Identifying Dyslexic Students: Designing a computer-based dyslexia screening test-prototype for higher education

Alexia Casale

University of Essex

Abstract

This article discusses the challenges involved in developing a computer-based dyslexia screening test for widespread HE usage. These challenges are illustrated through preliminary results from pilot-testing a prototype instrument designed to fill the gap in provision. A number of promising avenues, which mainly relate to two key issues (complexity and processing/reading speed), are outlined in this exploratory research; however, as with all exploratory studies, far larger sample sizes would be required in order to assess reliability and generalisability of findings. First, complex tasks, tapping multiple deficits (thereby circumventing dyslexic students' coping strategies), tend to be more sensitive than tasks which involve a single area of deficit. One of the reasons that existing tests for use with HE students do not seem to hold up under scrutiny (see Casale 2009) may be that the tasks involved tap only a single area of *deficit* when they tap multiple cognitive domains at all. Second, data analysis indicates that time taken to complete cognitive tasks (i.e. processing-speed), rather than accuracy, is a more effective measure for identifying dyslexic HE students. These findings have far-reaching implications for the types of 'reasonable adjustments' that the Disability Discrimination Act requires that dyslexic students be afforded in order to allow them to study and perform without disadvantage compared to non-dyslexics. However, this deficit is not significant merely in relation to extra time in exams: it potentially has much more serious implications in relation to dyslexic students' ability to cope with the volume of reading, turn-over of essays and other work at university.

Keywords: Dyslexia, computer-based test, higher education.

Identifying Dyslexic Students: Designing a computer-based dyslexia screening test-prototype for higher education

Introduction[1]

Developing a computer-based dyslexia screening test-prototype specifically targeted at university students is a challenging and complex endeavour, not least due to the lack of an accepted definition of dyslexia and the general difficulties of creating effective cognitive tests for high-performing adults. This exploratory research illustrates the importance of operationalising competing definitions and theories of dyslexia, in constructing and piloting a new instrument; by operationalising a range of definitions and theories, it is possible to explore and, thus, determine the types of tasks that reliably identify dyslexics in the HE population. Thus, a number of promising avenues for future work are identified and discussed.

In attempting to develop a prototype test to screen students for dyslexia, the working definition employed adopts and adds to the World Federation of Neurology definition: *'These fundamental cognitive difficulties, manifested in difficulties with reading, writing and related skills, may include phonological, visual, Short Term Memory (STM) and automatisation deficits. Other soft signs may include laterality discrimination and sequencing difficulties. Dyslexia is a heterogeneous disorder with individuals evidencing unique profiles of strengths and weaknesses'*[5].

Test Design

A range of theories were operationalised in designing the screening test prototype in order to avoid *a priori* theoretical assumptions that might bias the test towards identifying one particular pattern of difficulties. If dyslexia is best characterised as consisting of a range of fundamental deficits, operationalising tasks which focus on one (or even several) deficit theories means ignoring other deficits. Adopting a Grounded Theory Methodology (Bryman & Burgess, 1993; Glaser & Strauss, 1967) allows theory to emerge from data; by operationalising various deficit theories in a wide range of tasks, it is possible to allow the data to show which tasks and task-measures (e.g. accuracy, time

[1] For a discussion of the need for a new computer-based dyslexia screening test in HE, an analysis of existing screening tests (both computer- and paper-based), and an exploration of the problems associated with the many different and, indeed, competing definitions (and associated theories) of dyslexia, see Casale, Alexia. (2009). Identifying Dyslexic Students: The need for computer-based dyslexia screening in higher education. *Estro*, 1(1), 119-143.

DOI: https://doi.org/10.5526/esj147

taken, etc) are most effective in identifying dyslexic students. Research hypotheses can then be constructed, based on the adoption of the deficit theory/theories that the data have shown to most accurately characterise the phenomenon. These hypotheses can be tested by collecting new data, then the theoretical position of the research, and its hypotheses, can be refined, leading to a further stage of data collection and analysis. A key aim in the development of the test-prototype was to operationalise the range of deficit theories as comprehensibly[6] as possible, given practical limitations.

Drawing on the literature on IQ, and ability tests relevant to dyslexia (e.g. Weschler Adult IQ test [Weschler, 1955], Raven's matrices [Raven et al., 1962], and Wide Range Ability Test [Wilkinson, 1993]), a bank of 34 tasks (including 77 subtasks) was compiled. To organise this large number of tasks, six (non-analytical, organisational) categories were identified: (1) Short Term Memory (e.g. digit span), (2) sequencing (e.g. anagrams), (3) verbal reasoning (e.g. word-based arithmetic problems, identifying semantic similarities) and non-verbal reasoning (e.g. matrices), (4) phonological awareness and written language skills (e.g. non-word identification), (5) visuo-spatial abilities (e.g. block completion), and (6) self-report questionnaire. As some tasks were relevant to more than one dyslexia deficit theory, categorisation was determined by the primary cognitive domain tapped by the tasks.

