Computerised screening for dyslexia in adults

Chris Singleton and Joanna Horne
University of Hull, UK

Fiona Simmons
Liverpool John Moores University, UK

Identifying dyslexia in adulthood presents particular challenges because of complicating factors such as acquisition of compensatory strategies, differing degrees of intervention and the problem of distinguishing dyslexic adults from those whose literacy difficulties have non-cognitive causes. One of the implications is that conventional literacy measures, per se, do not provide a satisfactory basis for screening for dyslexia in adulthood as some dyslexic adults have above-average literacy skills and some non-dyslexic adults have very poor literacy skills. This study examined an alternative approach to dyslexia screening, using three tests that depend heavily on phonological processing, lexical access and working memory, but which are not conventional measures of literacy. Using these tests, which are computer delivered, 70 dyslexic adults from three different types of educational institutions were compared with 69 non-dyslexic adults from the same institutions. The results showed that the dyslexic and non-dyslexic groups were significantly different on all three computer-based tests, with an average effect size of 1.55. Adaptive versions of these tests were then created to reduce overall administration time for the suite to about 15 minutes. Analysis showed that the combined scaled scores from the adaptive versions of the three tests significantly discriminated the dyslexic from the non-dyslexic group with an increased effect size of 2.07 and with a sensitivity rate of 90.6% and a specificity rate of 90.0%. It was concluded that this approach is a valid and useful method of identifying dyslexia in adulthood, which, given the ease of administration to large numbers of adults, has noted advantages for education and employment.

The use of computers in the identification of children with dyslexia is now well established in UK schools, with a number of different programs available for teachers, ranging from screening software such as the Dyslexia Screener (Turner & Smith, 2003), Lucid Rapid Dyslexia Screening (Singleton, Horne, Leedale & Thomas, 2003) and the Dyscalelia Screener (Butterworth, 2004) to more detailed computerised assessment suites such as Lucid CoPS (Cognitive Profiling System) (Singleton, Thomas & Leedale, 1996), LASS Junior (Thomas, Singleton & Horne, 2001) and LASS Secondary (Horne, Singleton & Thomas, 1999). Most of these systems rely on assessing reading and spelling attainment together with cognitive abilities such as phonological awareness and verbal
memory, which underpin literacy development and which, in general, are good predictors of dyslexia (Snowling, 2000; Vellutino, Fletcher, Snowling & Scanlon, 2004).

When it comes to identifying dyslexia in adulthood, however, there are additional factors that must be taken into consideration, which necessarily complicate any process of screening or assessment. Although they may have struggled at school or may have been late in learning to read, by the time they reach adulthood, people with dyslexia may not necessarily show obvious or profound difficulties with literacy (Beaton, McDougall & Singleton, 1997; Fitzgibbon & O’Connor, 2002). Most will have developed personal strategies to compensate for dyslexic weaknesses, which can mask impairments (McLoughlin, Leather & Stringer, 2002; Stacey, 1997), and some will have received additional help while at school, probably enabling them to improve reading, at least, to within (or approaching) normal levels, even though writing and spelling skills may still be rather weak (Rack, 2004; Riddick, Farmer & Sterling, 1997). Adults such as these can easily remain undetected (false negatives).

A considerable minority of adults have poor literacy skills; the Skills for Life Survey (DfES, 2003) reported that 16% of adults aged 16–65 in England had literacy levels at entry level or below, that is, no better than that expected of an 11-year-old. Five per cent of adults in England were at entry level 2 or below, which means that they are likely to experience difficulty with understanding all but the simplest texts, and are unable to obtain information from everyday sources. The corresponding figures were higher for Scotland, Wales and Northern Ireland. By no means all of these adults who have low standards of literacy will have dyslexia, but they can be difficult to distinguish from less able dyslexic adults (especially if the latter have disadvantaged backgrounds) and so can easily be misclassified in dyslexia screening (false positives). When assessment is being carried out by a psychologist or specially trained teacher, these factors can be taken into account (see Bartlett & Moody, 2000; Kirk, McLoughlin & Reid, 2001; McLoughlin et al., 2002; Turner, 1997), but in the development of computerised assessment or screening tools such factors present particular challenges because broad and often poorly defined factors such as educational background, compensatory strategies and social deprivation cannot easily be accommodated.

Beaton et al. (1997) highlighted some of the hazards that confront psychologists attempting to identify dyslexia in adults who do not necessarily have very poor literacy skills. Dyslexic university students, for example, can often read and spell fairly well, although they are usually slow readers and generally have significant problems with tasks such as taking notes in lectures, writing essays and sitting written examinations (Singleton, 1999; Sterling, Farmer, Riddick, Morgan & Matthews, 1998). One of the earliest cases of this type reported, that of RE (Campbell & Butterworth, 1985), was a female undergraduate who, despite having superficially normal literacy skills, displayed marked deficits in verbal memory, nonword reading and phonological processing. It appears that she learned to read almost entirely by visual strategies, and employed no phonic decoding strategies. Since then, intellectually capable adults with very similar profiles to RE have been reported in the literature (e.g. Funnell & Davison, 1989; Howard & Best, 1997). Hanley (1997) reported results suggesting that dyslexics with superior vocabulary skills may be better able to compensate for their reading problems than those with smaller vocabularies. A further complication arises from the fact that some university entrants – particularly those in the ‘mature’ category – may perform rather poorer than expected on literacy tests because their experience of reading and writing since they left school has been limited.

Most adults with dyslexia have strengths in order to compensate for their reading difficulties. For example, E was significantly faster than controls), and F performed better in tasks requiring visual memory and visuo-spatial skills. These two individuals would be regarded as having dyslexia with a visual-spatial compensation. Johnston & Anderson (1995) found that dyslexics performed less well than controls on a test which measured phonological awareness, a skill which they thought dyslexics would be better able to exploit. When they repeated the test with a version that was less reliant on phonological processing, they performed better.

