

Assessing Literacy in Children and Adolescents

Jane Hurry & Estelle Doctor

Psychology and Human Development, Institute of Education, 25 Woburn Square, London

WC1H 0AA. E-mail: j.hurry@ioe.ac.uk; e.doctor@ioe.ac.uk

Correspondence should be addressed to Jane Hurry

This review provides a framework for examining the assessment of literacy underpinned by current theories of reading and writing. Our choice of assessment tools was influenced by a desire to identify tests that are suitable for diagnosing students of all ages. The key dimensions identified were printed word recognition (lexical and nonlexical), comprehension and writing. We conclude that there is no single test currently available that provides a comprehensive profile of literacy difficulties. The assessor will require critical discrimination in choosing tests that are theoretically sound and diagnostically useful.

Keywords: Assessment; reading; writing; diagnostic; age range

Our purpose in this paper is to present a theoretical framework for the selection of tests of literacy suitable for students of all ages and to review some assessment packages suitable for diagnostic usage. Many tests of reading and spelling are a-theoretical. They provide a reading or spelling 'age' but not a deeper understanding of the individual's literacy difficulties. Our focus here is to identify tests that can inform us of the underlying reasons for delays in literacy and provide insights into ways of addressing students' difficulties. To do this we match assessment packages to theories of reading development. We limit ourselves, where possible, to reporting on standardized tests which can be used across the age range, that is, the *same* test can be used to assess a 6-yr old and an undergraduate. We meet this challenge, firstly, because of the growing number of adults requiring diagnostic assessment as they pursue further and higher education; secondly, because diagnosticians and researchers may need to assess the same student at different ages; and thirdly to simplify the test selection process for diagnosticians. The tests we describe focus on what we argue are the three critical dimensions of literacy: printed word recognition, reading comprehension, and writing.

Printed word recognition

Recently, tests have been developed that take account of theoretical developments in the psychology of reading. One such theory, the dual-route cascaded (DRC) model (Jackson & Coltheart, 2001) has been particularly effective in accounting for various forms of reading difficulties across the age range: acquired dyslexia, beginning reading, developmental reading difficulties, and precocious reading (Laxon, Masterson, & Coltheart, 1991; Stuart & Coltheart, 1988; Stainthorp & Hughes, 1998). Dual route theories are concerned primarily with the reading skills required for "pronouncing or recognizing the meaning of individual words, or of word-like strings of letters" (Jackson & Coltheart, 2001, p. 2). The two routes referred to are: the automatic lexical route that links printed words directly to their meanings and the nonlexical route that converts letters to sounds (Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001; Zorzi, Houghton, & Butterworth, 1998). The lexical route is used successfully to read known words but fails to process novel words. The nonlexical route copes with novel items, provided that there is a transparent relationship between their spelling and their sound, but fails with irregular words.

There is now little dispute about the role of the nonlexical route in reading. One of the most common causes of literacy difficulties is a deficit in phonological processing, critical to the decoding of the nonlexical route (Hulme & Snowling, 1992; Stanovich & Siegel, 1994; Frith, 1995; Jackson & Coltheart, 2001; Rack, Snowling, & Olson, 1992; Siegel, 1989). These difficulties are associated with problems in the identification, classification, segmentation, blending and manipulation of phoneme segments in words. Many children who still have reading difficulties at the end of primary school do so because they have failed to understand the alphabetic principle and have difficulty applying phonological strategies.

Wagner, Torgesen, and Rashotte (1999) have distinguished three components of phonological processing: phonological awareness, phonological memory, and rapid naming. Phonological awareness involves both segmentation and blending. Lack of awareness of the phonological structure of oral language and of the relations between written and spoken language explains why some children find learning to read difficult. Phonological memory is the ability to store sound-based information for short periods of time in a part of memory called the phonological loop and this improves the reader's chances of decoding words not seen before, or long words. While poor phonological memory is unlikely to impair the reader's comprehension of simple sentences, it is likely to impair both reading and listening comprehension for more complex sentences and text. Rapid naming tasks are designed to assess speed of phonological production, involving the retrieval of phonological information from long term memory which would affect reading fluency. These tasks are a standard element of diagnostic reading assessments, however, there is a lack of clarity about exactly what they measure (Cattling and Denckla, 2001; Torgesen et al, 1997), and this is the subject of current research (Stuart and Stainthorp, personal communication, Institute of Education, January 2006). Poor performance on this type of test has been found to be characteristic of those who have been diagnosed as dyslexic (Badian, 1994; Wolf and Obregón, 1992; note Olson, 1995 failed to replicate this result) and it has been proposed that naming speed is a measure phonological production speed (Frederickson, Frith and Reason, 1997). Alternatively, rapid naming tasks may simply measure general speed of processing (Kail and Hall, 1994; Kail et al, 1999).

The non-lexical route alone is not enough to support good reading and spelling. Many people with literacy difficulties are hampered by an over-reliance on sounding out during reading and miss-spell words using phonic strategies. The other weakness in printed word recognition involves the lexical route. Many poor readers of all ages have a severely restricted vocabulary of words they can recognise fluently and automatically by sight. They have difficulty matching a word on the page with its representation stored in long-term memory. When little effort is required to identify sight words, the reader is able to concentrate more effectively on the complex processes involved in constructing the meaning of text (Gough, 1996).