Category 1 tasks mainly operationalise STM deficit theories, while Category 4 tasks are primarily relevant to phonological deficit theories.

Category 2 (sequencing) tasks relate primarily to automatisation theories. The anagrams/acronyms, digit-coding (adapted from the Weschler Adult IQ test [Weschler, 1955]) and sequencing tasks require participants to recognise and relate words and letters, symbols and numbers, and sequences of letters/numbers, respectively. The sequences task is one of several which has two components (i.e. sequence pattern recognition and sequence construction); such tasks were included to investigate whether dyslexic students compensate effectively for a single dyslexic deficit, but fail to compensate when two types of deficit are tapped simultaneously. Traditionally, automatisation deficit theories have been operationalised by naming speed tasks (Fawcett & Nicolson, 1994; Wolf & Bowers, 1999), which require verbal answers: a response modality unsuitable for computer-based tests due to current technical limitations. Other tasks designed to provide insight into possible

DOI: https://doi.org/10.5526/esj147

automatisation deficits[7] include a word identification task in which participants must select which one of six simultaneously presented words is a real word (the others are non-words); this is similar to the LADS word recognition task (Singleton et al., 2002b).

Category 3 (verbal and non-verbal reasoning) tasks do not operationalise deficit theories; however, the literature indicated that they might provide useful data. If dyslexics evidenced poorer results on verbal versus non-verbal reasoning tasks, this might indicate that the tasks were, in fact, dyslexia sensitive; the tasks involve similar skills, the main difference being the additional language loading on the verbal task, which might be expected to present difficulties for dyslexics. If dyslexics did not evidence deficits in these tasks, this would indicate that their difficulties were specific (i.e. related to dyslexia), rather than to a general learning disability or other co-morbid factor. This is intended to reassure students who are identified as probably dyslexic by a screening test. Exploring avenues of mitigating possible distress, a key ethical consideration, is often overlooked in test development research. However, following discussions and consultations with the Cambridge University Disability Resource Centre, this issue assumed a position of central importance in ensuring that the research adopted a sound ethical approach. As a benchmark of general ability, to which performance on tasks tapping dyslexic deficits could be compared, the data would also be useful for exploring within-individual performance discrepancies.

It has been suggested that dyslexics have increased skills (as opposed to deficits) associated with the right hemisphere, such as visuo-spatial abilities (Geschwind & Galaburda, 1987; Orton, 1925; West, 1991). Not only would such abilities provide a positive aspect to being identified as dyslexic, but would also be useful in developing compensatory strategies. Research has produced mixed results to date; Von Karolyi (2001) and Von Karolyi et al. (2003) demonstrated that dyslexics performed faster on an impossible figures test, though Winner et al. (2001) showed that dyslexics performed equal to, or worse than, controls on a number of other visuo-spatial tasks. Category 5 tasks (visuo-spatial abilities) were included to investigate this issue in the HE population.

Category 6 (self-report questionnaire) items were included to investigate other possible indicators of dyslexia which could not be objectively measured by a computer-based psychometric instrument (e.g. difficulty with everyday activities like writing cheques), and to give a cursory indication of whether there might be co-morbid difficulties, such as dyspraxia. This category provides the type

DOI: https://doi.org/10.5526/esj147

of detailed, qualitative information that an interview would tap. As such, this category was not directly linked to any key deficit theory. A learning styles questionnaire was also included, due to its potential to provide suggestions (albeit it basic ones) regarding compensatory strategies (to help mitigate distress in students newly identified as probably dyslexic), despite the fact that meta-analyses by Kavale et al. (1987, 1998), of 39 and 36 studies respectively, indicate that these approaches are not effective.

Methodology

Volunteer participants were recruited from Cambridge University in response to advertisements. Prior to participation, volunteers received information sheets, and signed an informed consent form. Due to the number of tasks involved, the material was divided into three one-hour testing-block sessions (administered in a consistent order). Following each test block, a brief, semi-structured interview was conducted to explore the functionality of the test and the usability of the interface.[8] To ensure anonymity, all data were stored under code-numbers.