On the other hand, I received special education in a special school and has an average age range by age range. In order to compensate for their reading difficulties, these individuals may have developed compensatory strategies that allow them to perform well in other areas. For example, A is particularly skilled at visual-spatial tasks, and B is good at recalling visual information. These compensatory strategies can mask underlying reading difficulties, making it difficult for psychologists to identify dyslexia in adulthood.

The presence of compensatory strategies and variable profiles highlights the complexity of identifying dyslexia in adulthood. It is important to consider a range of factors, including individual strengths and weaknesses, in order to accurately diagnose dyslexia. By taking into account these factors, psychologists can better identify dyslexia in adults and provide appropriate support to those who need it.
Most adults (and especially brighter ones) will tend naturally to develop strategies in order to compensate for limitations. In particular, there is evidence that many dyslexics have strengths in visual cognition that could be utilised in a compensatory manner. For example, Ellis, McDougall and Monk (1996) found that dyslexic children were significantly faster on visual processing tasks than other groups (including reading age controls), and Palmer (2000) found that dyslexic teenagers significantly outperformed reading age controls in visual–spatial memory tasks. Poor readers have been found to perform better than reading age controls on word recognition tasks when the words are paired with visually similar cues (Holligan & Johnston, 1988; Rack, 1987). Visual strengths can be utilised when learning to read and spell, which in turn may enhance skills and strategies in visual memory at the expense of those in verbal memory (see Johnston & Anderson, 1998; Singleton, Thomas & Horne, 2000). Conversely, children with good phonological skills, who can more readily form graphophonemic links, need less reliance on visual memory (Stuart, Masterson & Dixon, 2000). In adulthood, dyslexics have been observed to be overrepresented in careers in which visual skills are advantageous, such as architecture, design, engineering and the graphic arts (West, 1997).

On the other hand, it is perhaps because increasing numbers of dyslexic students have received special tuition while at school (although they may not necessarily have been formally identified as having dyslexia) that their literacy skills are often found to be in the average range by the time they reach university, depending on which measure of literacy is employed. Horne and Singleton (1997) reported that of a sample of 72 dyslexic students (age range 18–46, mean age 20.2, mean IQ 108.6) only 21% had single word reading scores that fell below the average range for the general population. This study found that dyslexic students were much more likely to be impaired in spelling and reading comprehension (40% had below average spelling, and 58% had below average reading comprehension). These findings are comparable with those of Everatt (1997), who found that, compared with non-dyslexic students, those with dyslexia tend to show similar performance on tests of single word reading but poorer performance on tests of spelling and reading comprehension. Simmons and Singleton (2000) also found evidence for impaired reading comprehension in dyslexic, compared with non-dyslexic students, on a task involving a lengthy, complex text. This task was believed to have greater ecological validity for studying university students than conventional comprehension tests, which typically employ short passages of text. The dyslexic students took longer to carry out the task, and experienced greater difficulty in processing inferential information, as opposed to literal information. The fact that the dyslexic students could comprehend the literal information just as well as the non-dyslexic students confirms that their word decoding skills were not significantly impaired. It was suggested that these results could be attributed to working memory deficiencies and perhaps a general impairment in comprehending verbal information as a result of less exposure to complex texts.

It is apparent from the foregoing that, although adults with dyslexia typically experience some problems with aspects of reading, writing and/or spelling, nevertheless literacy measures, per se, are inappropriate tools for identifying the individuals who have dyslexia and are particularly unsuited to screening, which, by definition, normally requires a swift, algorithmic process (Singleton, 1997a). One reason for this is that the differences in literacy skills between dyslexic and non-dyslexic persons are largely quantitative rather than qualitative in nature, that is, they tend to score lower on psychometric tests of literacy, and are slower at reading and writing than non-dyslexic.
persons, but do not, in general, display qualitative distinguishing features of dyslexia in their reading, writing and spelling, such as unique types of errors (Rack, 1997). A further reason is that the development of compensatory strategies (especially by individuals with higher intelligence) often masks literacy differences between dyslexic and non-dyslexic persons. If one were to take a group of dyslexic and a group of non-dyslexic adults, statistically significant differences between these groups in reading, writing and spelling would almost certainly be found (e.g. Hanley, 1997; Reid, Szczerbinski, Iskierka-Kasperek & Hansen, 2007; Snowling, Nation, Moxham, Gallagher & Frith, 1997). But if one took an individual adult with dyslexia, he or she might have literacy skills in the average range, while an individual adult who does not have dyslexia might have below-average literacy skills. A screening test has to be capable of identifying which individuals do, and do not, have dyslexia to a reasonable degree of accuracy, and for that task measures other than (or additional to) literacy tests are required. However, there is extensive evidence that, in adulthood, persons with dyslexia (even dyslexic university students) still exhibit limitations in tasks involving phonological processing, lexical access (i.e. tasks involving lexical decision or rapid recall of lexical information) and working memory (see Beaton et al., 1997). Arguably, therefore, these cognitive domains offer better prospects for effective dyslexia screening in adults. Snowling et al. (1997) compared dyslexic and non-dyslexic university students on a range of cognitive and literacy tasks, and found that the strongest and most consistent differences between the groups were on tests of phonological processing, including nonword reading (effect size 3.64), spoonerism accuracy (effect size 1.70), spoonerism speed (effect size 3.91), phonemic fluency (effect size 1.96) and phoneme deletion (effect size 1.90). Hanley (1997) also compared dyslexic university students with their non-dyslexic peers, finding that the former were significantly poorer at lexical decision tasks, nonword reading and nonword spelling, and working memory (digit span). Gottardo, Siegel and Stanovich (1997) compared adults (mean age 33 years) who had average reading ability with those having poor reading ability, using standardised measures of literacy and intelligence, together with several experimental measures of phonological processing (pseudoword repetition, phoneme deletion, Pig Latin) and working memory (digit span). These authors found that the phonological processing and working memory measures significantly predicted word and nonword reading ability.