Parallel distributed models (Plaut, McClelland, Seidenberg & Patterson, 1996; Seidenberg & McClelland, 1989), in common with dual route models, are concerned with printed word recognition, rather than comprehension. In contrast to the DRC model, they claim to account for reading in terms of a single processing mechanism which manages reading input based on associations built up from past experience (not unlike Skinnerian behaviourism). However, similar to the DRC model, this model also characterises reading as involving two pathways, a phonological pathway and a semantic pathway (Plaut, McClelland, Seidenberg & Patterson, 1996), where the phonological pathway is dominant in processing phonically regular words and the semantic pathway contributes more to the reading of exception words. Jackson and Coltheart (2001) argue that the DRC is superior to many of the so-called one route theories of reading because these models cannot explain a range of findings relating to exception and pseudoword reading, lexical decision making, or acquired surface dyslexia (see also Rastle and Coltheart, 1999a, 1999b and Stuart, 2002). Furthermore, some of the “connectionist” models are essentially dual-route models, with only one route so far implemented (from print to phonology).

Whilst dual route theories of reading offer a powerful way of understanding decoding, they do not address comprehension, a dimension of literacy that becomes increasingly important as readers mature.

Comprehension

Efficient word reading skills are critically important for effective reading comprehension, but not sufficient (Perfetti, Marron and Foltz, 1996; Oakhill and Yuill, 1996, Bishop and Snowling, 2004). Yuill and Oakhill (1991) estimate that as many as 10% of children have difficulties with comprehending text *despite* being able to decode the words. The difficulties experienced by these children are easier to overlook than problems with decoding, and teachers frequently consider such children to be good readers (Nation, Clarke, & Snowling, 2002).

According to Gough (1996) reading comprehension (RC) is the product of word reading efficiency (WRE) and the ability to comprehend language (LC) or $RC = WRE \times LC$. This conceptualisation of comprehension complements well the DRC model for printed word recognition. However, whilst this is a sound view, it is a little too simplistic for our purpose of identifying a framework for assessment. Children with poor language skills tend to have problems both with word recognition and reading comprehension (Bishop and Snowling, 2004). Bishop and Snowling (2004) suggest that those who can decode but cannot comprehend have good phonological skills but impaired semantic and syntactic skills, such as weak vocabulary and limitations of semantic knowledge (Nation and Snowling, 1998, 1999). However, the identification of these lower level skills fails to capture the essence of the poor reading comprehender.

Poor reading comprehenders (who can decode adequately) do not appear to have difficulties with all aspects of language comprehension (eg. receptive grammar, Oakhill and Yuill, 1996) and while they tend to be deficient in lower level skills, such as vocabulary knowledge, impairments in higher level cognitive skills are particularly characteristic. Poor comprehension skills are associated with poor working memory and this interferes with all aspects of comprehension recall (Oakhill and Yuill, 1996; Perfetti, Marron and Foltz, 1996). However, when given access to the text, poor comprehenders have particular difficulties with making inferences, rather than answering literal questions, with integrating information to form a coherent understanding and with comprehension monitoring (Cain, Oakhill, Barnes and Bryant, 2001; Rosenshine, 1980; Oakhill, 1994; Perfetti, Marron & Foltz, 1996; Cataldo &

Cornoldi, 1998). Difficulties with inferences manifest themselves both in reading and listening tasks. Poor comprehenders experience difficulties both in 'coherence' inferences, which require the reader to integrate information given within a text, and 'elaborative' inferences, which require the integration of information from the text with prior or general knowledge. Possibly as a consequence of their comprehension problems, they tend to have gaps in their world or domain knowledge (Perfetti, Marron & Foltz, 1996).

Finally, there are differences between listening and reading comprehension tasks. There are significant differences between spoken and written language, with printed texts being more formal, containing longer and more complex sentences and more dense information (Perfetti, 1985). Certain skills may be particular to each form of comprehension. For example, Cataldo and Cornoldi (1998) discuss the importance of being able to reflect back to the text in a reading task, and indeed criticise reading comprehension tests that do not allow the reader this opportunity.

Tests measuring reading comprehension should ideally focus particularly on these higher order skills. Unlike tests of word recognition, comprehension is to some extent a negotiation between author and reader. It is also more influenced by knowledge of the context. This means that tests of comprehension tend to be more culturally specific than tests of decoding, and therefore less easily transportable across cultures and time.

Writing

The subskills involved in writing differ from those in reading (Stainthorp, 2002). Berninger et al.'s model (1995) defines the processes in writing as planning, translating and reviewing. In addition, good writers need to be able to form their letters, words, numerals and sentences in a legible manner. They should generate meaningful sentences that express their thoughts, feelings and opinions adequately. Compliance with accepted standards of style, especially those governing punctuation, capitalisation and spelling is expected and they should be able to use acceptable syntactic, morphological and semantic elements to express their ideas.