The sample was divided into three groups: non-dyslexic, potentially-dyslexic and dyslexic students. The *dyslexic* group comprised students previously diagnosed as dyslexic by a formal assessment with an educational psychologist. Educational psychologists have no standardised method of formal assessment; individual assessors assume different definitions of dyslexia and employ different tests. Although desirable, it was not possible to assess all participants (or even re-assess all formally assessed 'dyslexic' participants) to reliably establish groupings; however, previous diagnosis by an educational psychologist is considered 'gold standard' evidence of dyslexia by most researchers and disability-support professionals. The *potentially-dyslexic* group comprised students who indicated that they might be dyslexic. Participants assigned to this group self-reported that they had similar problems to dyslexic family members, common dyslexic difficulties, and/or had been told by teachers/other professionals that they might be dyslexic, i.e. these participants had not been assessed by an educational psychologist, but had reason to believe that they might be dyslexic. The *non-dyslexic* group comprised students who believed they were not dyslexic, i.e. they had no reason to believe that they might be dyslexic.

DOI: https://doi.org/10.5526/esj147

Although relatively crude, these criteria for assigning participants to experimental groups are considered adequate in most dyslexia studies, and certainly are in the case of exploratory research. The addition of a *potentially-dyslexic* group refines these basic categories by separating cases likely to prove problematic.

Table 1: Socio-demographic data for pilot participants

	Non-dyslexic Group	Dyslexic Group	Potentially-Dyslexic Group
N	22	10	7
Male	2	6	3
Female	20	4	4
Mean age in years (3sf)	21.1	20.9	21.0
N attended comprehensive/ state school	12	6	5
N attended private/ independent school	10	4	2
Undergraduate	12	6	6
Postgraduate	10	4	1

Note. With regard to ethnicity: one of the dyslexic group answered 'Other', one of the potentially-dyslexic group answered 'Chinese', and all others answered 'White'.

Data are available for 10 dyslexic, 7 potentially-dyslexic and 22 non-dyslexic students for Block 1[9]. The three groups are roughly equivalent in terms of age, assumed cognitive ability and ethnicity, but not equivalent in terms of sex ratio or SES[10]. The sample is roughly representative of the University, with the exception of ethnicity and sex ratio (male to female ratio is approximately 2:3 as opposed to 1:1). Cambridge University is not representative of UK universities in general. General cognitive ability may be assumed to be particularly high at Cambridge University. Similarly, due to the bias towards students from private and independent schools, it may be assumed that more Cambridge students come from families with high SES.

In order to draw any firm conclusions about the validity or reliability of the test, or the individual tasks, the sample would need to be representative of the university, and greater equivalence between the testing groups would be necessary. For any results to be generalisable to the university population as a whole,

DOI: https://doi.org/10.5526/esj147

data would need to be gathered from a range of universities and other HE institutions across the UK. However, as the following analyses are intended to be exploratory, the limitations of the sample (with regard to validity, generalisability and reliability) merit a degree of caution but do not invalidate the usefulness of the data for this purpose.

Results

Three Block 1[11] tasks (see Tables 1, 2 and 3 respectively) will be used as examples to illustrate the most promising findings of the research: Sentence Construction (Which of the six words completes the sentence?), Sequencing (What number/letter completes the given logical sequence?), and Semantic Similarities (Which of the six words means the same as [e.g. cold]?).

Table 2:SentenceConstruction Task: Dyslexic versus Non-dyslexic Students

	Condition 1	Condition 2	Condition 3
n_{dys}	10	17	10
n_{nondys}	29	22	22
N	39	39	32
$Mean_{dys}$ time per item (secs)	9.50	9.07	9.50
$Mean_{nondys}$ time per item (secs)	7.26	6.88	6.88
SD_{dys} time (secs)	1.57	1.92	1.57
SD_{nondys} time (secs)	2.00	1.77	1.77
d.f.	37	37	30
t	3.20	3.69	4.00
Significance level: T-test (equal variances assumed)	0.003**	0.001***	0.000***
Sum of $Ranks_{dys}$	296	459	248

DOI: https://doi.org/10.5526/esj147

Sum of Ranks$_{nondys}$	485	322	281
Mann-Whitney U	49.5	68.5	27.5
Significance level: Mann-Whitney	0.001***	0.000***	0.000***

Note. n_{dys} represents the number of students in the dyslexic group. n_{nondys} represents the number of students in the non-dyslexic group. The assignment of students to groups is determined by the rules of the relevant experimental condition. Mean$_{dys}$ and SD$_{dys}$ for Condition 1 is the same as for Condition 3, i.e. n = 10, only diagnosed dyslexics are included in the dyslexic group. Mean$_{nondys}$ and SD$_{nondys}$ for Condition 2 is the same as for Condition 3, i.e. n = 22, the 7 potentially-dyslexic participants are not grouped with the non-dyslexic group in these Conditions. All decimal figures are given to 3 significant figures. This note applies to Tables 2, 3 and 4.