The term ‘screening’ may be used for any routine procedure that facilitates rapid and easy sorting into categories. However, in fields other than medicine there is usually the acceptance that some misclassification is inevitable: a certain level of error is traded for increased speed and practicality (Singleton, 1997a; Streiner, 2003). In education the purpose of screening is usually to identify children or adults who have special educational needs (e.g. because of dyslexia). The accuracy of any screening system is indicated by the frequency of misclassifications that it makes. False positives and false negatives are the two types of screening misclassifications. A false positive is a ‘false alarm’, that is, an instance in which a person has been classified as ‘at risk’ (in this case, possibly dyslexic) when, in fact, this is not the case. A false negative is a ‘miss’, that is, an instance in which a person has been classified as ‘not at risk’ (in this case, not dyslexic) when, in fact, the opposite is the case (they are dyslexic). The value of any screening test depends on having low frequencies of false positives and false negatives. False negatives are usually regarded as being more serious than false positives because although a high frequency of false positives may result in the diverting of resources to dealing with cases that do not actually require attention (i.e. a resource cost), a high frequency of false negatives results in critical cases be individuals because another important exhaustiveness of the conditions of developmental co-occurrence and (b) the polygenic family relationship dyslexic nature on the family of Mutet & Carroll, 2001, into a ‘borderline’ (i.e. the cases) that are not meeting those criteria, is an interesting distinction that they will later in this article.

In developing a taps (i.e. lowest possible level of the measurement administrative effort true false negatives are a problem; Singleton, 1983). Metrics of sensitivity, specificity, percentage of false positives and false negatives are satisfactory. However, judgement has to be made in some cases if a test is not picked up in a particular population in false positive cases.

The advantages of the method are clear but the method is also expensive and difficult to implement. Taps (i.e. screening) is a relatively straightforward approach, allowing for the assessment of a population without the need for a full assessment. However, the advantages of this method are apparent only if the test is applied in a rigorous and systematic manner. The method is also relatively easy to implement, allowing for the assessment of a population without the need for a full assessment. However, the advantages of this method are apparent only if the test is applied in a rigorous and systematic manner.
in critical cases being overlooked, which may have deleterious consequences for those individuals because they are likely to be denied the treatment or attention they need. Another important issue is that categorisation usually presupposes exclusive and exhaustive categories to which all individuals may be allocated; in the case of dyslexia this is not possible because of (a) the high comorbidity between dyslexic symptoms and symptoms of other developmental conditions, such as specific language impairment, developmental coordination disorder and attention deficit disorders (Snowling, 2008), and (b) the polygenetic basis of dyslexia, which results in individuals with a biological family relationship to a dyslexic person being more likely to display mild symptoms of a dyslexic nature even though they may not satisfy full criteria for dyslexia (Snowling, Muter & Carroll, 2007). Consequently, there are likely to be some individuals who fall into a ‘borderline’ or ‘grey area’ between clearly meeting criteria for dyslexia and clearly not meeting those criteria. In deciding what should be the best response to the needs of such cases, it is arguably better to look more closely at them as individuals rather than insisting that they be squeezed into one category or the other. This issue will be revisited later in this article.

In developing screening instruments, attempting to reduce misclassifications to the lowest possible levels almost always results in increasing the complexity and sophistication of the measurement, which will tend to reduce the practicality of screening and increase administrative effort and costs. In general, levels <25% for false positives and <25% for false negatives are advocated for effective educational screening (see Jansky, 1977; Potton, 1983; Singleton, 1997a). Screening tests may also be evaluated using the alternative metrics of sensitivity (the percentage of true cases correctly identified) and specificity (the percentage of false cases correctly identified). Glascoe and Byrne (1993) argue that sensitivity rates should be at least 80% and specificity at least 90% to regard the test as satisfactory. However, Grimes and Shultz (2002) point out that in medical screening a judgement has to be made regarding the relative costs of false negatives and false positives, and in some cases the costs of false negatives (e.g. early death because a critical condition was not picked up by screening) may be sufficiently high to accept a corresponding increase in false positives and the resource costs that this implies.

The advantages of computerised assessment in educational settings has been explored by Singleton (see Singleton, 1997b, 2001). Computers provide more precise measurement, especially when complex cognitive skills are being assessed. Tests are administered in an entirely standardised manner for all persons taking the test, which enhances reliability of measurement. Timings and presentation speeds can be controlled precisely. The subjective judgement of the test administrator does not affect the test outcome as it can in conventional tests. Results are available immediately, and when assessing older children and adults, assessment can be largely self-administered; both of these factors help to reduce administrative load and avoid delays. Adults, as well as children, often prefer computerised assessment to conventional assessment (Horne, 2002; Singleton, 2001). Adults often find conventional assessment by another person stressful and anxiety provoking, particularly if they have reading or writing problems or experienced difficulties at school, and particularly when the assessor is perceived as ‘teacher’ or some equivalent professional. By contrast, they are generally more relaxed and confident when taking computerised tests, and less worried about making errors (see BDA, 2005).