Their thoughts should be expressed in a creative and mature manner. The challenge remains to develop a test which incorporates all these processes. |

Selection of tests for review

To summarise, the ideal literacy assessment package should include tests of both lexical and nonlexical processing, comprehension, and writing. We review eight tests which cover these key dimensions of literacy across a broad age-range. They are shown in table 1. In selecting tests to review we favoured those that covered the age range from childhood to young adulthood. We included the Phonological Assessment Battery despite the fact that it is only standardised up to 14:11 years because it is the best available test of phonological processing with UK norms. Neither of the reading comprehension tests included are standardised for those aged 17 years or more because there are no tests for this older age group which allow the assessment of reading comprehension unconfounded with printed word recognition. Tests suitable for young adults which measure reading comprehension but which confound comprehension with printed word recognition are Progress in English (Kispail, Hagues and Ruddock, 1994) and the Adult Test of Reading and Writing (NRDC, 2004).

Table 1 around here

Assessments of printed word recognition

Comprehensive Test of Phonological Processing [CTOPP] (Wagner, Torgesen, & Rashotte, 1999)

Key Stages 1-4+

The CTOPP is suitable for ages 5:00 to 24:11 years. It measures three reading related phonological skills (phonological awareness, phonological memory and rapid naming) and provides an index of phonological processing strengths and weaknesses.

The test takes approximately 30 minutes to administer. Some subtests are read to the examinee by the examiner; others are relayed via an audio recorder. All ages take the same

tests of Elision, Blending Words, Memory for Digits and Nonword repetition. There is a test of sound matching for 5- and 6-year olds. Examinees younger and older than 7 have their own tests of Rapid Naming: Rapid Colour Naming and Rapid Object Naming for 5- and 6-year olds, and Rapid Digit Naming and Rapid Letter Naming for older examinees. There is a supplemental Nonword Blending task for all examinees, and supplemental tests of Phoneme Reversal and Segmentation of words and nonwords for older examinees.

Results are calculated in terms of standard scores, percentiles and ages and are given for individual subtests as well as composites for Phonological Awareness, Phonological Memory and Rapid Naming.

The test allows the examiner to:

- identify individuals who are significantly below their peers in important phonological abilities
- identify reception and KS1 pupils whose deficiencies in phonological processing could leave them at risk for reading failure
- use the outcome of the identification to develop early intervention designed to promote phonological awareness and to reduce the incidence of reading failure
- determine in older students strengths and weaknesses among developed phonological processes
- document individual's progress in phonological processing as a consequence of special intervention programmes

The CTOPP was normed in the USA between 1997-8 on a sample of 1,656 participants. Norms are presented in terms of standardised scores. Composite scores (Phonological Awareness, Phonological Memory, Rapid Naming) are based on a combination of standardised scores. Percentiles, age and school grade equivalents are also provided.

Its reliability coefficients are between .7 and .9, and the manual also demonstrates satisfactory validity, but it is an American test and this raises a problem because several of the subtests are played to the examinee via a tape recorder or CD. The speakers have

American accents. Older examinees who are not American seem to adjust to the accent, however, in an ongoing study (Stuart and Stainthorp, personal communication), it has been found that children are performing below the test norms on those sub-tests involving audio recording and users should be aware of this issue when using the CTOPP.

The CTOPP slots well into the DRC framework. Phonological processing is an essential component of literacy and, as discussed above, the phonological processes assessed by this test are essential precursors of literacy.

*Phonological Assessment Battery [PhAB] (Frederickson, Frith & Reason, 1997)**Key Stages 1-3+*

The PhAB offers a UK normed alternative to the CTOPP for children and young people aged 6:00 to 14:11. The assessment battery includes measures of phonological awareness/blending (Alliteration Test, Rhyme Test and Spoonerisms Test), semantic/phonological memory (Fluency Test, semantic, alliteration and rhyme), rapid naming (Naming Speed Test) and phonological decoding (The Non-Word Reading Test). In the standardisation sample, performance on the Alliteration Test was approaching ceiling at age 10 and this test would be most useful for younger children. The Non-Word Reading Test is a transparent test of decoding with no other factors such as short-term memory influencing performance. It achieved the highest correlations with the Neale Analysis of Reading – Revised (NARA II) in the standardisation sample (.72 with Neale Accuracy and .60 with Neale Comprehension). The Spoonerism Test showed the next highest correlations with the Neale (.63 with Neale Accuracy and .61 with Neale Comprehension). The Spoonerism Test measures analysis and synthesis of onsets and rimes but it cannot be assumed that it measures aural phonological awareness, examinees often appear to be constructing words in their head which they then 'read' back to the examiner. As with the CTOPP, it gives insight into the non-lexical dimension of the DRC framework. To administer the full battery of tests takes between 30-40 minutes.

The test was standardised in conjunction with the NARA II. The PhAB was administered to a sub-sample of 629 children across the age range (Yrs 1-9). Reliability as measured by internal consistency was adequate with most alpha coefficients above .8. It is difficult to demonstrate the validity of such a test but it can be estimated by correlations with the NARA II, which were reasonably high, but not perfect, to be expected as the two assessments measure different dimensions of reading.

