



Identifying dyslexia at the university: assessing phonological coding is not enough

Helle Fredslund Ottosen¹ · Katrine H. Bønnerup² · Ethan Weed³ · Rauno Parrila⁴

Received: 16 February 2021 / Accepted: 7 November 2021 / Published online: 10 March 2022
© The International Dyslexia Association 2022

Abstract

A dyslexia diagnosis in Denmark can have significant consequences for individuals, as support is not available to others with reading difficulties. Currently, the diagnosis is given solely on the basis of an electronically administered test consisting of two tasks assessing grapheme-phoneme correspondences. To examine whether the Danish diagnostic test is sufficient to identify university students with dyslexia, we compared the performance of 239 Danish university students who reported literacy difficulties and were tested for dyslexia with the Danish diagnostic test on three word-level tests (low-frequency word reading, high-frequency word reading and spelling to dictation) with the performance of separate control groups for each test: 220, 212 and 218 students, respectively. The results showed that 61% of students labelled “not dyslexic” by the Danish diagnostic test performed significantly worse than controls on at least two out of three word-level tests. In terms of self-report of literacy difficulties, students labelled “not dyslexic” by the diagnostic test were indistinguishable from those labelled “dyslexic.” These findings suggest that the current method of diagnosing dyslexia in Denmark is too narrow and that adding a few simple tests of word reading and spelling would minimize the risk of overlooking students in need of literacy support.

Keywords Diagnosing dyslexia · Literacy difficulties · Self-report · University students · Word-level difficulties

The research was completed while Helle Fredslund Ottosen was affiliated with the Counselling and Support Centre at Aarhus University, Denmark.

✉ Helle Fredslund Ottosen
helle22_7@hotmail.com

¹ Independent Scholar, Odense, Denmark

² Counselling and Support Centre, Aarhus University, Aarhus, Denmark

³ Department of Linguistics, and Language As a Tool for Learning Center, Aarhus University, Aarhus, Denmark

⁴ School of Education, and Macquarie Centre for Reading, Macquarie University, Sydney, Australia

Introduction

Dyslexia assessment should meet the highest possible standards of validity and reliability and produce as few false negatives and false positives as possible. In practice, however, there are additional factors to take into consideration. Often the reason for testing for dyslexia is to assess eligibility for support, and thus, the question of assessment also becomes a financial issue. As a result, the assessment situation must be as efficient as possible to minimize the time and resources spent on the assessment, as these resources can be better spent on support. According to Pennington et al. (2012), there has been a tendency towards basing a diagnosis of dyslexia on the presence of a phonological deficit. In line with this, Denmark released in 2015 a state-mandated, electronically administered, diagnostic test (henceforward *the Danish Diagnostic Test; DDT*) targeting only phonological coding. In this paper, we will examine whether this assessment method is sufficient for diagnosing dyslexia accurately at the university level.

The question of how to diagnose dyslexia is closely intertwined with how dyslexia is defined and explained. Internationally, definitions of dyslexia do not deviate substantially. The International Dyslexia Association (IDA) has adopted a definition stemming from the work of Lyon et al. (2003), who describe dyslexia as “[...] difficulties with accurate and/or fluent word recognition and [...] poor spelling and decoding abilities [...] which typically result from a deficit in the phonological component of language.” The American Psychiatric Association’s (APA) DSM-5 adopts Lyon et al.’s definition nearly word for word, describing dyslexia as “[...] a pattern of learning difficulties characterized by problems with accurate or fluent word recognition, poor decoding, and poor spelling abilities” (APA, 2013). Parrila and Protopapas (2017) also focus on the symptom level, defining dyslexia more specifically as “a persistent and unexpected difficulty in developing age- and experience-appropriate word reading skills” (p. 333), stressing word level as the primary affected component in the reading process. They further define this as including both accuracy and efficiency in reading words.

In contrast to these international definitions focusing on the behavioral level of the disability, the DDT takes as its vantage point the cognitive level. Thus, in the official guidelines to the DDT, dyslexia is described as difficulty “connecting letters and letter sequences to their normal sounds. This basic difficulty in learning to use the fundamental phonemic principle renders the reading of especially new words inaccurate and slow, and it leads to long-term problems with spelling. Difficulties at this basic level are called dyslexia” (Center for Læseforskning, 2020, our translation). This definition is based on a purely phonological view of dyslexia, following the tradition that dyslexia is explained by a phonological processing deficit.

The phonological deficit theory (e.g. Snowling, 2000) poses the widely accepted and empirically supported hypothesis that dyslexia is caused by deficient phonological representations, which render the acquisition of phonological skills involved in the reading process (phonological awareness and decoding) difficult (Vellutino & Fletcher, 2005). It should be noted, however, that this theory builds primarily on empirical studies involving children (Snowling et al., 1997). Even though the phonological deficit theory of dyslexia is predominant, several other possible explanations of dyslexia have been proposed and tested empirically. Among the prevailing theories are the subtyping theory, distinguishing between phonological and surface dyslexia (Castles & Coltheart, 1993), and the double-deficit theory, according to which either a phonological deficit, a naming speed deficit or a combination of both can cause reading difficulties (Wolf & Bowers, 1999). The subtype

of “surface dyslexia” proposed by Castles and Coltheart (1993) is characterized by a deficit in reading irregularly spelled words, as opposed to “phonological dyslexia” which is characterized by a deficit in reading nonwords. However, the empirical evidence of these subtypes is conflicting (see Manis et al., 1996, 1999; Peterson et al., 2013, 2014; Stanovich et al., 1997). Double-deficit theories have found support in some studies (Wolf & Bowers, 1999; Kirby et al., 2003; see review in Kirby et al., 2010) showing that naming problems predict reading fluency and reading accuracy, and according to Parrila and Protopapas (2017), “performance on RAN [rapid automatized naming] tasks helps identify children with reading difficulty that cannot be attributed to phonological deficits” (p. 337).

Parrila and Protopapas (2017) argue, however, that neither the existing single- nor double-deficit models of dyslexia provide sufficient explanation of the disorder resulting in word reading difficulties. This is in line with Pennington et al. (2012) and Erskine and Seymour (2005). In the extensive study by Pennington et al. (2012) of test results from children with and without dyslexia (defined as inaccurate or dysfluent decoding of words) from several different countries and orthographies, the analyses indicated that no single-deficit model could account for the majority of cases of dyslexia. Instead, Pennington et al. (2012) found that a so-called hybrid model, which allows for several different routes to dyslexia, was the best model to account for the data. This resonates with the findings of Erskine and Seymour (2005) who suggested a “cognitive mosaic model” to describe reading impairments, based on a study showing how individual university students with dyslexia displayed varying patterns of performance on a range of cognitive tests. Pennington et al. (2012) concluded that “using the presence of a PA [phonological awareness] deficit either to screen for dyslexia or to confirm a dyslexia diagnosis would miss about half the cases of dyslexia” (p. 221). In sum, it is contestable whether a phonological deficit is the sole possible explanation of dyslexia, although it is widely accepted as a common cause.

Regardless of how one explains dyslexia at the cognitive level, testing will always have to target the behavioral level. Common to the definitions highlighted in the previous paragraph is the concept of dyslexia as an impairment in word reading (and spelling). It is important to note that “dyslexic behaviour” (in terms of reading and writing abilities) is not simply the product of a specific genetic makeup resulting in clear-cut neurological structures that, in turn, cause specific cognitive processes. Many other factors, both cognitive and environmental, have the potential to shape an individual’s reading and writing skills (Erskine & Seymour, 2005; Frith, 1999; Pennington, 2006). In line with this, assessment guidelines usually take a broad approach, recommending that a diagnosis of dyslexia should be based on multiple measures. For instance, assessment guidelines of the IDA, the Swedish Dyslexia Association, the Norwegian Dyslexia Association, and the British Specific Learning Difficulties Assessment Standards Committee (SASC) recommend assessing areas such as RAN, word reading, vocabulary, oral reading fluency, reading speed, reading comprehension, spelling and writing in addition to phonological measures such as phonological awareness, phonological decoding and phonological memory (Aakermann et al., 2017; Engh et al., 2017; IDA, 2017; SASC, 2005). In Great Britain, the SASC decides which test batteries are suitable for the assessment of students applying for Disabled Students Allowances and recommends, for example the standardized assessment battery WIAT-III, which allows the clinician to base a diagnosis on a combination of test standard scores, qualitative analyses and the student’s history of reading and writing difficulties.