Time efficiencies are considerable when computer-based assessment is adaptive, that is, where the difficulty of items selected from a test bank is varied in response to the person’s progress on the test. The term ‘adaptive testing’ refers to any technique that
modifies the nature of the test in response to the performance of the test taker. This can be achieved in a variety of ways, although typically it is used in connection with tests that are developed by means of Item Response Theory (Hambleton & Swaminathan, 1985). Conventional tests are static instruments, fixed in their item content, item order and duration; by contrast, computer-based assessment can be dynamic. Since the computer can score performance at the same time as item presentation, it can modify the test accordingly, tailoring it more precisely to the capabilities of the person taking the test. In conventional tests, for some part of the time, the person’s abilities are not being assessed with any great precision because the items are either too difficult or too easy (which can easily lead to frustration and/or boredom). In a computerised adaptive test, however, because the program contains information about the difficulty level of every item in the item bank (based on pass rates in the standardisation population) the individual taking the test can be moved swiftly to that zone of the test that will most efficiently discriminate his or her capabilities. This makes the whole process speedier, more reliable, more efficient and often more acceptable to the person being tested. It has been shown that an adaptive, computer-based assessment can take only a quarter of the time taken to administer an equivalent conventional test (Olsen, 1990).

The present investigation addressed the effectiveness of computerised screening for dyslexia in adults, using three specially designed tests, which are not conventional tests of reading and spelling but which depend heavily on phonological processing, lexical access and working memory.

**Method**

**Participants**

The participants in the study totalled 139 adults varying in age from 16 to 74 years (mean 32.83 years; SD 14.37 years). They were drawn from three types of educational institution: 42 were from two universities, 48 were from three colleges of further education and 49 were from three adult learning centres. This varied mix of participants was chosen in order to assess the validity of the dyslexia screening instruments across a broad range of abilities, educational levels and socioeconomic backgrounds. Seventy of the participants had received a prior diagnosis of dyslexia from an educational psychologist adhering to the guidelines drawn up by the Department for Education and Skills Working Group on Specific Learning Difficulties in Higher Education (Patoss, 2005). These comprised the dyslexic group. The following list of diagnostic criteria has been extracted from the DfES guidelines, which are now standard in higher education in England and are applied when considering applications for Disabled Students’ Allowances (DSAs) from students with dyslexia.

1. A history of difficulty with the acquisition of literacy skills.
2. Persisting literacy difficulty (e.g. slow, inaccurate reading, poor writing and spelling).
3. Evidence of an underlying cognitive deficit in phonological processing speed, phonological awareness, visual and/or auditory working memory.
4. Exclusion of other factors (e.g. sensory impairment, English as a second or additional language, educational experience and opportunities for learning).
5. Underlying ability/achievement differentials. Although a discrepancy between underlying ability and attainment in literacy skills is not regarded as a diagnostic...
criterion in these guidelines, where such discrepancies do exist, they provide further supporting evidence (Frederickson & Reason, 1995; Howe, 1997; Miles, 1996; Siegel, 1999; Stanovich & Stanovich, 1997).

The remaining 69 participants comprised the non-dyslexic group; these participants were randomly recruited from the same institutions as the dyslexic participants and were checked by educational psychologists to ensure that they did not meet the criteria for dyslexia set out above.

Materials

Word Recognition test. Word Recognition is a test of lexical access involving speeded discrimination of real words from pseudowords. In each item, five words appear on the computer screen in random positions. Only one of these five is a real word; the other four are pseudowords, which in some cases may be pseudo-homophones (i.e. letter strings that if pronounced sound like a real word). An example is: toad (target word), tode, trood, todday, toap (distractors). Using the computer mouse, the person taking the test has to click on the real word as quickly as possible. If no response is made within 30 seconds, the program automatically moves on to the next item. The test begins with two practice items, which are accompanied by spoken instructions. The practice items have to be passed successfully before the test proper commences. There were 120 items in the full (as opposed to adaptive) version of this test.

Word Construction test. Word Construction is a test of lexical encoding of pseudowords from syllables. Using digitally pre-recorded speech, the computer speaks a three-syllable pseudoword (e.g. 'subromast') and the person taking the test has to click on the syllables that make up this pseudoword in the correct order, selecting them from a bank of nine different syllables displayed on the screen in a 3 x 3 grid. As each syllable is clicked on, it appears in a template at the top of the screen. If a mistake is made, the person can click on the template to undo the error, so permitting another choice to be made. When the person is satisfied with their choice of syllables, they click on an arrow at the bottom right-hand corner of the screen, and the program proceeds to the next item. The test begins with two practice items, which are accompanied by spoken instructions. The practice items have to be passed successfully before the test proper commences. The task has to be completed as swiftly as possible. If no response is made within 30 seconds, the program automatically moves on to the next item. There were 50 items in the full version of this test. This task, which resembles a type of nonword spelling test, was designed primarily to tap phonological coding ability (i.e. mapping sounds on to letter sequences) but it also draws upon working memory in order to retain activation of the sound of the target word while checking this against the options that appear on the screen.

Working Memory test. Working Memory is a test of backwards digit span. Using digitally pre-recorded speech, a sequence of digits is spoken by the computer, and the person taking the test is required immediately to recall these in reverse order, entering the digits using the computer keyboard. The test begins with two practice items accompanied by spoken instructions. The practice items have to be passed successfully before the test proper commences. The test then proceeds as in a conventionally delivered digit span task, commencing with items of two digits in sequence, followed by items of three digits, and so on up to nine digits in sequence. At each level two items are presented. If correct
responses (i.e., all digits correct and in the correct sequence) are made to one or both of these items then the program proceeds to the next level, in which there will be one more digit than the previous level. If both items are incorrect, then the program terminates (except if the error occurs at the first level). The program allows a limited time for each item; this is a function of the number of digits in the item and varies from 14 seconds up to a maximum of 28 seconds.