Test of Word Reading Efficiency [TOWRE] (Torgesen, Wagner & Rashotte, 1999)

Key Stages 1-4+

The TOWRE, suitable for ages 6:00 to 24:11 years, measures two kinds of word reading skill that, according to the DRC model, are critical in the development of overall reading ability: the ability to sound out words quickly and accurately, and the ability to recognize familiar words by sight. These abilities are associated with the lexical route. The test is rapidly administered: after a brief practise trial, the examinee is requested to read through a list of words, for 45 seconds. The procedure is repeated for nonwords. Scores are recorded in terms of standard scores, percentiles and ages.

The TOWRE provides a quick probe of weaknesses in these word reading and decoding skills that are particularly diagnostic of reading problems at every age level. One of its strengths is that it provides an indication of reading under time constraints, indicative of how the reader will cope under exam conditions or when there are deadlines to meet. It can also be used:

- For early identification and to prevent the emergence of serious reading difficulties and the development of negative attitudes to reading;
- To devise intervention that is structured, systematic and explicit, including training in
 - Phonemic awareness
 - Letter-sound correspondences
 - Blending skills
 - Rules
 - Use of context to help specify a word once it is partially or completely phonemically decoded
 - Strategies for dealing with multi-syllabic words
 - Automatic recognition of high frequency irregular words.

All the above should be embedded in opportunities for meaningful reading and writing experiences.

Deficiencies in fluency may sometimes occur independently of problems with accuracy. Older children who have reading disabilities may read accurately, but not fluently. Intervention programmes that improve accuracy may not improve fluency. The TOWRE should always be administered alongside an untimed test of word recognition, such as the WORD or WRAT3 (see below), to ensure that both reading accuracy and latency are addressed.

The TOWRE was normed on the same sample as the CTOPP. As with the CTOPP, norms are presented in terms of standardised scores. Percentiles, age and school grade equivalents are also provided.

Reliability of the test was measured in terms of content, time and scorer and is above .9.

There are two comparable forms of the test. Validity is also satisfactory.

The TOWRE slots well into the DRC framework. A poor score on the measure of phonemic decoding relative to the sight word efficiency score is indicative of difficulties with the sublexical route (sometimes referred to as phonological dyslexia); a poor score on the measure of sight word efficiency relative to phonemic decoding indicates difficulties with the lexical route (sometimes referred to as surface dyslexia). If both are weak then there is a double deficit in sight word recognition and phonological processing.

Wide Range Achievement Test 3 [WRAT 3] (Wilkinson, 1993)

Key Stages 1-4+

The WRAT 3, suitable for ages 5:00 to 74:11 years, is an untimed single word test of reading and spelling (also arithmetic, not relevant in this context). It was developed in the USA and standardised on a sample of 4433 people. The candidate reads aloud or spells a list of words that are then scored for accuracy. The test has two parallel forms and scores are recorded in terms of standard scores, percentiles and grade equivalents. The test is a-theoretical: most of its words are regular in letter-sound correspondence. There is also, however, a confounding of regularity with frequency which makes it difficult to interpret the results within the dual route framework and draw diagnostic conclusions about processing strategies. The value of the WRAT 3 is that it assesses how well the examinee can function when there are no time

constraints and therefore provides a useful additional dimension of single word reading to supplement the TOWRE

Reliability of the WRAT 3 was measured using internal consistency, comparison of parallel forms and test re-test. Values of reliability were all above .9. Validity was also explored in a number of ways and found to be satisfactory.

The WRAT was not designed with any current model of reading as a framework, but a theoretical interpretation may be superimposed on the findings.

Assessments of comprehension

Neale Analysis of Reading Ability – Second Revised British Edition [NARA II] (Neale, 1997)

Key Stage 1-2+

The NARA II is standardised for ages 6:00 to 12:11 years, but is described in the manual as being suitable for older readers if used diagnostically. However, the story content would be unappealing to most older readers.

The test takes about 30 minutes to administer. The main test consists of six graded passages which are used to assess reading accuracy, rate (sometimes referred to as fluency), and comprehension. These are offered in two parallel forms. The reading accuracy test differs from tests of word reading as in the TOWRE or the British Ability Scales, because words are read in context. However, the particular reason why the NARA II is covered in this review is because it offers a measure of comprehension which is fairly unique. The assessment of comprehension is plagued by the fact that it is frequently confounded with decoding. This is true of all the group administered reading tests which claim to measure comprehension, and also of some individually administered tests. In the NARA II, as the child

reads the passage, the examiner provides the words that the child cannot decode. The child's own reading is of course used to score reading accuracy, however, the prompting of the correct word allows the child to hear the whole passage. The questions at the end of the passage can then measure comprehension, not decoding. If the child makes more than 16 errors (20 for the final passage) comprehension should not be assessed. Whilst in most circumstances this is sensible, the tester needs to be cautious where a child has difficulty reading aloud, perhaps because of some articulatory problem. In this case, weakness in word reading accuracy places an artificial ceiling on comprehension. The weakness of the NARA II is in its content validity. The questions are fairly limited in the aspects of comprehension assessed. Around 65% of the questions involve the simplest level of comprehension, literal comprehension (e.g. Text: A black cat came to my house. Question: What came to the little boy's house?). The remaining questions address only fairly simple inferences (e.g. Text: A surprise parcel arrived...Peter looked at the strange stamps. Question: How do you know the parcel came from another country). There is almost no measurement of the important skills of integration, comprehension monitoring or identifying the main idea of a passage.