In contrast, a diagnosis of dyslexia in Denmark is based on a person’s standard scores from two subtests targeting phonological coding. To the best of our knowledge, Denmark is the only country where an official diagnosis of dyslexia is based solely on performance on phonological coding (see Mather et al., 2020, for a recent review of assessment and

intervention methods around the world). The Danish diagnostic test (officially named *Ordblindetesten* [word-blindness test]) was developed by Møller et al. (2014) as an electronic test with the purpose of standardizing and disambiguating assessment for dyslexia in Denmark (Poulsen et al., 2016). The test was developed and validated for various educational levels and is currently used throughout the country and across educational levels as the basis for determining whether a student has dyslexia and is thereby eligible for special educational support. A combined standard score from the two phonological coding subtests places the student in one of three categories, based on cutoff points that take into account the student's educational level. Thus, higher education students whose standard score is below 82 are categorized as "dyslexic" (graphically represented with a red colour in the DDT report); higher education students with a standard score between 82 and 88 are categorized as having "inaccurate phonological coding" (represented with yellow) and higher education students with a standard score above 88 are categorized as "not dyslexic" (represented with green). Henceforward, instead of using "dyslexic," "inaccurate phonological coding," and "not dyslexic" to label the three categories in DDT, we will refer to the categories as red, yellow and green in order to avoid any confusion with the general terms "dyslexic" and "not dyslexic." Students falling into the red category qualify directly for state-financed support, while students in the yellow category receive further assessment in order to evaluate the severity of the difficulties in the context of the student's educational level and, hence, the student's eligibility for support. This further assessment can include tests of simple and complex word reading, spelling and reading comprehension and considers the students' history of literacy difficulties. Green students do not qualify to receive support.

A substantial number of students with dyslexia are currently enrolled in Danish higher education institutions. Exact numbers are unknown, but of the approximately 262,000 students enrolled in higher education across the country in 2018 (Danmarks Statistik, 2020a, b, c, d), 6268 received special educational support due to dyslexia (Styrelsen for Undervisning og Kvalitet, 2020, personal communication), amounting to 2.4% of the student population. At Aarhus University in 2018, 527 students with dyslexia received support, amounting to 1.6% of the total student population of 33,112 (Aarhus University, 2018). However, both these numbers underestimate the prevalence of dyslexia in higher education as they disregard students who have an earlier dyslexia diagnosis but do not currently receive any support, as well as students who have not yet been diagnosed with dyslexia.

To help students with disabilities, the Danish state has invested resources in providing services and support to assist in overcoming their difficulties, so that the state, instead of the individual educational institution, bears the expense of the support. Thus, students in higher education diagnosed with dyslexia have access to assistive technology, such as text-to-speech, speech-to-text, word prediction and OCR software, as well as recurring personalized literacy counselling that supports these students in developing strategies to improve their academic reading and writing. In addition, a diagnosis of dyslexia makes it possible for them to apply for exam accommodations such as extra time and the use of their assistive technology. While this support system is of very high standard internationally, access to the support system is limited to students who are diagnosed with dyslexia based on the DDT.

Because the assessment method used in Denmark deviates from international common practice, it is important to investigate the sensitivity of the DDT. This is particularly important in the context of university students. Several studies across languages (Bønnerup et al., 2019; Hatcher et al., 2002; Re et al., 2011; Tops et al., 2012) have shown that students in higher education with developmental dyslexia continue to experience problems with reading and writing, including decoding, word recognition, reading fluency, spelling and

writing speed. Tops et al. (2012) found that testing word reading, word spelling and phonological awareness was sufficient to identify students with dyslexia. In line with this, Bønnerup et al. (2019) found that, as a group, Danish higher education students with dyslexia had phonological decoding difficulties but also that they were impaired in spelling and in reading morphologically complex words and complex text. Broggi et al. (2019) found that a few phonological processing tasks (nonword reading, a letter exchange task and syllable exchange task) could be effective as an instrument for screening for reading disabilities at universities, but the authors also make the point that this small battery of phonemic tasks “should be used as a screening tool and not the only source of information for RD screening” (p. 263), and stress the importance of self-reported difficulties with tasks related to reading and writing.

However, having made it into tertiary education without being diagnosed, university students with reading impairment are individuals who have significant cognitive resources at their disposal and have been able to compensate for their underlying difficulties for a large portion of their formal education. A few studies (Cavalli et al., 2016; Parrila et al., 2007) have shed light on the strengths of high-functioning dyslexics, which allow them to compensate for their reading and writing difficulties to some extent. For example, Parrila et al. (2007) found that university students with a history of reading difficulties were able to compensate when presented with phonological tasks, using other strategies than the ones meant to be measured by the test. In turn, Hebert et al. (2018) found no differences between students with and without a history of reading difficulties on measures of reading comprehension, whereas significant differences on word reading and reading times for both words and text were present. Bønnerup et al. (2019) reported that Danish university students with dyslexia have the capacity to at least partially compensate for the effects of a phonological deficit by using other cognitive resources. This study also found that dyslexic students in 5- to 6-year university programs performed significantly better on measures of reading and spelling than dyslexic students studying for a specific profession in 3.5- to 4-year professional bachelor programs (e.g. nursing, engineering and teaching). The fact that even among these two groups of higher education students literacy difficulties were expressed differently suggests that university students are likely a unique population.

Each year, approximately 120 higher education students who self-report literacy difficulties are assessed for dyslexia at Aarhus University. Of the assessed students, approximately 25% fall into the green category on the DDT and therefore do not qualify for support. For some of these students, the assessment interview clearly suggests other plausible explanations for their difficulties, such as reading comprehension difficulties unrelated to word reading, attention deficits or psychiatric disorders. However, a considerable number of these “green” students report a history of literacy difficulties similar to students diagnosed with dyslexia. As several studies have suggested that self-report of reading acquisition difficulties in itself is often a strong indicator of dyslexia (e.g. Deacon et al., 2012; Parrila et al., 2007), this casts some doubt on the sensitivity of the DDT.

Given the literature reviewed above suggesting that (1) we risk overlooking individuals with dyslexia when testing only for phonological difficulties, (2) self-report is in itself often a marker of dyslexia and (3) students with dyslexia at the university level have often developed extensive compensatory strategies, together with (4) international recommendations that the assessment of dyslexia at all levels includes a wide range of measures, we hypothesize that a significant number of students with self-reported literacy difficulties testing in the normal range on the DDT will perform poorly on other measures targeting the key behavioural symptoms of dyslexia: poor word reading and poor spelling. Thus, the purpose of the present study is to investigate

whether it is sufficient to test students at university for dyslexia by means of the phonology-based diagnostic test currently used in Denmark. We examine this by comparing the performances of university students from the three diagnostic categories—red, yellow and green—from the DDT on three additional tests targeting word-level difficulties with each other and with the performance of control students.

A few characteristics of Danish are worth mentioning. Danish has a moderately inflective morphology, and in the context of European languages, it can be characterized as having a complex syllabic structure and a relatively deep orthography with a high number of inconsistent phoneme-grapheme correspondences and morphological influences on spelling (Seymour et al., 2003). According to Juul and Sigurðsson (2004), roughly 60% of the five hundred most common words have perfect phoneme-grapheme correspondences. However, this is often not the case for less frequent words, e.g. the academic and subject-specific vocabulary in curricula in higher education, many of which stem from other orthographies. Words from French, Latin and Greek have typically been adapted to the Danish orthography to some extent although many have retained parts of the original spelling. This means that even if a Danish speaker knows a so-called “foreign” word and its original spelling, it is seldom possible to spell the word correctly in Danish by using that original spelling (e.g. French spelling) because the spelling of parts of the word has changed. However, it is also not possible to spell the entire word correctly by applying native Danish orthographic rules because the original spelling has been retained for other parts of the word. Examples of this are *fysioterapeut* (in English “physiotherapist”) from the Greek *phýsis* and *therapeutés* or *buschauffør* (“bus driver”) from the French *chauffeur*.

Method

Participants

The participants were two groups of university students at Aarhus University: students with self-reported literacy difficulties, who were referred and assessed for dyslexia, and control students who reported not experiencing literacy difficulties. Only participants with Danish as their first language and normal or corrected-to-normal vision were included in this study. We have no individual data on the participants’ proficiency in foreign languages, but all Danish higher education students have received substantial instruction in the English language through 9–13 years schooling prior to entering university, as well as 5–9 years of instruction in German and/or French.

The data from students with literacy difficulties was collected as part of a clinical dyslexia assessment, and the data from control students were collected for the purpose of obtaining norms for the battery of supplementary tests used at Aarhus university to further assess the students who are classified as yellow. Consequently, the control group data consists of data from four different groups of students without literacy difficulties, as described below. Furthermore, due to government guidelines for the use of the DDT, we were not allowed to use the test to assess students with no reported literacy difficulties but compared students with literacy difficulties to the existing norms. This, of course, has consequences for the array of analyses we were able to conduct (see our discussion of limitations).