Wide Range Achievement Test (WRAT-3). The Reading and Spelling tests from WRAT-3 (Jastak & Wilkinson, 1993) were administered to check differences in conventional literacy skills between the groups.

Results

The means of the WRAT-3 Reading and Spelling tests for the dyslexic and non-dyslexic groups were significantly different, Reading: dyslexic mean 40.31 (SD 6.43), non-dyslexic mean 51.26 (SD 4.80), F(1, 129) = 129.41, p < .001; Spelling: dyslexic mean 31.12 (SD 6.29), non-dyslexic mean 46.38 (SD 4.43), F(1, 129) = 306.65, p < .001. Note that of the 139 participants in the study, WRAT-3 data were available for only 130 participants.

Table 1 shows the raw scores for the three computerised tests, categorised by group (dyslexic vs. non-dyslexic) and type of institution (university vs. further education college vs. adult learning centre). For all three tests, the difference between the dyslexic and the non-dyslexic groups was highly significant. Word Recognition: $F(1, 138) = 157.15; p < .001$; effect size (Cohen's $d$) = 2.11; Word Construction: $F(1, 138) = 111.37; p < .001; d = 1.77$; Working Memory: $F(1, 133) = 19.01; p < .001; d = 0.78$. (Note that in calculating Cohen’s $d$, the divisor was the pooled estimate of the population standard deviation; see Cohen, 1988, p. 44.) There was a significant effect of type of institution on Word Construction, $F(2, 138) = 6.58; p = .002$; post hoc analysis (Tukey) indicated that the means of the university and further education groups were both significantly higher than that of the adult learning group ($p < .05$), but there was no significant difference between the means of the university and further education groups. There was no significant effect of type of institution on Word Recognition, $F(2, 138) = 2.73; p = .069$

<table>
<thead>
<tr>
<th>Group</th>
<th>Universities</th>
<th>Further education colleges</th>
<th>Basic skills centres</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dyslexic</td>
<td>21</td>
<td>20</td>
<td>29</td>
<td>70</td>
</tr>
<tr>
<td>Non-dyslexic</td>
<td>21</td>
<td>28</td>
<td>20</td>
<td>69</td>
</tr>
<tr>
<td>All</td>
<td>42</td>
<td>48</td>
<td>49</td>
<td>139</td>
</tr>
<tr>
<td>Word</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Recognition</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dyslexic</td>
<td>83.52 (18.25)</td>
<td>71.75 (21.83)</td>
<td>68.55 (16.76)</td>
<td>73.96 (19.58)</td>
</tr>
<tr>
<td>Non-dyslexic</td>
<td>107.71 (9.17)</td>
<td>105.39 (13.05)</td>
<td>110.20 (9.44)</td>
<td>107.49 (11.01)</td>
</tr>
<tr>
<td>All</td>
<td>95.62 (18.80)</td>
<td>91.38 (23.90)</td>
<td>85.55 (25.04)</td>
<td>90.60 (23.12)</td>
</tr>
<tr>
<td>Word</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Construction</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dyslexic</td>
<td>31.60 (10.42)</td>
<td>22.90 (10.43)</td>
<td>18.85 (8.99)</td>
<td>23.71 (10.98)</td>
</tr>
<tr>
<td>Non-dyslexic</td>
<td>42.14 (6.05)</td>
<td>41.04 (6.59)</td>
<td>40.15 (12.43)</td>
<td>41.12 (8.47)</td>
</tr>
<tr>
<td>All</td>
<td>36.76 (10.06)</td>
<td>33.27 (12.33)</td>
<td>27.91 (14.92)</td>
<td>32.47 (13.10)</td>
</tr>
<tr>
<td>Working</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Memory</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dyslexic</td>
<td>4.25 (1.68)</td>
<td>3.55 (2.59)</td>
<td>2.30 (1.92)</td>
<td>3.25 (2.21)</td>
</tr>
<tr>
<td>Non-dyslexic</td>
<td>5.57 (3.39)</td>
<td>4.81 (2.42)</td>
<td>5.21 (2.18)</td>
<td>5.16 (2.68)</td>
</tr>
<tr>
<td>All</td>
<td>4.93 (2.74)</td>
<td>4.28 (2.54)</td>
<td>3.50 (2.47)</td>
<td>4.21 (2.63)</td>
</tr>
</tbody>
</table>

© United Kingdom Literacy Association 2009
or Working Memory, \( F(2, 133) = 2.55; p = .086 \). There were significant interactions between group and type of institution for both Word Recognition, \( F(2, 138) = 3.58; p = .031 \) and Word Construction, \( F(2, 138) = 3.46; p = .034 \) but not for Working Memory, \( F(2, 133) = 1.67; p = .19 \). Post hoc analysis indicated that the dyslexics in both the further education and adult learning groups were relatively more impaired on both Word Recognition and Word Construction than the non-dyslexic groups in these institutions, compared with the degree of difference in impairment on these measures for the two university groups (\( p < .05 \)).

There were 58 males and 81 females in the sample. There were no significant main effects of gender, except in the Word Recognition test, in which the females (mean 94.45; \( SD = 20.34 \)) were found to score significantly higher than the males (mean 83.81; \( SD = 22.70 \)), \( F(1, 138) = 7.55; p = .007 \). There were no significant interactions between gender and any other variables.

**Development of the adaptive versions of the Word Recognition and Word Construction tests**

Adaptive versions of the Word Recognition and Word Construction tests were created by calculating difficulty levels for every item and selecting a reduced set of items using the Computerised Adaptive Sequential Testing method (Drasgow & Olson-Buchanan, 1999). In this approach, which is based on Item Response Theory (Embretson & Reise, 2000), blocks of items of known difficulty are administered in an adaptive sequence, and the results were converted to a scaled score from 1 to 9 where higher scores signify increased risk of dyslexia. This resulted in adaptive versions of these two tests, in which the person attempts between 10 and 40 items, significantly reducing administration time.