The second revised British edition was standardised in 1996 on a sample of 3,474 children aged 6 to 12 years. Reliability was good for accuracy and comprehension but slightly less satisfactory for rate (correlation between parallel forms = .66). Validity was also satisfactory.

Wechsler Objective Reading Dimensions (WORD) (Rust, Golombok, & Trickey, 1993)

Key Stages 1-4

The WORD was designed to offer an overall reading measure that was easily comparable with the Wechsler Intelligence Scale for Children, third edition UK (WISC-III^{UK}; Wechsler, 1992).

The WORD is designed for the assessment of children aged 6:00 to 16:11 years and takes about 20 minutes to administer. It has three subtests: Basic Reading, Spelling and Reading

Comprehension, thus covering all our headings to some degree. The Basic Reading test begins with four items assessing the match between letters and beginning and end sounds, three items matching pictures to words and then 48-item word reading test with a fairly steep gradation of difficulty, as might be expected from a short test spanning a wide age group. The Spelling sub-test has 6 items assessing the ability to write letters and 44 items assessing spelling. The word reading and spelling components of the WORD are similar to a number of other tests, such as those offered by the British Ability Scales. However, the Reading Comprehension component is more unusual. As with the NARA II, reading comprehension is assessed by asking the examinee a set of passages, graded by difficulty, before asking a series of questions. There are 38 such questions. The WORD has a number of advantages over the NARA II and some disadvantages. The advantages are that it is designed to work alongside an IQ test, that it covers quite a wide age range, that the person being assessed is allowed to read the text silently if they wish, arguably more 'authentic' than being required to read aloud and that it tests a wider range of comprehension skills than the NARA II. Around 30% of the items assess literal comprehension skills, the other 70% involve simple or complex inference or identifying the main idea of the story. In the WORD, Reading Comprehension is confounded with word reading. The person taking the test can read silently or out loud and has access to the text whilst being questioned, but is not given help with decoding, as in the NARA II. This limits the use of the comprehension sub-scale as a stand-alone measure of reading comprehension, though comparison with the Basic Reading test allows the identification of discrepancies in levels of skill. Finally, as the WORD offers assessment for a fairly wide age range with a limited number of items, assessment at the lowest and highest ends of the distribution are fairly crude, a floor and ceiling effect can be observed.

The UK edition was standardised in 1991, using a sample of 794 children aged 6 to 16 years and closely approximated socio-economic and ethnic composition in the UK at that time. The reliability and validity have been thoughtfully explored. Coefficients for internal consistency and test-retest reliability are above .9. The WORD correlates highly with the appropriate dimensions of a similar test, the Differential Ability Scale (DAS, an American adaptation of the British Ability Scales; Elliott, 1990). It is useful to note that whilst intercorrelations between the

two tests on word-reading are .82, and between the two tests on spelling are .86, correlation between comprehension on WORD and word reading on the DAS are much lower at .42, confirming that comprehension does not follow inevitably from word reading (Shankweiler, 1999).

Test of Adolescent and Adult Language (TOAL-3) Hammill, Brown, Larsen, & Wiederholt, 1996)

Key Stages 3-4+

The test is suitable for ages 12:00-24:11 years. The aims of this test are:

- to identify students whose scores indicate that they are significantly below their peers and who might benefit from intervention;
- to determine areas of relative strength and weakness in literacy and language abilities;
- and to document progress in literacy and language development as a consequence of intervention.

More specifically the test investigates aspects of spoken and written language such as vocabulary, grammar, receptive language (reading and listening) and expressive language (writing and speaking).

The TOAL-3 was normed in 1993 on a sample of 3,056 persons in 26 states of the USA. Norms for the subtests are presented in terms of standard scores. Composite scores are available for General Language, Written Language, Spoken Language, Vocabulary, Grammar, Receptive Language, Expressive Language, Listening, Speaking, Reading and Writing.

The test meets the necessary criteria for reliability within the American context and appears to have content validity, criterion-related validity, and construct validity.

Assessments of writing

Test of Written Language [TOWL] (Hammill, & Larsen, 1996).