Students tested for dyslexia

All university students tested for dyslexia with the DDT at Aarhus University from 2015 to 2019, who consented to their test results being used for scientific purposes, and who completed all three additional tests, were included in this study. A total of 239 students were categorized by the DDT into three categories: red (127 students), yellow (48 students) and green (64 students). Because the testing of students with literacy difficulties was part of an assessment in which individual concerns were prioritized, only a subgroup of students tested for dyslexia with the DDT also completed a questionnaire about their reading history.

Control students

To recruit control students for the word reading tasks, we approached students in cafeterias and public study areas and asked them if they would be willing to complete a reading task for us. The students who, prior to testing, answered “no” when asked if they experienced any difficulties with reading, writing or spelling in their studies were then included in the final sample. The spelling data were collected from students who volunteered to complete the spelling task after their lecture and who also answered “no” to a question about experiencing reading, writing and spelling difficulties. The ARHQ-R-DA data were also collected from students who volunteered to complete the online questionnaire after their lecture. Data from these students were collected one test at a time as part of a preliminary validation of each test in order to minimize the time the testing took for each control student as we could not offer them any compensation for their time. As a result, the number of control participants varied slightly between the three tests, and each control participant only completed one test. Thus, the control groups consisted of 218, 220 and 212 students on the three tests respectively, and the control groups for the ARHQ-R-DA consisted of 336 students (primary school¹ scale) and 113 students (current status scale) (Table 1). No control participants completed DDT (see above).

Table 1 shows the distribution of gender, age and fields of study for all groups. The gender distribution in the green, yellow and red groups reflects the population that seeks help from the literacy support centre at Aarhus University, as well as the fact that more women than men pursue a university education in Denmark.² The overall ratio of male to female university students is also reflected in the gender distribution in the control group. Although there was a relatively small number of older participants in our sample, we limited the data presented here to students between the ages of 18–30. The full data set is available at https://osf.io/h5vgm/?view_only=c05ed06114e8425eb655c7669a3efa14.

¹ In Denmark, primary school refers to grades 0 through 9, usually corresponding to ages 6–16.

² According to data from the Danish Ministry of Higher Education and Science, 57% of new students admitted to bachelor programs at Danish universities in 2021 were women (Uddannelses- og forskningsministeriet, 2021), and according to unpublished data from 2020–2021 from the literacy support unit at the university, approximately 70% of students seeking assessment or literacy support at Aarhus University were women.

Table 1 Background information on students tested for dyslexia and control students

Additional tests	Tested for dyslexia			Control groups		Low-frequency word reading	High-frequency word reading
	Red	Yellow	Green	Spelling	ARHQ-R-DA primary		
Number	127	48	64	218	336	220	212
Gender F/M	92/35	32/16	44/22	90/128	203/133	130/90	119/93
Age M (SD)	23 (2.24)	23.67 (2.06)	23.78 (2.50)	22.19 (1.70)	22.2 (2.2)	22.35 (2.12)	22.45 (2.13)
Fields of study	22	12	27	18	148	41	27
Arts	52	19	18	144	63	87	126
Business and Social Sciences	21	7	7	0	37	44	23
Health	32	10	12	56	88	48	36
Science and Technology							
Reading history questionnaire	Tested for dyslexia			Control groups		ARHQ-R-DA current	
Number	43	21	17		336		113
Gender F/M	32/11	13/8	10/7		203/133		69/44
Age M (SD)	22.8 (2.20)	23.48 (1.69)	25.06 (2.43)		22.2 (2.2)		22.23 (1.7)
Fields of study	8	5	6		148		1
Arts	15	7	3		63		27
Business and Social Sciences	10	4	4		37		37
Health	10	5	4		88		48
Science and Technology							

ARHQ-R-DA, Adult Reading History Questionnaire – Revised – Danish adaptation

Materials

The Danish Diagnostic Test (DDT) consists of three subtests. In the first subtest, the student hears a pseudoword and is asked to choose an appropriate spelling from five alternatives. In order to do so, the student is expected to decode the five possible pseudoword spellings and match them with the heard pseudoword. The student has 5 min to complete as many of the 40 items as possible. The second subtest is a phonological choice test in which the student is asked to identify which of five pseudoword sounds like a real word when pronounced. The student has 7 min to complete as many of the 44 items as possible. For each of these two phonological subtests, a standard subscore is automatically calculated as an efficiency score based on accuracy (corrected for guesses) and time (corrected for extremes). An overall standard score is given as the mean of the two subscores. The third subtest is a vocabulary test (giving no subscore), the purpose of which is to identify students whose Danish vocabulary is not sufficient for the result of the phonological subtests to be valid, but because this test is too easy for students in higher education with Danish as a first language, we do not use this test in our analyses.

Spelling was assessed using a dictation test consisting of 30 words. The test was developed at Aarhus University and originally tested with a group of 282 university students with no experienced reading or writing difficulties. Cronbach's α was 0.84. The words were between four and 20 letters ($M=10.2$), one and seven syllables ($M=3.7$) and one and six morphemes ($M=2.9$) long (see Appendix 1). Sixteen words were of foreign origin (12 from French and Latin, two from Greek, one from German and one from English) and include spelling patterns not found in words of Danish origin. Each word was read aloud in the context of a sentence, and the target word was then repeated. The student was asked to spell the target word using pen and paper. Spelling accuracy was scored as the number of correctly spelled words. We chose to include this spelling in the dictation task because the DDT does not target spelling production, which is the task that students are faced with within their studies. In addition, the items of this spelling test mirror the morphological and orthographic complexity frequently found in academic Danish vocabulary.

Low-frequency word reading was assessed using a list of 30-long, morphologically complex, low-frequency words (see Appendix 2); 26 of these were academic words of foreign origin (13 from Latin and French, 12 from Greek and one from English). This test was also developed at Aarhus University for use with university students. The words consisted of seven to 24 letters ($M=11.7$), three to eight syllables ($M=4.4$) and one to six morphemes ($M=2.8$). The student was asked to read the words aloud as accurately and quickly as possible. Readings were recorded and subsequently reading time and accuracy were scored and reading rate (number of words read correctly per ten seconds) calculated. We chose to include this test to measure the word reading accuracy and fluency with words typically found in academic texts.

High-frequency word reading was assessed using a standardized list of 40 short morphologically simple high-frequency words, all of which were of Danish origin (see Appendix 3). This test is a widely used word reading test, developed at the Center for Læseforskning (Elbro, n.d.). The words are between two and nine letters ($M=3.8$), one and three syllables ($M=1.4$) and one and three morphemes ($M=1.2$) long. The student was asked to read the words aloud as accurately and quickly as possible. Readings were recorded, reading time and accuracy scored and reading rate calculated. This test was included to assess word reading fluency with familiar non-academic words.

Adult Reading History Questionnaire – Revised – Danish adaptation (ARHQ-R-DA) was used as a measure of self-reported literacy difficulties. The original ARHQ by Lefly and Pennington (1991) was revised by Parrila et al. (2003). A Danish version was tested on a control group (370 for the primary school scale and 167 for the current status scale) consisting of university students who did not report having literacy difficulties. Questions on the primary school scale focus on reading and spelling acquisition problems when the participants were in primary school, whereas questions of the current status scale focus on current literacy problems. An analysis of internal consistency reliability for each subset of questions led to a reduction of the number of questions, so that the final Danish version consists of 11 questions for each scale, with reliability (Cronbach's α) of 0.93 for the primary school scale and 0.85 for the current status scale. For each question, students were asked to rate themselves on a scale from 0 to 4, with 0 corresponding to no difficulties and 4 corresponding to great difficulties (see Appendix 4). For each scale, a student's score was calculated as the total number of points divided by the maximum number of points possible.

Procedure

All students tested for dyslexia were assessed individually by a literacy counsellor as part of a 1.5-h-long assessment for dyslexia, which also included a background interview. In addition, these students were asked if they would fill out a questionnaire about their difficulties, and if they consented, the questionnaire was subsequently sent to them electronically to complete in their own time.

The control students were tested individually on the two-word reading tests in libraries, cafeterias and other study areas at the university, whereas the spelling test was administered to larger groups of students in auditoriums following lectures. The testing was done by literacy counsellors and student assistants who had been trained to perform the testing.

