UNIVERSITY OF AMSTERDAM

UvA-DARE (Digital Academic Repository)

Six factors of adult dyslexia assesed by cognitive tests and self-report questions: Very high predictive validity

Tamboer, P.; Vorst, H.C.M.; de Jong, P.F.

DOI
10.1016/j.ridd.2017.09.010

Publication date
2017

Document Version
Final published version

Published in
Research in Developmental Disabilities

License
Article 25fa Dutch Copyright Act (https://www.openaccess.nl/en/policies/open-access-in-dutch-copyright-law-taverne-amendment)

Link to publication

Citation for published version (APA):
Tamboer, P., Vorst, H. C. M., & de Jong, P. F. (2017). Six factors of adult dyslexia assesed by cognitive tests and self-report questions: Very high predictive validity. *Research in Developmental Disabilities*, *71*, 143-168. https://doi.org/10.1016/j.ridd.2017.09.010

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, P.O. Box 19185, 1000 GD Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (https://dare.uva.nl)

Download date:21 Apr 2026

Research in Developmental Disabilities 71 (2017) 143–168

Contents lists available at ScienceDirect

Research in Developmental Disabilities

journal homepage: www.elsevier.com/locate/redevdis

Six factors of adult dyslexia assessed by cognitive tests and self-report questions: Very high predictive validity

Peter Tamboer*, Harrie C.M. Vorst, Peter F. de Jong

University of Amsterdam, Faculty of Social and Behavioural Sciences, Nieuwe Achtergracht 127, 1018WS, Amsterdam, The Netherlands

ARTICLE INFO

Number of reviews completed is 2

Keywords:
Dyslexia
Digital assessment
Self-report questions
Discriminant analysis
Principal components analysis

ABSTRACT

The Multiple Diagnostic Digital Dyslexia Test for Adults (MDDDT-A) consists of 12 newly developed tests and self-report questions in the Dutch language. Predictive validity and construct validity were investigated and compared with validity of a standard test battery of dyslexia (STB) in a sample of 154 students.

There are three main results. First, various analyses of principal components showed that six or more factors of dyslexia can be distinguished (rapid naming, spelling, reading, short-term memory, confusion, phonology, attention, complexity). All factors are represented by the MDDDT-A. Second, various discriminant analyses showed good predictive validity for both the tests of the MDDDT-A (90%) and the STB (90%). However, predictive validity of the questionnaire was highest (97%). Third, we analysed the best predictors of dyslexia and found that predictive validity is higher when construct validity is high, that is when a set of predictors represents many characteristics of dyslexia.

The main conclusion is that a digital test battery can be a reliable screening instrument for dyslexia in students, especially when it is accompanied by self-report questions. A theoretical conclusion is that dyslexia is characterized by at least six cognitive impairments in a complex way. In students, this structure may be modulated by high intelligence and good schooling through various compensation strategies. It is therefore recommended to include assessments of all characteristics of dyslexia to achieve the most reliable diagnoses in different samples and in different countries.

1. Introduction

In the Netherlands, various methods for diagnosing dyslexia have been developed for children. Reliable methods for diagnosing dyslexia in adults are however sparse and expensive. Although many young adults with dyslexia have been tested during their school days, an unknown but probably substantial portion of them was never tested at all. Of older adults with dyslexia, a majority was never tested because when they were young, diagnosing dyslexia at school was not a widespread practice like it is nowadays. This resulted in the present situation where there are still many adult people who are not aware of having dyslexia. They may experience various difficulties at work while they do not know that dyslexia may be the underlying cause. But even when they come up with the idea of having themselves tested, prohibitive costs may prevent them following through on it.