Key Stage 2-4+

Most of the writing abilities mentioned in Berninger et al's model can be assessed by using the TOWL. This test is suitable for ages 7:00-17:11 years. The subtests are as follows::

1. *Vocabulary* measures knowledge of word meanings and classes through the writing of meaningful sentences.
2. *Spelling* measures the ability to spell dictated words.
3. *Style* measures the ability to punctuate sentences and capitalize words properly.
4. *Logical Sentences* measures the ability to recognize and correct through rewriting the illogicalities existing in stimulus sentences, such as: *I ate the water during lunch.*
5. *Sentence Combining* measures the ability to incorporate the meanings of several sentences into a comprehensive single sentence containing embedded phrases, introductory clauses, adjective sequences, and so on.
6. *Contextual Conventions* measures the ability to spell words properly and to apply the rules governing punctuation of sentences and capitalization of words in a spontaneously written composition.
7. *Contextual Language* measures the ability to use mature words that represent a variety of parts of speech; complex sentences comprised of introductory and concluding clauses, embedded phrases, and adjective sequences and grammatical forms such as subject-verb agreements.
8. *Story Construction* measures the ability to write in a logical, organized fashion; to generate a specified theme or plot to develop a character's personality; and to employ an interesting and engaging prose.

Usage of these tests allows the examiner to evaluate written language strengths and weaknesses and to develop intervention programmes suitable for individual needs.

The TOWL-3 was last standardised in 1993 on a sample of 3,056 persons residing in 26 states of the USA. Norms for the subtests are presented in terms of standardised scores. Quotients are used to derive composite scores for Overall Writing, Contrived Writing and Spontaneous Writing.

Percentiles and Age and School Grade equivalents are also provided.

The test meets the necessary criteria for reliability (.90) within the American context and appears to have content validity, criterion-related validity, and construct validity.

Discussion

We have reviewed an array of tests selected on the basis of providing a theoretical coherent coverage. None of them cover all the processes involved in literacy. Some of the tests look specifically at a particular aspect of the literacy process, such as the TOWRE which focuses on latency of word recognition (lexical and nonlexical), or the CTOPP which focuses on underlying phonological processes. Others include a mixture. The WRAT3, for example, assesses word recognition and spelling but does not distinguish lexical and nonlexical processes. Nor does the WORD which also assesses word recognition, spelling and comprehension. The drawback of the WORD comprehension test is that comprehension is confounded with word recognition. The NARA II which assesses word recognition, reading rate, and comprehension avoids this problem, but predominantly assesses literal comprehension skills and penalizes candidates who have difficulty reading aloud. The TOAL assesses expressive and receptive language. The picture format used for the listening vocabulary task is difficult for some candidates to process; the reading tasks do not measure lexical and nonlexical processes; the prose tests evaluate syntactic processing but do not address free recall or recognition memory. The TOWL measures writing ability but does not address all the criteria delineated in the Berninger et al. model (1995). At present, in order to

thoroughly assess the key theoretical dimensions of literacy, the assessor must utilize a wide array of tests and exercise a degree of critical discrimination.

Some students have had difficulty building first an auditory and later a reading vocabulary. Blending and/or segmentation difficulties may be paramount, forcing them to rely on their visual memory for what words look like. In some cases this further restricts their word recognition abilities. Other students lack awareness of the phonological structure of oral language and of the relations between written and spoken language. For all these students, their poorly developed phonological awareness, coupled, possibly with early speech difficulties and language difficulties, explains why they find or have found learning to read and spell difficult. Their printed word identification difficulties are likely to arise in childhood and usually result in a restricted sight vocabulary of words they can read fluently and automatically. Ultimately it is this difficulty in rapid word recognition that limits their comprehension. In terms of the DRC model they lack the phonemic decoding skills that would enable them to identify words they do not recognize by sounding them out.

Some students have difficulty comprehending texts despite being able to read single words accurately. Their difficulties tend to be particularly located in the higher order skills of making inferences, making a coherent sense of a text and monitoring their comprehension. Unfortunately, these skills are currently not adequately assessed in comprehension tests. Also, we have identified an absence of tests suitable for adults which measure comprehension skills independently of word recognition skills.

Accurate diagnostic assessment that offers a profile of a person's current cognitive skills is important for two reasons. First, demonstration of literacy difficulties, particularly specific reading difficulties, is the gateway to a range of valuable services: a statement of special educational need, a right to additional educational provision, access to resources such as subsidized computers, special examination conditions, etc. In this case, it is important to identify the degree of difficulty and standardized norms are useful for this purpose. It may also be necessary to identify a specific condition, such as dyslexia, and this involves the

interpretation of the assessment profile. The typical profile of a person with dyslexia would be particularly poor phonological awareness, as measured by the CTOPP and particularly slow word and nonword reading, as measured by the TOWRE (see Jackson & Coltheart, 2001). One might expect comprehension, as measured by WORD or NARA II to be strong relative to level of word reading accuracy.

A diagnosis based on cognitive profiling is surely preferable to the traditional diagnosis based on the discrepancy between IQ and reading level. Frith (1995) has pointed out that if there is a biological base to dyslexia, for example a dyslexia gene, then dyslexia should vary independently of IQ. This undermines confidence in 'discrepancy' techniques of diagnosis and makes the interpretation of a cognitive profile more important.