For the Word Recognition test, a correlation of -.95 was obtained between the full version and the adaptive version, while for the Word Construction test, the correlation was -.96. Both these correlations were statistically significant (\( p < .001 \)). Data from the Working Memory test were also converted to the same scaled scores to provide comparability of scores with the other two tests, and a correlation of -.85 was obtained between the original scores version and the scaled scores, which is also statistically significant (\( p < .001 \)). To conserve space, descriptive statistics of these data are not shown, but statistical analysis showed that the conversion to the adaptive versions using scaled scores did not alter the significance level of any of the differences between the means or the interactions in the data set, indicating comparability with the original forms of these tests.

The discriminant power of the adaptive versions of the three tests may be increased by amalgamating their scaled scores, to give a combined score ranging from 3 to 27. The data for the combined scores are shown in Table 2. Statistical analysis of these data revealed significant main effects of group, \( F(1, 133) = 155.79; p < .001; d = 2.07 \), and institution type, \( F(2, 133) = 7.01; p = .001 \) and a significant interaction of Group × Institution type, \( F(2, 138) = 5.00; p = .008 \). Post hoc analysis (Tukey) indicated that the means of the university and further education groups were both significantly lower than that of the adult learning group (\( p < .001 \)).

In order to further establish the validity of the tests as a dyslexia screening device, a discriminant function analysis was carried out on the combined scaled scores. It should be noted that five participants were absent for the Working Memory test and hence the overall number of participants in this analysis was 134. Analysis revealed that the
Table 2. Means (and standard deviations) of the composite scaled score derived from the three tests, categorised by group and type of institution.

<table>
<thead>
<tr>
<th>Group</th>
<th>Universities</th>
<th>Further education colleges</th>
<th>Basic skills centres</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dyslexic</td>
<td>15.15 (4.25)</td>
<td>18.20 (5.27)</td>
<td>21.30 (3.82)</td>
<td>18.54 (5.06)</td>
</tr>
<tr>
<td>Non-dyslexic</td>
<td>8.19 (3.74)</td>
<td>9.85 (3.47)</td>
<td>8.79 (5.14)</td>
<td>9.03 (4.09)</td>
</tr>
<tr>
<td>All</td>
<td>11.59 (5.29)</td>
<td>13.40 (5.97)</td>
<td>16.13 (7.60)</td>
<td>13.78 (6.61)</td>
</tr>
</tbody>
</table>

Table 3. Discriminant function analysis of results for the composite scaled score on the computerised screening test.

<table>
<thead>
<tr>
<th>Score range</th>
<th>Category</th>
<th>Dyslexic</th>
<th>Non-dyslexic</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>3–12</td>
<td>Low risk</td>
<td>11</td>
<td>59</td>
<td>70</td>
</tr>
<tr>
<td>14–27</td>
<td>High risk</td>
<td>56</td>
<td>8</td>
<td>64</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>67</td>
<td>67</td>
<td>134</td>
</tr>
</tbody>
</table>

Table 4. Discriminant function analysis of results for the composite scaled score on the computerised screening test with the inclusion of a borderline category.

<table>
<thead>
<tr>
<th>Score range</th>
<th>Category</th>
<th>Dyslexic</th>
<th>Non-dyslexic</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>3–11</td>
<td>Low risk</td>
<td>6</td>
<td>48</td>
<td>54</td>
</tr>
<tr>
<td>12–14</td>
<td>Borderline</td>
<td>7</td>
<td>14</td>
<td>21</td>
</tr>
<tr>
<td>15–27</td>
<td>High risk</td>
<td>54</td>
<td>5</td>
<td>59</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>67</td>
<td>67</td>
<td>134</td>
</tr>
</tbody>
</table>

Table 5. Discriminant function analyses for Wide Range Achievement Test (WRAT-3) Reading and Spelling scores (N = 130).

<table>
<thead>
<tr>
<th>Test</th>
<th>Category</th>
<th>Dyslexic</th>
<th>Non-dyslexic</th>
</tr>
</thead>
<tbody>
<tr>
<td>WRAT-3 Reading</td>
<td>Below average</td>
<td>9</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Average or above</td>
<td>56</td>
<td>65</td>
</tr>
<tr>
<td>WRAT-3 Spelling</td>
<td>Below average</td>
<td>16</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Average or above</td>
<td>49</td>
<td>65</td>
</tr>
</tbody>
</table>

The optimum cut-off was obtained by regarding scores in the range 3–13 as indicative of low risk of dyslexia and scores in the range 14–27 as indicative of high risk of dyslexia (see Table 3). The sensitivity measure for this cut-off was 83.6%, and specificity 88.1%. However, it was found that sensitivity and specificity were improved by the inclusion of a 'borderline' category for scores in the range 12–14. The results are shown in Table 4. Analysis of the data without the borderline category gave a sensitivity rate of 90.6%, and a specificity rate of 90.0%.

To compare the predictive accuracy of the computerised tests with that of the WRAT-3 tests, discriminant function analyses were also carried out on the data obtained from the WRAT-3 Reading and Spelling tests. For both of these tests, the cut-off used was one standard deviation below the mean, dividing the sample into two categories: (i) below average and (ii) average or above. The results of these analyses are provided in Table 5, which shows that specificity for both tests is 100%. However, sensitivity rates are very low (13.9% for Reading and 24.6% for Spelling). In other words, although WRAT-3
Reading and Spelling tests do not misclassify any non-dyslexics as having dyslexia, they misclassify most dyslexics as non-dyslexics (false negatives).