Results

We compared results from the three additional tests and the two ARHQ-R-DA scales across the four groups (red, yellow, green and control) in order to examine differences between groups. Data for each test was modelled as a multiple linear regression, with test performance predicted by group, age, gender and academic department. Post hoc pairwise comparisons were done on the dependent measure, using Tukey's HSD to calculate adjusted p -values. Beta estimates for the regressions are reported in Table 2, and Table 3 reports significance values from the subsequent post hoc pairwise comparisons. Because different groups of control participants were used for each task, and these control groups were slightly different in sizes; the residual degrees of freedom were also different for each reported regression. For the two-word reading tests, only the reading rate (correct per 10 s) is reported, as there is a considerable ceiling effect for reading accuracy on both tests. The full dataset and R code for conducting all analyses, as well as tables containing results of models with all predictors and all interactions are available at the online repository: https://osf.io/h5vgm/?view_only=c05ed06114e8425eb655c7669a3efa14.

Table 2 Results of multiple linear regression with scores on each of the three tests (spelling, high-frequency word reading rate, and low-frequency word reading rate) predicted by group, age, gender and academic department. Beta estimates for the three diagnostic groups (green, red, and yellow) are shown relative to the scores of control participants, gender is shown relative to female participants and the three Academic Departments are shown relative to Arts. Standard error is shown in parentheses

	Spelling	HF words	LF words
Green	-4.93 *** (0.68)	-4.77 *** (0.63)	-2.48 *** (0.22)
Red	-10.01 *** (0.54)	-8.05 *** (0.49)	-4.10 *** (0.17)
Yellow	-7.77 *** (0.74)	-6.88 *** (0.69)	-3.51 *** (0.24)
Age	0.24* (0.10)	-0.10 (0.09)	0.06 (0.03)
Gender	-1.71 *** (0.44)	0.23 (0.42)	0.37* (0.15)
Business and Social Sciences	0.99 (0.62)	-0.31 (0.55)	-0.07 (0.19)
Health	1.03 (0.91)	1.22 (0.72)	0.42 (0.23)
Science and Technology	-0.58 (0.70)	-1.17 (0.66)	-0.43* (0.21)
<i>N</i>	457	451	459
<i>R</i> ²	0.49	0.44	0.62

****p* < 0.001; ***p* < 0.01; **p* < 0.05

Table 3 Adjusted *p*-values from post hoc pairwise comparison (Tukey HSD) of the groups on the three tests (spelling, low-frequency word reading rate, and high-frequency word reading rate). *p*-values less than 0.05 are highlighted in bold

Test parameter	Groups	Tested for dyslexia		
		Red	Yellow	Green
Spelling	Yellow	<0.05	-	-
	Green	<0.001	<0.01	-
	Control	<0.001	<0.001	<0.001
LF words	Yellow	0.042	-	-
	Green	<0.001	<0.01	-
	Control	<0.001	<0.001	<0.001
HF words	Yellow	0.37	-	-
	Green	<0.001	<0.05	-
	Control	<0.001	<0.001	<0.001
ARHQ-R-DA – primary school	Yellow	0.96	-	-
	Green	0.31	0.22	-
	Control	<0.001	<0.001	<0.001
ARHQ-R-DA – current status	Yellow	0.99	-	-
	Green	0.34	0.40	-
	Control	<0.001	<0.001	<0.001

Spelling

The model predicting performance on the spelling task was significant ($F(8,448) = 54.38, p < 0.01$) with an adjusted R^2 of 0.48. Diagnostic group, age and gender were all

significant predictors of spelling performance, with the diagnostic group as the strongest predictor (see Table 2 for beta estimates). Post hoc pairwise comparisons showed significant differences between all diagnostic groups (Table 3).

Low-frequency word reading

The model predicting performance on the low-frequency word reading task was significant ($F(8,450)=92.39$, $p<0.01$) with an adjusted R^2 of 0.61. The diagnostic group was the strongest predictor (see Table 2 for beta estimates). Pairwise comparisons showed significant differences between all diagnostic groups (Table 3).

High-frequency word reading

The model predicting performance on the high-frequency word reading task was significant ($F(8,442)=43.95$, $p<0.01$) with an adjusted R^2 of 0.43. Only the diagnostic group was a significant predictor of high-frequency word reading (see Table 2 for beta estimates), and pairwise comparisons indicated that all differences between the pairs were significant, except for the yellow and red groups (Table 3).

Adult History Reading Questionnaire – Revised, Danish version (ARHQ-R-DA)

Data from the reading questionnaire's two subscales measuring primary school and current reading experiences were modelled as two separate multiple linear regressions, with diagnostic group, age, gender and academic department as predictor variables.

The model predicting primary school experience was significant ($F(8,522)=24.65$, $p<0.01$) with an adjusted R^2 of 0.26. The pairwise comparisons showed that the control group reported significantly fewer reading acquisition problems than the other three groups ($p<0.001$), while the red, yellow and green groups did not differ significantly from each other.

Similarly, the model predicting the current status scale was significant ($F(8,298)$, $p<0.01$) with an adjusted R^2 of 0.42. As above, the pairwise comparisons showed that the control group reported significantly fewer current literacy problems than the other three groups ($p<0.001$). The red, yellow and green groups did not differ significantly from each other.

Within-group variability in the performances

The boxplots in Fig. 1 show the distribution of results from the four groups for each additional test and the two scales of the ARHQ-R-DA.

While illustrating the overall correlation between the word-level test results and the categories made by the DDT (a correlation analysis showed significant correlations between the classification made by the DDT and results from all the additional tests as well as both ARHQ-R-DA scales; see correlation table in Table 4 of Appendix 5), the boxplots also show that for each test there are a considerable number of green students whose scores are at a level similar to red and yellow students and markedly below the control mean. This is also the case for the ARHQ-R-DA data, with the boxplot showing that very few individuals in the green group score at a level comparable to that of the controls.

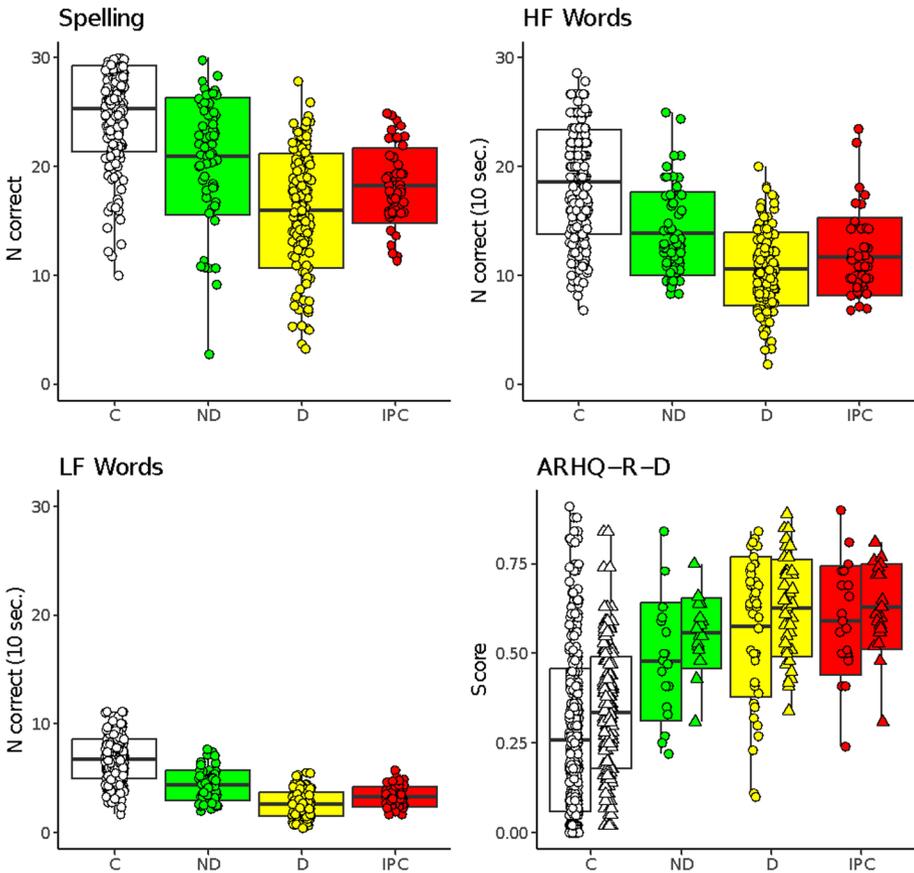


Fig. 1 Boxplots of distributions for test and questionnaire results. Note: HF Words = high-frequency word reading rate; LF = low-frequency word reading rate. For all plots, each point represents 1 participant’s score. Colors represent the categories assigned by the Danish diagnostic test: green/ND (= ‘not dyslexic’), yellow/IPC (‘inaccurate phonological coding’), and red/D (‘dyslexic’). Control participants’ (C) scores are represented with a white box. The middle line in each box represents the group’s mean, and the top and bottom of each box represents 1 standard deviation from the mean. The “whiskers” show the range of the data (highest and lowest scores in each group). Points have been jittered slightly to avoid overlap. In the bottom right figure (ARHQ-R-DA), circles indicate elementary school questions and triangles indicate current status questions

We then examined the number of students in each DDT category whose spelling and word reading rate scores were more than one standard deviation below the mean of the control group (henceforward a low score). This analysis was done in order to examine whether the low performers in each group were the same individuals across the three additional tests or whether these individuals merely performed poorly on a single test. This analysis is illustrated in Fig. 2.