Secondly, understanding a student's strengths and weaknesses informs teaching decisions. At its simplest, the teacher identifies what the student already knows and what remains to be learnt. To make decisions about what to teach and how to teach it is also necessary to interpret the information provided by diagnostic assessment. We have suggested a framework which guides not only testing but also instruction by identifying the central cognitive skills. For example, one student, referred for a Disabled Students Allowance (DSA) diagnostic assessment, was tested on the CTOPP. Although her IQ was in the normal range she struggled on all parts of the CTOPP. Her poor score on rapid naming indicated that she was slow at retrieving phonological information from long-term memory and executing a sequence of operations quickly and repeatedly. Intervention should focus on improving the speed and accuracy with which she recognizes words required for her course of study. The causes of reading difficulty are also likely to have a bearing on instruction (Jackson & Coltheart, 2001; Morton & Frith, 1995). For example, one person may experience reading difficulties because of a genetically based dyslexia, another because they missed out on good teaching. It is likely that these two people will require different instruction. In the case of those with a genetically based deficit, it will be necessary to identify the area of difficulty and develop strategies to compensate. For example, they may need to rely more heavily on the lexical route if they have difficulty with phonological decoding. A person who has not been

systematically taught would benefit from general support across the range of key skills. Whilst diagnostic assessment does not reveal the causes of all reading difficulties, it informs about how these manifest themselves. As such it is crucial for a better understanding of reading difficulties. For diagnostic analysis to be more reliable we need to develop tests that are more transparently linked to theory. This is the challenge.

Word count 5,806

References

- Badian, N. (1994). Do dyslexic and other poor readers differ in reading related cognitive skills? *Reading and Writing*, 6, 45-63.
- Berninger, V. W., Abbott, R. D., Whitaker, D., Sylverster, L., & Nolen, S. B. (1995). Integrating low- and high- level skills in instruction protocols for writing disabilities. *Learning Disability Quarterly*, 18, 293-309.
- Bishop, D. V. M., & Snowling, M. J. (2004). Developmental dyslexia and specific language impairment: same or different? *Psychological Bulletin*, 130, 858-886
- Cain, K., Oakhill, J. V., Barnes, M. A., & Bryant, P. E. (2001). Comprehension skill, inference making ability and their relation to knowledge. *Memory and Cognition*, 29, 850-859.
- Cataldo, M. G., & Cornoldi, C. (1998). Self-monitoring in poor and good reading comprehenders and their use of strategy. *British Journal of Developmental Psychology*, 16, 155-165.
- Cutting, L. E., & Denckla, M. B. (2001) The relationship of rapid serial naming and word reading in normally developing readers: An explanatory model. *Reading and Writing*, 14, 673-705.
- Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review*, 108, 204-256.

- Elliott, C. D. (1990). *Differential Ability Scales: Introductory and technical handbook*. San Antonio, TX: Psychological Corporation.
- Frith, U. (1995). Dyslexia: can we have a shared theoretical framework? *Educational and Child Psychology, 12*, 6-17.
- Fredrickson, N., Frith, U., & Reason, R. (1997), *Phonological Assessment Battery- Standardised Edition*. London: NFER-Nelson.
- Gough, P. (1996) How children learn to read and why they fail. *Annals of Dyslexia, 46*, 3-20.
- Hammill, D. D., & Larsen, S. C. (1996). *Test of Written Language. Third Edition*. Austin Texas: Proed.
- Hammill, D. D., Brown, V. L., Larsen, S. C., & Wiederholt, J. L. (1996). *Test of Adolescent and Adult Language. Third Edition*. Austin Texas: Proed.
- Hulme, C., & Snowling, M. (1992). Phonological deficits in dyslexia: A "Sound" reappraisal of the verbal deficit hypothesis? In N. Singh, & I. Beale (Eds.), *Learning disabilities: Nature, theory and treatment* (pp. 270-301). New York: Springer Verlag.
- Jackson, N. W., & Coltheart, M. (2001). *Routes to Reading Success and Failure*. New York: Psychology Press.
- Kail, R., & Hall, L. K. (1994) Processing speed, naming speed, and reading. *Developmental Psychology, 30*, 949-954
- Kail, R.; Hall, L.K. & Caskey, B.J. (1999). Processing speed, exposure to print and naming speed. *Applied Psycholinguistics, 20*, 303-314.
- Laxon, V. Masterson, J., & Coltheart, V. (1991). Some bodies are easier to read: The effect of consistency and regularity on children's reading. effect of consistency and regularity on children's reading. *Quarterly Journal of Experimental Psychology. Human Experimental Psychology, 43A*, 793-824.