** p < .05 ** p < .01 ***p < .001*

Table 3:Sequencing Task: Dyslexic versus Non-dyslexic Students

	Condition 1	Condition 2	Condition 3
n_{dys}	10	17	10
n_{nondys}	29	22	22
N	39	39	32
Mean$_{dys}$ time per item (secs)	28.8	25.5	28.8
Mean$_{nondys}$ time per item (secs)	20.8	20.8	20.8
SD$_{dys}$ time (secs)	6.88	8.39	6.88
SD$_{nondys}$ time (secs)	5.56	4.53	4.53
d.f.	37	23.1	30
t	3.69	2.08	3.92
Significance level: T-test (equal variances assumed)	0.001***	[a] 0.049*	0.000***

DOI: https://doi.org/10.5526/esj147

Sum of Ranks$_{dys}$	296	406	240
Sum of Ranks$_{nondys}$	484	374	288
Mann-Whitney U	49.0	121	35.0
Significance level: Mann-Whitney	0.001***	0.063	0.002**

[a] Levene's test produced significant results, therefore equal variances not assumed formula employed.

** p < .05 ** p < .01 ***p < .001*

Table 4:SemanticSimilarities Task: Dyslexic versus Non-dyslexic Students

	Condition 1	Condition 2	Condition 3
n_{dys}	10	17	10
n_{nondys}	29	22	22
N	39	39	32
Mean$_{dys}$ time per item (secs)	14.8	13.6	14.8
Mean$_{nondys}$ time per item (secs)	11.7	11.7	11.7
SD$_{dys}$ time (secs)	2.98	3.53	2.98
SD$_{nondys}$ time (secs)	2.83	2.59	2.59
d.f.	37	37	30
t	2.96	1.97	3.05
Significance level: T-test (equal variances assumed)	0.005**	0.057	0.005**
Sum of Ranks$_{dys}$	281	413	229
Sum of Ranks$_{nondys}$	499	367	299
Mann-Whitney U	64.0	114	46.0

DOI: https://doi.org/10.5526/esj147

Significance level: Mann-Whitney	0.008**	0.039*	0.008**

* p < .05 ** p < .01 ***p < .001

It was not possible to discriminate dyslexic and non-dyslexic students on the basis of absolute scores on Block 1 tasks, though dyslexics required significantly longer than non-dyslexics to complete many of the tasks, mirroring results from earlier usability testing (see footnote 8).

The Sentence Completion task taps skills (comprehension and grammatical knowledge) not expected to be deficient in dyslexic university students; as anticipated, few participants answered any questions incorrectly, though dyslexics took significantly longer on the task than non-dyslexics. A highly significant difference was found between the groups (at the 0.003 and 0.001 levels respectively) using both parametric (between-samples T-test) and non-parametric (Mann-Whitney) tests.

On the Sequencing task, which taps areas of dyslexic difficulty – specifically, sequencing and automatisation skills (i.e. automatic knowledge of the alphabet/numerical system) - similar results were obtained, with no significant difference between groups in terms of absolute score, although dyslexics required significantly longer (at the 0.001 level for both T-test and Mann-Whitney) to complete the task than non-dyslexics.

Results from the Semantic Similarities task, which taps vocabulary and comprehension, followed the same pattern, with dyslexics completing the task significantly more slowly than non-dyslexic students (at the 0.005 level for T-test and 0.008 level using Mann-Whitney).

These analyses (Condition 1) employed the standard practice of assigning research participants to the dyslexic sample on the basis of previous dyslexic diagnosis, and considering all other participants non-dyslexic, including the potentially-dyslexic group (i.e. non-dyslexic and potentially-dyslexic groups are combined).

Grouping potentially-dyslexic students with formally diagnosed dyslexics (Condition 2), still resulted in a significant difference in task completion time between the dyslexic and non-dyslexic groups (Sentence Completion: at the 0.001 level T-test, and 0.000 level Mann-Whitney; Sequencing: at the 0.049 level T-test, and 0.063 level Mann-Whitney; Semantic Similarities: at the 0.057 level T-test, and 0.039 level Mann-Whitney), though the significance was borderline for

DOI: https://doi.org/10.5526/esj147

Semantic Similarities on the T-test and non-significant (though approaching significance) for Sequencing on the Mann-Whitney.