Worldwide, most traditional diagnostic instruments are based on definitions of dyslexia, such as the one provided by the World Health Organization (2010). Dyslexia is generally defined as a specific reading disorder characterized by a specific and significant impairment in the development of reading skills, which is not due to problems with visual acuity, schooling, or overall mental

* Corresponding author at: Overtoom 247B, 1054HW, Amsterdam, The Netherlands.
E-mail addresses: petertamboer@kpnmail.nl (P. Tamboer), H.C.M.Vorst@uva.nl (H.C.M. Vorst), P.F.deJong@uva.nl (P.F. de Jong).

http://dx.doi.org/10.1016/j.ridd.2017.09.010
Received 17 October 2016; Received in revised form 5 September 2017; Accepted 16 September 2017
Available online 15 October 2017
0891-4222/ © 2017 Elsevier Ltd. All rights reserved.

development. It is also generally assumed that early learning delays cannot be overcome completely despite remedial teaching programs, and that these learning delays interfere with academic achievements into adulthood for most of the people with dyslexia. Prevalence estimates of dyslexia range from about 5% to about 15% of the population across countries.

Specific characterizations of dyslexia are provided by various theories which were developed in more than two decades of scientific research. Many theories have related reading difficulties of people with dyslexia to various underlying deficits (see for reviews: Elliott & Grigorenko, 2014; Peterson & Pennington, 2012). Of these, the most frequently reported are phonological deficits. The phonological deficit theory posits that dyslexia is caused by impairments in phonological information processing, probably caused by problems in the access to or fuzziness of phonological representations of spoken words (e.g. Shaywitz & Shaywitz, 2005; Snowling & Hulme, 2005; Vellutio, Fletcher, Snowling, & Scanlon, 2004). In recent years, also visual/attentional deficits have been frequently reported although their relevance as a cause of dyslexia is debated (e.g. Goswami, 2015; Lobier & Valdois, 2015). The most important hypotheses postulate deficits in visual attention span (e.g. Bosse, Tainturier, & Valdois, 2007; Lobier, Zoubrinetzky, & Valdois, 2012; Romani, Tsouknida, Betta, di, & Olson, 2010), temporal spatial attention (e.g. Facoetti, Ruffino, Peru, Paganoni, & Chelazzi, 2008; Facoetti et al., 2009), noise exclusion (e.g. Sperling, Lu, Manis, & Seidenberg, 2006) and visual crowding (e.g. Lorusso et al., 2004; Martelli, Di Filippo, Spinelli, & Zoccolotti, 2009).

Apart from phonological and visual/attentional deficits, other symptoms of dyslexia are also widely investigated and discussed. These include deficits of short-term memory, rapid naming, speed of processing and many more. None of these deficits is regarded as a necessary or sufficient cause of dyslexia. Instead, most researchers adhere to a multiple deficit view of dyslexia, that dyslexia is caused by multiple cognitive factors which operate probabilistically (Pennington, 2006; van Bergen, van der Leij, & de Jong, 2014).

In the present study, our goal is not to evaluate different theories of dyslexia, but to find the best way to diagnose dyslexia in adults. Especially for diagnosing dyslexia, it is crucial to understand that apart from reading and spelling deficits, many other deficits are related to dyslexia. Although the literature on adult dyslexia is sparse compared to that on dyslexia as a developmental disorder in children, various symptoms of dyslexia in adults were recently investigated. As children with dyslexia, adults with dyslexia experience difficulties with phoneme awareness, rapid automatized naming, reading, spelling, written word recognition and working memory (e.g. Cavalli et al., 2016; Kemp, Parrila & Kirby, 2009; Nergård-Nilssen & Hulme, 2014; Tops, Callens, Lammertyn, Van Hees, & Brysbaert, 2012; Vellutio et al., 2004). In addition, deficits in various executive functions have been reported (Smith-Spark, Henry, Messer, Edvardsdottir, & Ziecik, 2016).