Of the 64 students classified as green by the DDT, 88% obtained a low score on at least one of the additional tests: 27% on only one additional test, 35% on two additional tests, and 26% on all three additional tests. In the group of students classified as yellow, all students obtained a low score on at least one additional test: 4% on only one additional test, 42% on two additional tests and 54% on all three additional tests. Of the students

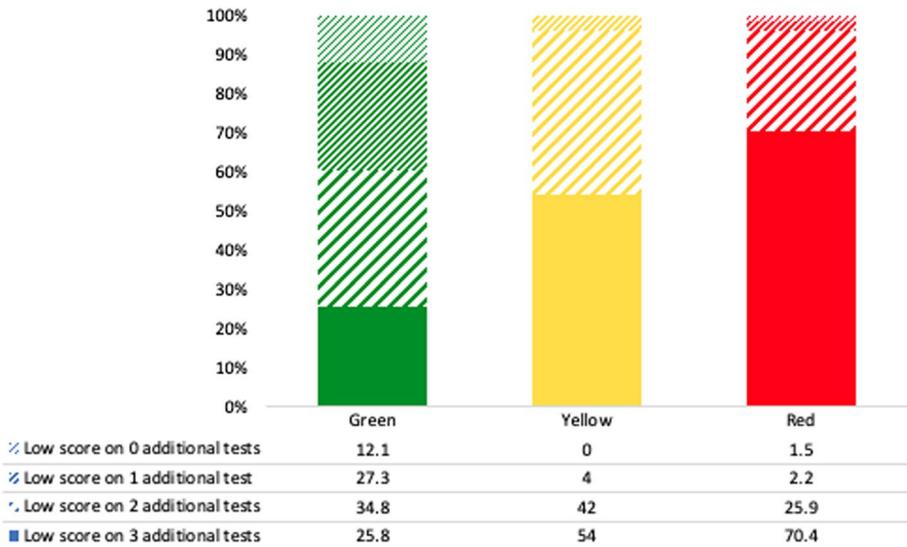


Fig. 2 Distribution of performances of the three groups from the DDT on additional tests. Each column shows the percentages of students in each of the three groups who had a low score on all three additional tests (full colour), two additional tests, one additional test and no additional tests respectively. Numbers in the rows below are the percentages

classified as red, 98% obtained a low score on at least one additional test: 2% on only one additional test, 26% on two additional tests and 70% on all three additional tests. Ideally, the same calculation should be done for the low-performing individuals in the control group. However, as mentioned in the “**Method**” section, each control participant only completed one test, rendering the comparison of performances across tests impossible.

Discussion

The purpose of this study was to investigate whether it is sufficient to test university students for dyslexia only by means of the phonology-based DDT that currently is the only accepted diagnostic tool in Denmark. We did this by comparing how university students from the three diagnostic categories from the DDT performed on three additional word-level tests—reading high-frequency words, reading low-frequency words and spelling to dictation—compared to university students who reported no literacy difficulties. While we found an overall correlation between the groups based on the DDT scores and performance on the additional tests, the results also reveal that students who self-report literacy difficulties, but are categorized as green by the DDT and therefore do not qualify for support, perform significantly poorer than a control group on all tests that did not show a significant ceiling effect.

In **spelling**, all three groups tested for dyslexia were significantly different from the control group. Given that Danish has an alphabetic orthography and that spelling (despite a deep orthography) thus builds on phonology, it makes sense that there would be a correlation between performance on phonological tests and a spelling test. This being said, it is of

interest that the group of green students spell significantly worse than the control group. It is worth noting that our spelling test did not control the time used on spelling correctly—a problem often articulated by these students at our assessment meetings and supported by Parrila and Protopapas (2017) who point out that extraordinary effort in completing a task should also be taken into consideration when diagnosing dyslexia.

The pattern on low-frequency word reading rate was the same as in spelling, with all three groups tested for dyslexia and the control group performing significantly different from each other. On the high-frequency word reading test, which intuitively should be the easiest of the three tests as it includes only short common words, the only significant difference was between the red and the yellow groups. It is somewhat surprising that students seeking assessment but scoring too high on the DDT to be diagnosed with dyslexia were significantly slower than controls at reading very simple words such as *pil* (“arrow”) and *du* (“you”) that should be accessed automatically.

The reading history questionnaire results show clearly that most university students seeking assessment for dyslexia at the university level have experienced and continue to experience high levels of difficulty with reading and writing, regardless of how they perform on the phonological measures in the DDT. Thus, there was no difference between students categorized as red, yellow or green, neither in terms of the difficulties they experienced in primary school nor in terms of their current difficulties, a finding in line with Parrila et al. (2007) and Deacon et al. (2012) who showed that self-report of literacy difficulties in itself is often a strong indicator of dyslexia.

A closer examination of the individual performances across the three additional tests showed that the vast majority of students classified by the DDT as either red or yellow (96% for both groups) scored more than one standard deviation below the control group mean on at least two out of three of the additional tests. Thus, our results suggest that low performance on word-level tests is a characteristic trait of university students who are given a diagnosis of dyslexia and who are thus eligible for special educational support in their studies. Interestingly, this characteristic is also true for 61% of students who experience literacy difficulties, but are classified by the Danish diagnostic test as green and, as a result, do not qualify for support.

If the Danish diagnostic test is sufficiently sensitive in measuring phonological difficulties at the university level, it would seem that a phonological deficit cannot explain the word-level difficulties we have observed in a large proportion of the green students. The low scores on the high-frequency and low-frequency word reading tests of the green students were due to a slow reading rate, not to accuracy problems. This finding might suggest that these students have a naming deficit (in line with Wolf & Bowers, 1999). Thus, it is possible that the green students with word reading difficulties would have performed poorly on tests of rapid automatized naming, but because this study did not include such tests, this remains speculative. A naming deficit could also provide an explanation for spelling difficulties (see e.g. Stainthorp et al., 2013) and we know that naming and spelling deficits are both prevalent in university students with dyslexia (e.g. Kemp et al., 2009). However, this suggestion is tentative at best given that multiple studies have failed to find an association between naming speed and orthographic processing (e.g. Georgiou et al., 2016).

The green students who performed poorly on either low-frequency word reading or spelling of morphologically complex words, both of which include a high proportion of words of foreign origin, possibly fit the description of the subtype of “surface dyslexia” proposed by Castles and Coltheart (1993) and characterized by a deficit in reading irregularly spelled words. Recognizing these long and uncommon words as a whole would be

difficult as many of the words are unfamiliar to many students. In that case, the words would have to be read either by means of phonological decoding or by recognizing morphemes or other orthographic chunks. If one manifestation of developmental surface dyslexia is poorer learning of orthographic patterns (e.g. Castles & Holmes, 1996; Wang et al., 2014), students with surface dyslexia would find the reading of these morphologically complex words with spelling patterns that do not follow Danish orthographic rules challenging. However, this interpretation is also tentative as the empirical evidence for this subtype is not strong.

If a double- or multiple-deficit model explanation is at work here, one might object that these students should have performed at a lower level on the grapheme-phoneme conversion tasks in the DDT. While this seems to be a valid point, there are counterarguments to be made. First, even if a phonological deficit were the explanation for dyslexia, the DDT targets this rather narrowly, building on the definition of dyslexia as a difficulty “connecting letters and letter sequences to their normal sounds” (Center for Læseforskning, 2020, our translation). Including different kinds of phonological tests might reveal phonological difficulties for students who score above the cutoff limit in the DDT. There is a possibility that the DDT, while doing an overall good job at identifying a large proportion of students with dyslexia, is too specific, leaving a group of university students with clear literacy difficulties as false negatives.