When potentially-dyslexic students were omitted from analysis (Condition 3), the difference between the groups was highly significant at the 0.000 level (T-test and Mann-Whitney) for Sentence Completion, at the 0.000 (T-test)/0.002 (Mann-Whitney) level for Sequencing, and the 0.005 (T-test)/ 0.008 (Mann-Whitney) level for Semantic Similarities.

Table 5:Reading Time for Task Instructions: Dyslexic versus Non-dyslexic Students (Condition 1 and Condition 3)

	Sentence Construction		Sequencing		Semantic Similarities	
	Cond. 1	Cond. 3	Cond. 1	Cond. 3	Cond. 1	Cond.3
n_{dys}	10	10	10	10	10	10
n_{nondys}	29	22	29	22	29	22
N	39	32	39	32	39	32
$Mean_{dys}$ time per item (secs)	9.21	9.21	38.3	38.3	42.0	42.0
$Mean_{nondys}$ time per item (secs)	8.89	8.38	30.8	30.8	27.3	26.8
SD_{dys} time (secs)	2.47	2.47	6.50	6.50	16.5	16.5
SD_{nondys} time (secs)	3.67	2.66	8.11	7.53	9.81	7.66
d.f.	37	30	37	30	11.3	10.8
t	0.258	0.839	2.64	2.69	2.66	2.79
Significance level: T-test (equal variances assumed)	0.798	0.408	0.012*	0.011*	[a] 0.022*	[a] 0.018*
Sum of $Ranks_{dys}$	224	191	280	228	288	234
Sum of $Ranks_{nondys}$	557	338	500	300	493	295
Mann-Whitney U	122	84.5	65.0	47.0	57.5	41.5

DOI: https://doi.org/10.5526/esj147

Significance level: Mann-Whitney	0.456	0.305	0.009**	0.009**	0.004**	0.004**

[a] Levene's test produced significant results, therefore equal variances not assumed formula employed.

** p < .05 ** p < .01 ***p < .001*

The time taken to complete tasks (i.e. processing-speed) appears to be a robust and consistent discriminator, with dyslexic students even requiring significantly longer to read instructions on Sequencing (at the 0.012 T-test/0.009 Mann-Whitney level for Condition 1; 0.011/0.009 for Condition 3) and Semantic Similarities (at the 0.022 T-test/0.004 Mann-Whitney level for Condition 1; 0.018/0.004 for Condition 3), though results did not reach significance for Sentence Completion for Condition 1 or 3.

Semantic Similarities taps comprehension, Sequencing taps sequencing skills for verbal and non-verbal information, and Sentence Construction taps knowledge of written language rules (grammar and spelling); therefore, the dyslexic students' apparent processing-speed deficit is found on verbal and non-verbal tasks, across a range of cognitive domains, and is not restricted to phonological processing or reading speed.

Discussion

These findings provide support for the argument that traditional adult dyslexia tests, which often rely on absolute scores, may be inappropriate for university students, since high-achieving dyslexics may obtain ceiling-level scores even on tasks tapping dyslexic difficulties. Instead, the time taken to complete items, tasks and even reading of task instructions appears to have a greater discriminatory power for this population.

The importance of timing variables can be conceptualised as a measure of processing speed; therefore, the data (especially from Sequencing) could be interpreted as supporting the automatisation deficit theory. Sentence Completion and Semantic Similarities can be grouped with

DOI: https://doi.org/10.5526/esj147

instruction reading as involving reading plus comprehension; it could be argued that in non-disabled, adequately educated adults, reading is usually an automatic skill, as is knowledge of the number system/alphabet (as in Sequencing). However, all deficit theories relate to posited fundamental processing difficulties which underlie problems in reading; therefore, the findings could, equally, be used to support these theories. Only by comparing results from a wide range of tasks, which effectively operationalise the different deficit theories, as well as other relevant areas of research, could this claim be lent significant weight.

Significance levels improved when potentially-dyslexic students were omitted from analyses (Condition 3), rather than grouped with either the non-dyslexics (Condition 1) or with the diagnosed dyslexics (Condition 2), suggesting that the potentially-dyslexic group is not homogenous but contains both non-dyslexic and (previously unidentified) dyslexic students. The fact that these tasks appear sensitive to the presence of undiagnosed dyslexics is particularly promising, as a screening test must identify such students rather than simply replicating experimental groupings, i.e. distinguishing students with previous, formal diagnoses of dyslexia from those without.