- Morton, J., & Frith, U. (1995). Causal Modelling: A Structural Approach to Developmental Psychopathology. In D. Cicchetti, & D. J. Cohen (Eds.), *Manual of Developmental Psychopathology*. Volume 1 (pp. 357-390). New York: Wiley.
- Nation, K., & Snowling, M.J. (1998). Semantic processing and the development of word recognition skills: Evidence from children with reading comprehension difficulties. *Journal of Memory and Language*, *39*, 85–101.
- Nation, K., & Snowling, M.J. (1999). Developmental differences in sensitivity to semantic relations among good and poor comprehenders: Evidence from semantic priming. *Cognition*, *70*, 1-13.
- Nation, K., Clarke, P., & Snowling, M. (2002). General cognitive ability in children with reading comprehension difficulties. *British Journal of Educational Psychology*, *72*, 549-560.
- Neale, M. (1997). *Neale Analysis of Reading Ability – Revised [NARA II]*. Windsor: NFER-Nelson.
- Oakhill, J. (1994). Individual Differences in Children’s Text Comprehension. In M.A Gernsbacher (Ed.), *Handbook of Psycholinguistics* (pp. 821–848). London: Academic Press.
- Oakhill, J., & Yuill, N. (1996). Higher order factors in comprehension disability: Processes and remediation. In C. Cornaldi & J. Oakhill (Eds.), *Reading comprehension difficulties: Processes and Intervention*. Mahwah, NJ: Erlbaum.
- Olson, R. K. (1995). Language deficits in “specific” reading disability. In M. Gernsbacher (Ed.). *Handbook of Psycholinguistics*. San Diego, CA: Academic Press.
- Perfetti, C., Marron, M., & Foltz, P.W. (1996). Sources of Comprehension Failure: Theoretical Perspectives and Case Studies. In C. Cornoldi & J. Oakhill (Eds.) *Reading Comprehension Difficulties: Processes and interventions*, 137-165. Hillsdale, NJ: Lawrence Erlbaum Assoc.

- Rack, J. P., Snowling, M. J., & Olson, R. K. (1992). The nonword reading deficit in developmental dyslexia: a review. *Reading Research Quarterly, 27*(1), 28-53.
- Rastle, K. & Coltheart, M. (1999a). Lexical and nonlexical phonological priming in reading aloud. *Journal of Experimental Psychology: Human Perception and Performance, 25*, 461-481.
- Rosenshine, B. V. (1980). Skill hierarchies in reading comprehension. In R. J. Spiro, B. C. Bruce, & W. F. Frewer (Eds.), *Theoretical issues in reading comprehension* (pp. 535-554). Hillsdale, NJ: Lawrence Erlbaum.
- Rust, J., Golombok, S., & Trickey, G. (1993). *Wechsler Objective Reading Dimensions (WORD)*. UK. Sidcup: The Psychological Corporation.
- Shankweiler, D. (1999). Words to meanings. *Scientific Studies of Reading, 3*, 113-127.
- Siegel, L. S. (1989). IQ is irrelevant to the definition of learning disabilities. *Journal of Learning Disabilities, 22*, 469-479.
- Stainthorp, R. (2002) Writing is hard. *Psychology of Education Review, 26*(1), 3-12.
- Stainthorp, R., & Hughes, D. (1998). Phonological sensitivity and reading: Evidence from precocious readers. *Journal of Research in Reading, 21*(1), 53-68.
- Stanovich, K. E., & Siegel, L. S. (1994). Phenotypic performance profile of children with reading disabilities: A regression-based test of the phonological-core variable-difference model. *Journal of Educational Psychology, 86*, 24-53.
- Stuart, M. & Coltheart, M. (1988). Does reading develop in a sequence of stages? *Cognition, 30*, 139-181.
- Stuart, M. (2002). Using the dual route cascade model as a framework for considering reading development. In R. Stainthorp & P. Tomlinson (Eds.). *Learning and teaching*

reading. Monograph Series II, British Journal of Educational Psychology, Leicester UK: British Psychological Society.

Torgesen, J. K., Wagner, R. K., Rashotte, C.A., Burgess, S., & Hecht, S. (1997).

Contributions of phonological awareness and rapid automatic naming ability to the growth of word-reading skills in second- to fifth-grade children. *Scientific Studies of Reading, 1*, 161-185.

Torgesen, J. K., Wagner, R. K., & Rashotte, C. A. (1999). *Test of Word Reading Efficiency*.

Austin: Texas: Pro-ed

Wagner, R. K., Torgesen, J. K., & Rashotte, C. A. (1999). *The Comprehensive Test of Phonological Processing*. Austin: Texas: Pro-ed.

Wilkinson, G. S. (1993). *Wide Range Achievement Test - 3 (WRAT 3)*. USA: Jastak Assessment Systems.

Wolf, M., & Obregón, M. (1992). Early naming deficits, developmental dyslexia and a specific deficit hypothesis. *Brain & Language, 42*, 219-247.

Yuill, N., & Oakhill, J. (1991). *Children's problems in text comprehension: An experimental investigation*. Cambridge, UK: Cambridge University Press. Zorzi, Houghton, & Butterworth, 1998

Zorzi, M., Houghton, G., & Butterworth, B. (1998). Two routes or one in reading aloud? A connectionist dual-process model. *Journal of Experimental Psychology: Human Perception and Performance, 24*, 1131-1161.

Table 1. Selection of tests

Dimension of Literacy	Test
Phonological processing	CTOPP, PhAB
Word recognition	TOWRE, WRAT3, WORD
Comprehension	WORD, NARA II & TOAL
Writing (including spelling)	TOWL, WRAT3, & WORD