A second point to be made is the level of compensatory skills which this population of university students with literacy difficulties have (cf. Bønnerup et al., 2019; Parrila et al., 2007); the DDT might simply be too easy for some of them. It is conceivable that some of our green students, with extraordinary effort, are able to compensate enough for an underlying phonological deficit to score above the cutoff when focusing on the grapheme-phoneme conversion tasks, e.g. by relying on pattern recognition, but when faced with the everyday literacy challenges of reading complex academic words in complex academic texts or spelling complex words in challenging written assignments, their cognitive resources fall short.

A notable limitation of our study is the fact that we were not able to test our control group with the DDT that defined our other three groups. We cannot rule out that some of our control participants could have potentially qualified as red or yellow despite initially answering “no” to a question about having ever experienced literacy difficulties. In addition to this, the green group is a heterogeneous group of students who seek an explanation for their experienced difficulties. Thus, this group likely represents both students with specific difficulties with reading and/or writing, as well as students whose issues are attributable to other diagnoses, such as ADD, anxiety or depression. It is conceivable that an exclusion of participants with other diagnoses would have allowed for a more homogeneous group of green students, who might be even more like the groups of yellow and red students in terms of their performance on our word-level tests. We also acknowledge that our focus on word reading and spelling may have excluded students with other kinds of literacy difficulties, such as reading comprehension or text writing difficulties. It is important to broaden the scope to identify students with literacy difficulties unrelated to word reading. However, as dyslexia is currently the only officially recognized reading impairment in Denmark, it seems paramount to us to start with addressing the issue of diagnosing this impairment appropriately.

An additional limitation is the fact that we have different control groups for each test, and we urge caution in the interpretation of the results. This shortcoming means, for example that further analyses of low-performing individuals in the control group are not possible. Future studies investigating the low performers in the group of university students who do not perceive themselves as having difficulties with reading or writing are needed in order to explore what characterizes this group in comparison to the group of green students. The DDT is not available for control participants, but it may be useful to use a test similar to the DDT to assess their phonological impairments as well.

Finally, different contexts for data collection for the control group and the groups tested for dyslexia may have played a role in their performance, with controls trying to read words as quickly as possible in a public space, and students tested for dyslexia reading the lists in a clinical setting. However, if the control students forced themselves to read at too high a pace, they would presumably have been more prone to making reading errors, which would be reflected in the reading rate score that is a combined measure of accuracy and speed. It is more likely that the non-optimal testing conditions for control participants resulted in a conservative estimate of their performance and thus may have reduced the differences between them and the students tested for dyslexia.

Future studies should include tests targeting other areas of literacy than word reading and spelling, including RAN, phonological awareness, morphological awareness, orthographic knowledge, reading comprehension and writing, in order to obtain a better understanding of the difficulties that underlie the green students' word reading and spelling problems. Also, looking closer at the qualitative side of results and analyzing the types of errors made in the respective groups could help to better understand whether the difficulties experienced by our green students are in fact phonologically based.

Our findings have some important implications for research with university students in general and for Danish dyslexia diagnosis in particular. Despite the fact that Danish is a relatively opaque orthography, word reading accuracy tasks were at the ceiling. While the use of speeded tasks is standard practice in consistent orthographies, this is not always the case in more opaque orthographies even with university students as participants (e.g. Deacon et al., 2006; Savill & Thierry, 2012). Our results suggest that if the sample is identified as having dyslexia on the basis of phonological coding and word reading accuracy measures, it is unlikely to be representative of all students with dyslexia that would be identified if slow and effortful word reading is used as a criterion instead of inaccurate reading. Instead, it is likely that accuracy-based sample selection will lead to biasing the sample towards individuals with more severe word reading problems, and will potentially increase the frequency of co-morbid deficits that often are of interest in these studies.

In terms of diagnosing Danish students with dyslexia, the findings of this study suggest that a few simple tests of word reading speed and spelling accuracy would add valuable information to the assessment of university students who report literacy difficulties. An easy way to find students with word-level difficulties who are categorized as green by the DDT would be to require a low score (e.g. one SD below the control mean) on at least two out of three tests. Even though it is possible that a few students would have severe difficulties only in one area—for example in the reading of academic words or spelling—this precaution would rule out students who performed poorly on just one test due to factors that may be unrelated to word-level problems, such as nervousness. While the Danish system of special educational support is of a very high standard internationally, our study suggests that the current state-mandated method of diagnosing dyslexia denies a significant number of university students with verifiable literacy difficulties access to the support. The Danish diagnostic test appears to be quite sensitive to the definition of dyslexia it is based on, but if we adopt a definition more in line with the internationally acknowledged definitions formulated by IDA and APA, the test seems to be too specific, leaving out dyslexic students whose word-level problems are either attributable to another deficit (such as a naming deficit) or who are too accomplished in compensating for their difficulties to be identified by the current phonological tasks. With a view to international definitions of dyslexia, international guidelines for the assessment for dyslexia, and substantial empirical evidence, a dyslexia diagnosis should not be the product of a single standard score based on a test assessing the knowledge of correspondences between phonemes and graphemes. Instead, it should be based on a clinical evaluation of a

combination of background information and results from a variety of tests in order to minimize the risk of overlooking a significant number of students with dyslexia in need of literacy support. A more fine-grained assessment is critical for an informed distribution of literacy support resources.

Appendix 1 Words in spelling test, 30 items

fedtet.

flypassagererne.

uciviliserede.

transporterede.

torskerogn.

lærerne.

erhvervslivet.

konkurrence.

kontrolleret.

beskadiget.

forargede.

ingeniører.

niveauet.

blokering.

torpederede.

observation.

konklusionsafsnittet.

ordentlige.

hierarki.

nødt.

øjenkrogen.

fedt.

schæferhund.

passager.

erindrer.

fysioterapeut.

buschaufføren.

erhverv.

bevidstløshed.

forarger.

Appendix 2 Low-frequency word list, 30 items

Irrelevans.

Eksplicit.

Supplement.

Delfortolkning.

Legitim.

Negation.

Observation.

Kriterier.
Formidlende.
Efterfølgeligt.
Processering.
Argumentatorisk.
Kvalitativt.
Kompilation.
Reservoir.
Centrerede.
Karakteristika.
Subjektsprædikat.
Værdiopfattelser.
Komplementaritetsprincip.
Oligarki.
Ionisere.
Traumatisk.
Gymnasiast.
Homogeniseret.
Hermeneutik.
Epistemologi.
Apoplektisk.
Metabolisme.
Allegoriseringer.

Appendix 3 High-frequency word list, 40 items

læ ti.
dusø.
ur at.
piltør.
sætjul.
køn vin.
træ abe.
fri glo.
øje stå.
håbe huse.
løbe gave.
dyne dåse.
birk aske.
skud bold.
gladblød.
æble gråt.
kistekasse.
trøst smagt.
strøm straf.
kærlighed kedeligt.

Appendix 4 Adult Reading History Questionnaire – Revised, Danish version (English translation)

Elementary school questions

1. How much difficulty did you have learning to read in elementary school? (0 = none, 4 = a great deal)
2. How much extra help did you need when learning to read in elementary school? (0 = no help, 4 = a great deal)
3. Did you ever reverse the order of letters or numbers when you were a child? (0 = no, 4 = a great deal)
4. Did you have difficulty learning letter and/or colour names when you were a child? (0 = no, 4 = a great deal)
5. How would you compare your reading skill to that of others in your elementary classes? (0 = above average, 4 = below average)
6. All students struggle from time to time in elementary school. In comparison to your classmates, how much did you struggle to complete your work? (0 = not at all, 4 = much more than most)
7. Which of the following most nearly describes your attitude toward reading as a child? (0 = very positive, 4 = very negative)
8. When you were in elementary school, how much reading did you do for pleasure? (0 = a great deal, 4 = none)
9. How would you compare your reading speed in elementary school with that of your classmates? (0 = above average, 4 = below average)
10. How much difficulty did you have learning to spell in elementary school? (0 = none, 4 = a great deal)
11. When you were in elementary school, how many books did you read for pleasure each year? (0 = more than 10, 4 = none)

Current status questions

1. How much difficulty do you currently have with reading? (0 = none, 4 = a great deal)
2. Have you received extra help during your postsecondary education? (0 = no help, 4 = tutors)
3. How would you compare your reading skill to that of others in your postsecondary classes? (0 = above average, 4 = below average)
4. All students struggle from time to time at the postsecondary level. In comparison to your classmates, how much do your struggle to complete your work? (0 = not at all, 4 = much more than most)
5. Have you experienced difficulty reading or writing in academic Danish? (0 = none, 4 = a great deal)
6. What is your current attitude toward reading? (0 = very positive, 4 = very negative)
7. How much reading do you do for pleasure? (0 = a great deal, 4 = none)
8. How would you compare your current reading speed with that of others with the same age and education? (0 = above average, 4 = below average)
9. How would you compare your current spelling to that of others with the same age and education? (0 = above average, 4 = below average)
10. Do you currently reverse the order of letters or numbers when you read or write? (0 = no, 4 = a great deal)
11. How many books do you read for pleasure each year? (0 = more than 10, 4 = none)