Combining the potentially-dyslexic group with the dyslexic group (Condition 2) increased the significance of results for Sentence Construction, but not Sequencing or Semantic Similarities. These finding might indicate that Sequencing and Semantic Similarities are less sensitive (significance levels were lower for all conditions in Semantic Similarities versus Sentence Construction). However, this could also be interpreted as indicating that not all the potentially-dyslexic students were actually dyslexic, and that these two tasks discriminate between those students in the potentially-dyslexic group who are, in fact, dyslexic and those who are not.

DOI: https://doi.org/10.5526/esj147

Table 6:Mean Task Times for Individual Participants in the Potentially-Dyslexic Group versus Mean Times for Dyslexic and Non-Dyslexic Groups

Participant Identifier	SentenceCompletion (secs)		Sequencing (secs)		SemanticSimilarities (secs)	
	Task	Instruction Reading	Task	Instruction Reading	Task	Instruction Reading
AA	17.7**	36.7*	12.2**	12.6**	27.8*	33.4
BB	5.33	9.60	4.40	4.10	9.99	15.4
CC	11.1	23.6	7.70	10.2**	18.8	47.8**
DD	12.6	20.4	9.43*	12.9**	17.0	29.4
EE	12.4	56.9**	8.90*	6.60	36.1**	23.8
FF	13.6*	35.6*	8.15	5.70	17.9	37.7*
GG	10.3	20.3	8.48*	21.3**	17.8	26.5
$Mean_{dys}$ (n=10)	14.8	42.0	9.50	9.21	28.8	38.3
$Mean_{nondys}$ (n=22)	11.7	26.8	6.88	8.39	20.8	30.8
Mid-point	13.3	34.4	8.19	8.80	24.8	34.6

Note. All decimal figures are given to 3 significant figures.

* *mean time of potentially-dyslexic participant is above the mid-point between the mean times of the dyslexic and non-dyslexic group*

** *mean time of potentially- dyslexic participant is higher than the mean time for the dyslexic group*

By examining the data from the 7 potentially-dyslexic students on a case-by-case basis, it is possible to strengthen the case for these interpretations. Participant AA's mean item times were more than those of the dyslexic group as a whole on Sentence Completion and Sequencing, and above the

DOI: https://doi.org/10.5526/esj147

midpoint (between the mean time of the dyslexic group and the non-dyslexic group[12]) on Semantic Similarities; instruction reading times were above the mid-point for Sentence Completion, and higher than the dyslexic group on Sequencing. Only with regard to the instruction reading time for Semantic Similarities were the data for Participant AA more similar to the non-dyslexic group. In a further Block 1 Clocks task (matching clock-faces to numeric or verbal descriptions of a specific time), which has face validity as a dyslexia sensitive measure, Participant AA's instruction reading time and mean task time were higher than the mean times for the dyslexic group. This appears a strong indication that Participant AA is dyslexic, though not yet formally diagnosed as such.

On the other hand, Participant BB's data is more similar to the non-dyslexic group than the dyslexic group on all counts, including the Clocks task; therefore, it is likely that Participant BB is not dyslexic, despite self-reporting concerns about being an undiagnosed dyslexic.

The pattern of data from the other 5 potentially-dyslexic participants (CC-GG) is less clear cut. This may be due to the fact that they are borderline cases, have performed in anomalous ways, or because the small samples sizes involved means that the dyslexic and non-dyslexic means have been skewed by individual anomalous results. The results do support the conclusion that the potentially-dyslexic group cannot be effectively grouped with either dyslexics or non-dyslexics.

This would explain why the significance of results increases when this group are omitted from analyses (Condition 3 versus Condition1) for all three tasks, including instruction reading. The fact that the potentially-dyslexic group is not cohesive also clarifies why it cannot be effectively combined with the dyslexic or non-dyslexic group, and why Conditions 1 and 2 evidence an inconsistent pattern of results; combining the potentially-dyslexic group and non-dyslexic groups (Condition 1) gave higher significance levels than combining the potentially-dyslexic group and dyslexic groups (Condition 2) on Sequencing and Semantic Similarities, and had little effect on significance for Sentence Construction. The fact that combing the potentially-dyslexic group with the dyslexic group (Condition 2) has a negative effect on significance levels may indicate that the 5 borderline cases (Participants CC-GG) are better grouped with Participant BB (probably non-dyslexic), i.e. as a whole these cases represent a more non-dyslexic, than dyslexic, profile of results. This might also explain why the group mean for non-dyslexics on Sequencing, Semantic Similarities, including instruction reading, does not change when potentially-dyslexic students are included.

DOI: https://doi.org/10.5526/esj147

Equally, a simpler explanation is that, as the non-dyslexic group is twice as large as the dyslexic group (n_{dys}=10, n_{nondys}= 22), the results of the potentially-dyslexic group exert a lesser effect on the non-dyslexic group than on the dyslexic group.

