the assessment for specific patterns of reading and spelling deficits. Our results showed evidence for different profiles in dyslexia such as isolated reading and/or spelling deficits, and isolated problems in reading and spelling words and/or pseudowords. More importantly, our findings support the claim that the DRM is applicable to Urdu, not only as a model for typical reading but also for impaired reading. Moreover, children showing isolated pseudoword deficits were twice as many as children showing isolated word deficits in our sample. Hence, this indicates that phonological deficits may be more common than orthographic deficits in Urdu struggling readers.

2.6.3 Education and Literacy Instruction in Pakistan

In a developing country like Pakistan, the school situation differs from that in developed countries. Parents, teachers and school management often lack an awareness and understanding of the learning process and behavioural problems. The poor quality of teaching and learning, un(der)trained teachers and outdated curricula compromise an adequate identification of children with specific learning disabilities like dyslexia and those that 'merely' lag behind in reading and writing. Clearly, the general public needs to be better informed, while parents, teachers and school-management staff need to be educated and trained in the recognition, assessment and remediation of literacy problems and dyslexia (Khan et al., 2011).

Other complicating factors are the scarcity in resources and lack of specialized, approved services that could perform targeted assessments, while existing remedial services