**Discussion**

The purpose of this study was to investigate the validity of a computer-based approach to screening for dyslexia in adults. The sample was deliberately chosen to give a wide range of ages, abilities and educational and socioeconomic backgrounds. The dyslexic and non-dyslexic groups were found to differ substantially in reading and spelling abilities. It was found that the mean scores of the dyslexic and non-dyslexic groups were significantly different on all three of the computerised tests, with an average effect size of 1.55, which is regarded as large (Cohen, 1988). However, the numbers of items in the Word Recognition ($n = 120$) and Word Construction ($n = 50$) tests would preclude their use in a screening device as the tests would take far too long. Consequently, adaptive versions of these two tests were created in order to reduce administration time; in fact, the average administration time was cut to under 5 minutes per test and the duration of the whole screening to about 15 minutes. It was found that the mean combined scaled scores of the adaptive versions for the dyslexic and non-dyslexic groups were significantly different, with an increased effect size of 2.07.

Discriminant function analysis indicated very satisfactory rates of 90% for both sensitivity and specificity. Altering cut-offs would result in an increase in one of these measures but a corresponding decrease in the other: that is, reducing the cut-off would decrease false negatives but increase false positives, and increasing the cut-off would decrease false positives but increase false negatives (Bland, 1999; Singleton, 1997a). Unless there was a very good reason to do this (e.g. where identification was so critical that even tiny numbers of false negatives would be unacceptable), the 90% rates should be regarded as optimal for this test. These sensitivity and specificity rates may be compared with those obtained for the WRAT-3 Reading and Spelling tests, for which specificity was faultless (100%) but sensitivity dangerously deficient (13.9% for Reading and 24.6% for Spelling). In other words, use of WRAT-3 Reading and Spelling tests as screening instruments (without any other measures) would miss the vast majority of the dyslexics, and consequently would be an unsatisfactory basis for screening for dyslexia in adults.

The 90% sensitivity and specificity rates for the computer screener were obtained by excluding a ‘borderline’ category; without this category sensitivity and specificity rates were still perfectly acceptable, being within the limit advocated by Glascoe and Byrne (1993) for sensitivity and not significantly different from the limit advocated by these authors for specificity. We suggest that a borderline categorisation should be regarded as a cue for the administrator to examine that person’s results more closely before deciding whether the risk of dyslexia is high or low and whether there are other factors that need to be taken into consideration. In the present study, the borderline category accounted for 21 participants (16% of the sample) and comprised one-third dyslexics and two-thirds non-dyslexics. Further analysis of this group indicated that on Word Recognition and Word Construction there were no significant differences between the borderline category and the low-risk category, but there was a significant difference ($p < .001$) between these categories on Working Memory. Inspection of the seven dyslexic cases in this category revealed a mean scaled score for Working Memory of 7.71 (SD 1.6), which is very high.

© United Kingdom Literacy Association 2009
Clearly, this is a subgroup of individuals who display very poor working memory but do not appear to have significant difficulties with the lexical and phonological components of the test (i.e. Word Recognition and Word Construction). This information is helpful to assessors and arguably justifies the adoption of this category.

It should not be assumed that computerised screening is necessarily more effective than conventional (i.e. non-computerised) screening. In particular, computerised tests are at present unable to cope reliably with oral responses, which some critics would regard as a serious deficiency when assessing abilities such as verbal memory and phonological awareness. An alternative view is that this inherent limitation of computerised tests simply underlines the requirement for high-quality validation of such tests. The best-known conventional screening device for dyslexia in adults is the Dyslexia Adult Screening Test (DAST) (Nicolson & Fawcett, 1998), which comprises 11 subtests that are delivered individually by an assessor. The range of DAST subtests includes measures of phonological processing and auditory memory, but also contains an assessment of postural stability, which reflects the authors’ particular theoretical perspective on dyslexia (see Nicolson, Fawcett & Dean, 2001). The test takes about 30 minutes to administer and performance on each subtest is transformed into an index score, with the sum of the indices being divided by the number of subtests to arrive at a dyslexia-at-risk quotient. Harrison and Nichols (2005), who have criticised the DAST on grounds of inadequate validation and standardisation, carried out an evaluation study of DAST that compared 117 dyslexic students with 121 non-dyslexic students, finding a sensitivity of 74% and a specificity of 84%, both of which are below the limits for practicable screening advocated by Glascoe and Byrne (1993). These authors also note that their findings ‘suggest that individuals with lower general intelligence are more likely to be identified as dyslexic with the DAST regardless of their level of reading proficiency’ (p. 431). A substantial number of students who rated their reading skills as good were classified by DAST as having high risk of dyslexia, although this finding should be interpreted with caution as it was based on subjective report. As pointed out earlier, there is ample evidence of students with superficially good literacy skills despite having dyslexia (see Beaton et al., 1997). However, it should be noted that the sensitivity and specificity rates for DAST obtained by Harrison and Nichols (2005) are inferior to those obtained for the computer-based screener, regardless of the adoption of a borderline category in the latter. This suggests that in terms of predictive validity a computer-based screener can be at least as good as, if not rather better than, a conventional screener, while also offering many other advantages that have already been outlined.