The results of this research support the view that university students use compensatory strategies and high general ability to perform at ceiling level, even on tasks previously identified by the literature as tapping areas of dyslexic deficit. Although dyslexic students may achieve equivalent scores to non-dyslexics, precise timing data from computer-based administration of the tasks demonstrate that it usually takes them significantly longer to do so. This speed deficit appears to be consistent across a range of tasks, including reading of task instructions on Sequencing, Semantic Similarities, though not Sentence Construction, perhaps because the instructions for this task were particularly short and simple.

These findings have far-reaching implications for the ability of dyslexic students to cope not merely with the volume of reading, but also turn-over of essays and other work, at university. Thus the reasonable adjustments (e.g. extra time in examinations) for disabled students, required by the Disability Discrimination Act, do not constitute an unfair advantage, as some people argue, but are vital in allowing dyslexics to study without disadvantage compared to non-dyslexic students.

Conclusion

This explorative research described the design and piloting of a dyslexia screening test-prototype for HE students. Instead of operationalising a single definition of dyslexia, this study employed a wide-range of (deficit and enhanced abilities) theories to examine which tasks were effective in identifying dyslexic students. A number of promising avenues for further work were revealed: processing speed appears to be a far stronger indicator of dyslexia than absolute scores in the HE population, and complex tasks, tapping multiple deficits (thereby circumventing dyslexic students' coping strategies), tend to be more sensitive than tasks which involve a single area of deficit. These results offer valuable insights into the challenges of identifying dyslexia in high-performing adults and provide a strong foundation for further work to implement a computer-based screening test for dyslexia suitable for the HE population.

DOI: https://doi.org/10.5526/esj147

Acknowledgements

The original research described in this article was conducted at, and financed by, the Centre for Applied Research for Educational Technologies (CARET), University of Cambridge, in collaboration with the Faculty of Education, University of Cambridge, and the Cambridge University Disability Resource Centre. Major contributors to the research include: Jem Rashbass, Martyn Rouse, Sue Danson, Graham Phillips, Heidi Tranberg, Chris Applegate, Tim Froggart, Jenny Haynes, Mary Marshall, and Judith Jesky.

The author presented a paper on the preliminary findings of the research at the Bangor Dyslexia Conference, 24-27th July, 2003.

References

Bryman, A. & Burgess, R.G. (Eds.) (1993). *Analyzing qualitative data.* London: Routledge.

Casale, Alexia. (2009). Identifying Dyslexic Students: The need for computer-based dyslexia screening in higher education. *Estro*, 1(1), 119-143.

Fawcett, A.J., & Nicolson, R.I. (1998). *Dyslexia Adult Screening Test.* London: Psychological Corporation.

Gerschwind, N. & Galaburda, A. (1987). *Cerebral lateralization.* Cambridge, MA: MIT Press.

Glaser, B.G. & Strauss, A.L. (1967). *The discovery of grounded theory: strategies for qualitative research.* Hawthorne, NY: Aldine de Gruyter.Kavale, K., & Forness, S. (1987). Style over substance: Assessing the effectiveness of modality testing and teaching. *Exceptional Children,* 54, 229-239.

Kavale, K., Hirshoren, A., & Forness, S.R. (1998). Meta-analytic validation of the Dunn and Dunn model of learning-style preferences: A critique of what was Dunn. *Learning Disabilities Research and Practice,* 13 (2), 75-80.

Orton, S.T. (1925). 'Word blindness' in school children. *Archives of Neurology and Psychiatry, 14,* 581-613.

DOI: https://doi.org/10.5526/esj147

Raven, J.C., Court, J.H., & Raven, J. (1962). *Advanced Progressive Matrices*. Oxford: Oxford Psychologists Press.

Singleton, C.H., Horne, J.K., & Thomas, K.V. (2002b). *LADS: Lucid Adult Dyslexia Screening.* Beverley, East Yorkshire: Lucid Research.

von Karolyi, C. (2001). Visuo-Spatial strength is dyslexia: Rapid discrimination of impossible figures. *Journal of Learning Disabilities, 34 (4),* 380-391.

von Karolyi, C., Winner, E., Gray, W., & Sherman, G.F. (2003). Dyslexia linked to talent: Global visual-spatial ability. *Brain and Language*, 85 (3), 427-431.

Weschler, D. (1955). *Weschler Adult Intelligence Scale.* Texas: Psychological Corporation.

West, T. (1991). *In the Mind's Eye.* Buffalo, N.Y: Prometheus Books.