Appendix 5

Table 4 Correlations (Spearman's rho) for study variables

Variable	1	2	3	4	5	6	7	8
1. Classification on the Dyslexia test	-	-	-	-	-	-	-	-
2. Spelling (correct)	-.685***	-	-	-	-	-	-	-
3. High-frequency word reading accuracy	-.342***	.484***	-	-	-	-	-	-
4. High-frequency word reading rate	-.671***	.278***	.263***	-	-	-	-	-
5. Low-frequency word reading accuracy	-.527***	.567***	.431***	.111	-	-	-	-
6. Low-frequency word reading rate	-.796***	.616***	.417***	.423***	.726***	-	-	-
7. ARHQ-R-DA—elementary school	.502***	-.432***	-.354***	-.274**	-.144	-.392***	-	-
8. ARHQ-R-DA—current status	.680***	-.145	-.109	-.197	.001	-.182	.721***	-

* $p < .05$, ** $p < .01$, *** $p < .001$

Acknowledgements The authors thank colleagues and student assistants in the Literacy Counselling Team at the Counselling and Support Centre of Aarhus University for the data collection and Anne Leth Pedersen for valuable feedback on early drafts of the paper. We also thank the reviewers for their comments and questions which helped us improve the paper.

Author contribution All authors contributed to the study conception and design. Material preparation was performed by Katrine H. Bønnerup, Helle Fredslund Ottosen and colleagues at the Counselling and Support Centre of Aarhus University under the supervision of Rauno Parrila. Data collection was performed by Katrine H. Bønnerup, Helle Fredslund Ottosen, colleagues and student assistants at the Counselling and Support Centre of Aarhus University. Analyses were performed by Ethan Weed and Helle Fredslund Ottosen. The first draft of the manuscript was written by Helle Fredslund Ottosen and Katrine H. Bønnerup, and all authors commented on and edited previous versions of the manuscript. All authors read and approved the final manuscript.

Funding The research was funded by our respective universities.

Availability of data and material All data are available online at https://osf.io/h5vgm/?view_only=c05ed06114e8425eb655c7669a3efa14.

Code availability Not applicable.

Declarations

Ethics approval We declare that we complied with the university guidelines that were in place in 2015–2019 when we collected the data.

Consent to participate All participants gave consent to participate in the study. The participants with literacy difficulties gave written consent, and the control participants, who completed only one task and provided no identifying information, gave oral consent.

Consent for publication All authors consented to submit the paper for publication in *Annals of Dyslexia*.

Conflict of interest The authors declare no competing interests.

References

- Aakermann, T., Bergseth, G., Berntsen, H., Dale, L., Helland, T., Høyen-Tengesdal, I., Lie, R., Lyster, S.-A. H., Melby-Lervåg, M., & Nygård, T. (2017). *Faglige retningslinjer for kartlegging, utredning og oppfølging av elever med dysleksi* [Professional guidelines for mapping, assessment and follow-up of students with dyslexia]. Dysleksi Norge.
- Aarhus University. (2018). *AU i tal 2018* [AU in numbers 2018]. Retrieved 25 November 2020 from https://www.au.dk/fileadmin/www.au.dk/AU_i_tal/AU_i_tal_2018/AU_i_tal_2018_DK_Nov_2019.pdf
- American Psychiatric Association (APA). (2013). *Diagnostic and statistical manual of mental disorders (DSM-5(R))*. American Psychiatric Pub.
- Bønnerup, K. H., Pedersen, A. L., Weed, E., & Parrila, R. (2019). Differences in the literacy skills of Danish dyslexic students in two types of higher education programmes. *Dyslexia*, 25(2), 173–189. <https://doi.org/10.1002/dys.1617>
- Broggi, M. J., Ready, R. E., & Moore, D. L. (2019). Screening for reading disability in university students with phonological processing and working memory tasks. *Dyslexia*, 25, 256–266. <https://doi.org/10.1002/dys.1625>
- Castles, A., & Coltheart, M. (1993). Varieties of developmental dyslexia. *Cognition*, 47(2), 149–180. [https://doi.org/10.1016/0010-0277\(93\)90003-E](https://doi.org/10.1016/0010-0277(93)90003-E)
- Castles, A., & Holmes, V. M. (1996). Subtypes of developmental dyslexia and lexical acquisition. *Australian Journal of Psychology*, 48, 130–135. <https://doi.org/10.1080/00049539608259519>
- Cavalli, E., Casalis, S., El Ahmadi, A., Zira, M., Poracchia-George, F., & Colé, P. (2016). Vocabulary skills are well developed in university students with dyslexia: Evidence from multiple case studies. *Research in Developmental Disabilities*, 51–52, 89–102. <https://doi.org/10.1016/j.ridd.2016.01.006>

- Center for Læseforskning, Institut for Uddannelse og Pædagogik, Socialstyrelsen, & Undervisningsministeriet. (2020). *Vejledning til Ordblindetesten (version 8)* [Guidelines for the Dyslexia test (8th ed.)]. Copenhagen: Børne- og Undervisningsministeriet, Styrelsen for Undervisning og Kvalitet.
- Danmarks Statistik. (2020a). *Korte videregående uddannelser* [Short-cycle higher education]. Retrieved 25 November 2020 at <https://www.dst.dk/da/Statistik/emner/uddannelse-og-viden/fuldtidsuddannelser/korte-videregaende-uddannelser>.
- Danmarks Statistik. (2020b). *Mellemlange videregående uddannelser* [Medium-cycle higher education]. Retrieved 25 November 2020 at <https://www.dst.dk/da/Statistik/emner/uddannelse-og-viden/fuldtidsuddannelser/mellemlange-videregaende-uddannelser>.
- Danmarks Statistik. (2020c). *Bacheloruddannelser* [Bachelor's programs]. Retrieved 25 November 2020 at <https://www.dst.dk/da/Statistik/emner/uddannelse-og-viden/fuldtidsuddannelser/bacheloruddannelser>.
- Danmarks Statistik. (2020d). *Lange videregående uddannelser* [Long-cycle higher education]. Retrieved 25 November 2020 at <https://www.dst.dk/da/Statistik/emner/uddannelse-og-viden/fuldtidsuddannelser/lange-videregaende-uddannelser>.
- Deacon, S. H., Parrila, R., & Kirby, J. (2006). Processing of derived forms in high-functioning dyslexics. *Annals of Dyslexia*, 56, 103–128. <https://doi.org/10.1007/s11881-006-0005-3>
- Deacon, S. H., Cook, K., & Parrila, R. (2012). Identifying high-functioning dyslexics: Is self-report of early reading problems enough? *Annals of Dyslexia*, 44, 205–226. <https://doi.org/10.1007/s11881-012-0068-2>
- Elbro, C. (n.d.). *Om "Elbros ordlister"* [About "Elbro's word lists"]. Retrieved 25 November 2020 from https://laes.hum.ku.dk/test/elbros_ordlister/om_ordlisterne/.
- Engh, G., Frylmark, A., Jacobsson, C., Näreskog, M., & Salo, G. (2017). *Modell för utredning av läs- och skrivsvårigheter* [Model for assessment of reading and writing difficulties]. Stockholm: Svenska Dyslexiföreningen.
- Erskine, J. M., & Seymour, P. H. K. (2005). Proximal analysis of developmental dyslexia in adulthood: The cognitive mosaic model. *Journal of Educational Psychology*, 97(3), 406–424. <https://doi.org/10.1348/000712603321661859>
- Frith, U. (1999). Paradoxes in the definition of dyslexia. *Dyslexia*, 5, 192–214. [https://doi.org/10.1002/\(SICI\)1099-0909\(199912\)5:4%3C192::AID-DYS144%3E3.0.CO;2-N](https://doi.org/10.1002/(SICI)1099-0909(199912)5:4%3C192::AID-DYS144%3E3.0.CO;2-N)
- Georgiou, G., Aro, M., Liao, C.-H., & Parrila, R. (2016). Modeling the relationship between rapid automatized naming and literacy skills across languages varying in orthographic consistency. *Journal of Experimental Child Psychology*, 143, 48–64. <https://doi.org/10.1016/j.jecp.2015.10.017>
- Hatcher, J., Snowling, M. J., & Griffiths, Y. M. (2002). Cognitive assessment of dyslexic students in higher education. *British Journal of Educational Psychology*, 72, 119–133. <https://doi.org/10.1348/000709902158801>
- Hebert, M., Zhang, X., & Parrila, R. (2018). Examining reading comprehension text and question answering time differences in university students with and without a history of reading difficulties. *Annals of Dyslexia*, 68(1), 15–24. <https://doi.org/10.1007/s11881-017-0153-7>
- International Dyslexia Association (IDA). (2017). *Dyslexia assessment: What is it and how can it help?* Baltimore: Author. Retrieved 25 November 2020 from <https://dyslexiaida.org/dyslexia-assessment-what-is-it-and-how-can-it-help/>.
- Kemp, N., Parrila, R. K., & Kirby, J. R. (2009). Phonological and orthographic spelling in high-functioning adult dyslexics. *Dyslexia*, 15, 105–128. <https://doi.org/10.1002/dys.364>
- Kirby, J. R., Parrila, R. K., & Pfeiffer, S. L. (2003). Naming speed and phonological awareness as predictors of reading development. *Journal of Educational Psychology*, 95(3), 453–464. <https://doi.org/10.1037/0022-0663.95.3.453>
- Kirby, J. R., Georgiou, G. K., Martinussen, R., & Parrila, R. (2010). Naming speed and reading: From prediction to instruction. *Reading Research Quarterly*, 45(3), 341–362. <https://doi.org/10.1598/RRQ.45.3.4>
- Lefly, D. L., & Pennington, B. F. (1991). Spelling errors and reading fluency in compensated adult dyslexics. *Annals of Dyslexia*, 41, 143–162. <https://doi.org/10.1007/BF02648083>
- Lyon, G. R., Shaywitz, S. E., & Shaywitz, B. A. (2003). A definition of dyslexia. *Annals of Dyslexia*, 53(1), 1–14. <https://doi.org/10.1007/s11881-003-0001-9>
- Manis, F. R., Seidenberg, M. S., Doi, L. M., McBride-Chang, C., & Petersen, A. (1996). On the bases of two subtypes of developmental dyslexia. *Cognition*, 58(2), 157–195. [https://doi.org/10.1016/0010-0277\(95\)00679-6](https://doi.org/10.1016/0010-0277(95)00679-6)
- Manis, F. R., Seidenberg, M. S., Stallings, L., Joanisse, M., Bailey, C., Freedman, L., Curtin, S., & Keating, P. (1999). Development of dyslexic subgroups: A one-year follow up. *Annals of Dyslexia*, 49, 105–134. <https://doi.org/10.1007/s11881-999-0021-1>
- Mather, N., White, J., & Youman, M. (2020). Dyslexia around the world: A snapshot. *Learning Disabilities: A Multidisciplinary Journal*, 25, 1–17. <https://doi.org/10.18666/LDMJ-2020-V25-11-9552>
- Møller, H. L., Arnbak, E., Petersen, D. K., Poulsen, M., Juul, H., & Elbro, C. (2014). *Teknisk rapport om Ordblindetesten* [Technical report on the Dyslexia Test]. Copenhagen: Center for Reading Research,