There are four principal situations in which computerised dyslexia screening for adults would be useful. The first is in higher education where, in the United Kingdom, financial support is available in the form of DSAs for dyslexic students to purchase assistive technology and study skills tuition to support their learning. Such students are also entitled to additional time in written examinations. A national survey of 195 institutions of higher education in the United Kingdom revealed that 43% of students recognised as having dyslexia were not formally identified until after admission to university (Singleton, 1999). However, DSAs and examination concessions are contingent upon having a report from an educational psychologist or other accredited assessor confirming that the student has dyslexia, and such assessments are expensive. A screening test enables institutions to give more informed advice to students who are considering formal assessment and helps to ensure that costs, overall, are minimised. The second situation is in further education, where dyslexia is recognised as a special educational need and additional teaching to target these courses from the identifying the 

The third situation, where approximately (2001) studied if the literacy over a period of time, 48% made significant declines in performance, which suggests that many students may have dyslexia. In these cases, lessons received ‘much of the input of tutors’ (p. 9). These students’ difficulties were found from these findings, which screened for dyslexia needs of dyslexic learners (e.g. Lo 1994; Venezy & the students found no evidence that given to other groups (2005), after item analysis, led to which dyslexia failure of Rice an

The fourth situation involves the impact on job progression (Fitzpatrick et al., 2006). The rights of dyslexic students to equality of opportunity, which together with the need for employment, education and training, puts them at a disadvantage. Many students are required to take courses that include making effective use of technology, the development of teaching materials in alternative formats, and strategies to enable teachers to overcome the barriers that students face in order to become more effective learners (Singleton, 1999). However, there are
additional teaching resources are funded to support identified dyslexic students. In order to target these resources appropriately to enable dyslexic students to cope with their courses from the outset, administrators require speedy and cost-effective methods of identifying the dyslexic students from amongst considerable numbers of new entrants.

The third situation in which dyslexia screening would be beneficial is in adult learning, where approximately half of students tend to have low standards of literacy. Brooks et al. (2001) studied the progress of over 1,200 adult learners with low or very low standards of literacy over a period of 2-6 months of basic literacy tuition, and found that although 48% made significant improvements (i.e. at least a quarter of a standard deviation), 30% declined in performance by a significant amount. A survey by Besser et al. (2004) suggests that most adult learners with literacy difficulties have poor phonological awareness, and in a detailed study of 53 adult learners these authors found that 16 had been diagnosed as having dyslexia and a further seven were considered by their tutors to have dyslexia (i.e. 43% estimated prevalence of dyslexia in that sample). Observing the lessons received by the adult students in their study, Besser et al. (2004) reported that 'much of the phonics teaching was done on the spur of the moment and there were instances of tutors' inadequate grasp of phonics leading to inaccurate phonics teaching' (p. 9). These authors concluded that teaching needed to be more intensive and focused on students' difficulties with word identification and decoding. It is tempting to conclude from these findings that adult literacy tuition might be more successful if such adults were screened for dyslexia and teaching methods were more closely related to the learning needs of dyslexic adults, especially as there is a literature documenting the benefits of this approach (e.g. Lee, 2000; Newman, 1994; Rack, 1997; Venezky, Bristow & Sabatini, 1994; Venezky & Sabatini, 2002). In a contrary view, a report by Rice and Brooks (2004) could find no evidence that adults with dyslexia require a different tuition approach from that given to other adults who have literacy difficulties. However, Siegel and Smythe (2005), after itemising numerous flaws in the Rice and Brooks report, observed that 'An in-depth analysis of the efficacy of intervention systems would have revealed the extent to which dyslexic adults respond to different environments' (p. 74), concluding that the failure of Rice and Brooks to do this undermines the validity of their judgement on the matter.

The fourth situation is in employment, where dyslexia can have a major negative impact on job opportunities, workplace efficiency, personal satisfaction and career progression ( Fitzgibbon & O'Connor, 2002; Morgan & Klein, 2000; Reid & Kirk, 2001). The rights of dyslexic people in the United Kingdom are protected by the Disability Discrimination Act 1975, and the Special Educational Needs and Disability Act 2001, which together make discrimination against people with disabilities unlawful in employment, education, training and the provision of services. If a person's dyslexia puts them at a substantial disadvantage compared with other citizens, responsible bodies are required to take reasonable steps to prevent that disadvantage. Such steps might include making changes to employment policies and workplace practices, changes to course requirements or work placements, the provision of support workers or assistive technology, the delivery of training in alternative ways and the provision of learning materials in alternative formats. Bartlett and Moody (2000) have detailed numerous interventions and strategies that can be applied in the workplace to help dyslexic employees to overcome the difficulties and hindrances created by dyslexia, enabling them to become more efficient, reducing their stress and increasing their job satisfaction. However, there are many dyslexic adults who have yet to be identified as having dyslexia.
and although they are probably aware of their limitations in the workplace or in the classroom they may well not understand the reasons for these difficulties and may not know how or where to seek help. Screening for dyslexia in the workplace provides a simple and cost-effective way to identify employees who have dyslexia so that appropriate steps can be taken by employers to support them, thus meeting their statutory obligations.

The benefits of computerised screening compared with conventional screening have already been outlined (see introduction), and this study has demonstrated that a computer-based approach using tests that tap phonological skills, lexical access and working memory offers a valid and practical solution to screening adults for dyslexia. The fact that computerised screening tests can be delivered speedily and, if necessary, simultaneously (via a computer network) to many hundreds of adults, without the requirement of trained administrators, supervision or expert interpretation, has considerable advantages. There are uses for this approach in education and employment, with the potential that large numbers of adults with hitherto unidentified dyslexia can be identified and hence begin to understand their condition and seek support that will enable them to overcome their difficulties.

References


© United Kingdom Literacy Association 2009


Received 9 May 2008; revised version received 4 September 2008.

Address for correspondence: Dr Chris Singleton, Department of Psychology, University of Hull, Hull HU6 7RX, UK. E-mail: c.singleton@hull.ac.uk

© United Kingdom Literacy Association 2009