Wilkinson, G.S. (1993). *Wide Range Achievement Test.* Wilmington, DE: Wide Range.

Wolf, M., & Bowers, P.G. (1999). The double-deficit hypothesis for the developmental dyslexias. *Journal of Educational Psychology, 91 (3),* 415-438.

World Federation of Neurology (1968). Report of research group on developmental dyslexia and world literacy. *Bulletin of the Orton Society, 18,* 22-22.

Footnotes

1 The use of *'…with individuals evidencing unique profiles of strengths and weaknesses'* picks up on the International Dyslexia Association (1988) definition, which states that *'People with dyslexia are unique; each having individual strengths and weaknesses'.* The British Dyslexia Association definition (2003) refers to a range of specific, everyday difficulties, such as '*difficulties with sequences'*, reflected in the current definition's inclusion of: '*Other soft signs may include laterality discrimination and sequencing difficulties*'.

2 It is not possible to operationalise the theories exhaustively, as there is a limit to how much time participants can be asked to commit to the research.

3 Another Category 4 word identification task that can be viewed as operationalising the automatisation deficit theory (as opposed to the phonological deficit theory) requires participants

DOI: https://doi.org/10.5526/esj147

to decide if a very briefly presented stimulus is a real word or a non-word. Category 5 includes a visual search task (participants must locate a given target in a display of distractors) and a mental rotation task (participants must rapidly decide whether or not a stimulus has been rotated).

4 Two sets of usability testing were carried out to: i) assess the functionality of the computer-based version (programmed in Flash Macromedia MX) of the tasks, ii) examine the effectiveness of the user interface, and iii) refine the tasks and individual task items with regard to issues such as wording of instructions, complexity, and time-limits. The first set of usability testing involved 13 non-dyslexic volunteers, while the second set involved 14 non-dyslexic and 4 dyslexic volunteers, recruited from among Cambridge University staff. On the basis of these data, an item analysis of all tasks was carried out and a number of individual items were excluded as ambiguous, confusing or of inappropriate complexity (i.e. all/no participants answered the item correctly). Minor refinements were also made to the user interface and the computer programme (e.g. an automatic cut-out function after multiple failures was added to progressive tasks such as digit span).

5 Block 1 was always administered first. Data from 39 participants was collected for Block 1 tasks. Only 11 (3 dyslexic, 2 potentially-dyslexics and 6 non-dyslexic) and 10 (1 dyslexic, 3 potentially-dyslexic and 6 non-dyslexic) participants, respectively, had completed Blocks 2 and 3 when the project ended. Therefore, as there was insufficient data to conduct any meaningful analyses – even exploratory ones – for Blocks 2 and 3, only Block 1 data was analysed in this paper.

6 The mean age for all groups (approximately 21 years) was representative of the University student population. The male to female ratio varied widely across the groups (for the university as a whole, the ratio is approximately 1:1). The 'type' of school a student attended before university can be viewed as a general measure of student's family's SES. Cambridge University accepts approximately 44% of its students from private and independent schools versus comprehensive and grammar schools; the dyslexic and potentially-dyslexic groups were roughly representative of the university population, though the data for non-dyslexics was skewed towards private and independent schools (i.e. indicating that these students came from families with high SES). The groups differed widely on the ratio of undergraduates to postgraduates. All students at Cambridge are expected to have at least three A's at A-level, indicating a very high level of academic achievement; this constitutes a reasonable measure of general cognitive ability, on which all groups were assumed to be equivalent. Beyond this, no further measures were taken to assess equivalence of academic and/or cognitive ability of participants; the number of undergraduates to postgraduates provides no further information about academic ability once it is considered that all the undergraduates were younger than the postgraduates. While approximately 19% of Cambridge students are from ethnic minorities, only two participants across all three testing groups gave an answer other than 'White'; therefore, the groups are approximately equivalent on this criterion, though they are not representative of the university population as a whole.

DOI: https://doi.org/10.5526/esj147

[7] It was considered unethical to provide participants with information about their performance during the pilot of the test prototype due to the fact that it was not possible to accurately or reliably explain the import of individual results.

[8] Mid-point figures refer to the time (in seconds) which is mid-way between the mean time of the dyslexic and non-dyslexic groups, i.e. $(Mean_{dys} - Mean_{nondys}) / 2 + Mean_{nondys}$. Times above the mid-point are closer to those of the dyslexic group than the non-dyslexic group. As standard deviations are large (probably due to the small sample sizes), using the mid-point gives a more effective measure of how similar a given time is to the dyslexic or non-dyslexic group than attempting to locate its position in standard deviations from either group.

DOI: https://doi.org/10.5526/esj147