- University of Copenhagen, & Danish School of Education, Aarhus University. Retrieved 25 November 2020 from <https://ordblindetest.nu/vejleder/pdf/TekniskRapportOrdblindetesten.pdf>
- Parrila, R. K., & Protopapas, A. (2017). Dyslexia and word reading problems. In K. Cain, D. L. Compton, & R. K. Parrila (Eds.), *Theories of reading development* (pp. 333–358). John Benjamins Publishing.
- Parrila, R., Georgiou, G., & Corkett, J. (2007). University Students with a significant history of reading difficulties: What is and is not compensated? *Exceptionality Education Canada*, 17(2), 195–220. <https://doi.org/10.5206/eei.v17i2.7604>
- Parrila, R., Corkett, J., Kirby, J., & Hein, S. (2003). *Adult Reading History Questionnaire-Revised*. (Unpublished questionnaire). University of Alberta.
- Pennington, B. F. (2006). From single to multiple deficit models of developmental disorders. *Cognition*, 101, 385–413. <https://doi.org/10.1016/j.cognition.2006.04.008>
- Pennington, B. F., Santerre-Lemmon, L., Rosenberg, J., MacDonald, B., Boada, R., Friend, A., Leopold, D., Samuelsson, S., Byrne, B., Willcutt, E. G., & Olson, R. K. (2012). Individual prediction of dyslexia by single vs. multiple deficit models. *Journal of Abnormal Psychology*, 121(1), 212–224. doi:<https://doi.org/10.1037/a0025823>
- Peterson, R. L., Pennington, B. F., & Olson, R. K. (2013). Subtypes of developmental dyslexia: Testing the predictions of the dual-route and connectionist frameworks. *Cognition*, 126(1), 20–38. <https://doi.org/10.1016/j.cognition.2012.08.007>
- Peterson, R. L., Pennington, B. F., Olson, R. K., & Wadsworth, S. J. (2014). Longitudinal stability of phonological and surface subtypes of developmental dyslexia. *Scientific Studies of Reading*, 18(5), 347–362. <https://doi.org/10.1080/10888438.2014.904870>
- Poulsen, M., Elbro, C., Møller, H. L., Juul, H., Petersen, D. K., & Arnbak, E. (2016). *The Danish Dyslexia test: Validity of a wide-range, web-based test for dyslexia*. (Conference paper). Retrieved 25 November 2020 from https://curis.ku.dk/ws/files/172140939/widerangedyslexiatest_bda_2016.pdf.
- Re, A. M., Tressoldi, P. E., Cornoldi, C., & Lucangeli, D. (2011). Which tasks best discriminate between dyslexic students and control in a transparent language? *Dyslexia*, 17(3), 227–241. <https://doi.org/10.1002/dys.431>
- Specific Learning Difficulties Assessment Standards Committee (SASC). (2005). SpLD Working Group 2005/DfES Guidelines. Retrieved 1 December 2020 from https://sasc.org.uk/SASCDocuments/SpLD_Working_Group_2005-DfES_Guidelines.pdf
- Savill, N. J., & Thierry, G. (2012). Decoding ability makes waves in reading: Deficient interactions between attention and phonological analysis in developmental dyslexia. *Neuropsychologia*, 50(7), 1553–1564. <https://doi.org/10.1016/j.neuropsychologia.2012.03.008>
- Snowling, M. J. (2000). *Dyslexia* (2nd ed.). Blackwell.
- Snowling, M. J., Nation, K., Moxham, P., Gallagher, A., & Frith, U. (1997). Phonological processing of dyslexic students in higher education: A preliminary report. *Journal of Research in Reading*, 20(1), 33–41. <https://doi.org/10.1111/1467-9817.00018>
- Stainthorp, R., Powell, D., & Stuart, M. (2013). The relationship between rapid naming and word spelling in English. *Journal of Research in Reading*, 36, 371–388.
- Stanovich, K. E., Siegel, L. S., & Gottardo, A. (1997). Converging evidence for phonological and surface subtypes of reading disability. *Journal of Educational Psychology*, 89, 114–127. <https://doi.org/10.1037/0022-0663.89.1.114>
- Tops, W., Callens, M., Lammertyn, J., Van Hees, V., & Brysbaert, M. (2012). Identifying students with dyslexia in higher education. *Annals of Dyslexia*, 62(3), 186–203. <https://doi.org/10.1007/s11881-012-0072-6>
- Uddannelses- og forskningsministeriet [Ministry of Higher Education and Science]. (2021). *Ansøgere og optagne fordelt på køn, alder og adgangsgrundlag* [Applicants and admissions by gender, age and admission basis]. Retrieved 17 October 2021 from <https://ufm.dk/uddannelse/statistik-og-analyser/sogning-og-optag-pa-videregaende-uddannelser/grundtal-om-sogning-og-optag/ansogere-og-optagne-fordelt-pa-kon-alder-og-adgangsgrundlag>
- Vellutino, F. R., & Fletcher, J. M. (2005). Developmental dyslexia. In M. J. Snowling & C. Hulme (Eds.), *The science of reading: A handbook* (pp. 362–378). Blackwell.
- Wang, H. C., Marinus, E., Nickels, L., & Castles, A. (2014). Tracking orthographic learning in children with different profiles of reading difficulty. *Frontiers in Human Neuroscience*, 8, 468. <https://doi.org/10.3389/fnhum.2014.00468>
- Wolf, M., & Bowers, P. G. (1999). The double deficit hypothesis for the developmental dyslexias. *Journal of Educational Psychology*, 91, 415–438. <https://doi.org/10.1037/0022-0663.91.3.415>