In summary, adults with dyslexia seem to experience most of the deficits that are also commonly reported in children with dyslexia. However, an issue to account for is that reading and spelling difficulties of adults with dyslexia may differ depending on the language. Some languages are known for their difficult spelling and related difficulties with phonological transitions (for instance English), while other languages are known for their transparency between their spelling and phonology (for instance Italian). Therefore, the best diagnosis of dyslexia in adults should comprise different measurements for different languages. The present study focuses on the Dutch language. Another issue to account for is that while in children with dyslexia reading and spelling difficulties may be the key symptom of dyslexia, in adults with dyslexia a complicating factor is that some well-educated people may have overcome difficulties such as these. Some deficits may have been compensated for, while other deficits may manifest themselves more severely when compared to children.

Which tests to be used to diagnose dyslexia in adults has hardly been studied. One example in the English language is the Dyslexia Adult Screening Test (DAST) (Nicolson and Fawcett, 1997, 1998), which includes measures of reading, spelling, writing, rapid naming, phonology, working memory, balance, and various verbal skills. In a validation study of the DAST (Harrison & Nichols, 2005), it was found that the spelling, reading, writing, and phonology tests contributed to the best classification accuracy, which however remained below 90 percent. Another example is a study by Tops et al. (2012). In this study, it was examined how a reliable diagnosis of dyslexia could be made in students in higher education with Dutch as their first language. Predictive validity of 53 subtests was examined covering abilities of reading and spelling, phonological awareness, general intelligence, vocabulary, rapid naming, memory, morphology and syntax, math, and speed of processing. Tops et al. recommended that a test battery for adults in higher education with Dutch as their first language should include measures of word reading, word dictation, proofreading, phonological awareness, rapid naming, and calculation. This is quite an extensive list of tests for a reliable diagnosis, and it requires a lot of time to administer all these test as well as the need for test assistants. As a result, the costs of a diagnosis will be substantial and this is a hurdle for persons who must pay for themselves. For these persons, a cheap and easily available screening battery for dyslexia could be of immense help. Two workable solutions are fully digitised tests with automatic reports and digitised self-assessment through self-reports.

One such battery in the English language is the Bangor Dyslexia Test (BDT) (Reynolds & Caravolas, 2016), which is a valid screening tool for dyslexia with a classification rate of 94%. This test battery measures various abilities such as verbal/phonological short-term working memory, spatial awareness, arithmetic skills, and executive functioning. By our knowledge no fully digitised test batteries are available for adults in the Dutch language, although the 'IDAA15 + ' (Schraeyen et al., 2009) was validated for young adults (16 years). A disadvantage of this test is that it only can be used by professional psychologists with relatively large costs for the client.

A second cheap solution for a reliable digitised diagnosis of dyslexia is a self-report questionnaire. In several previous studies, support was found for the reliability and validity of self-assessment of dyslexia (Snowling, Dawes, Nash, & Hulme, 2012; Tamboer & Vorst, 2015; Willcutt, Boada, Riddle, Chabilda, DeFries, & Pennington, 2011). For example, Tamboer and Vorst (2015) reported classification performances using various criteria of dyslexia of 94% or higher. This study also concerns the further validation of the questionnaire that was used by Tamboer and Vorst.

The prime aim of the present study was to validate the Multiple Diagnostic Digital Test Battery for Adults (MDDDT-A), a fully automated computerised battery of both tests and self-report questions for diagnosing dyslexia in adults. The battery can be administered without the help of test assistants. The tests and self-report questions of the MDDDT-A were derived from tests and questions of three previous studies (Tamboer, Vorst & Oort, 2014; Tamboer & Vorst, 2015; Tamboer et al., 2016). We investigated criterion and construct validity of the MDDDT-A in a new sample, by relating the MDDDT-A to a standard battery of tests. The aim was also to replicate previous findings from the three previous studies. The following three questions were asked in this study:

1. What is the relationship of the components of the MDDDT-A with a battery of standard tests (STB) covering the key features of dyslexia? (construct validity)
2. How well can the MDDDT-A battery predict dyslexia in students as compared to a standard test battery of dyslexia? (predictive validity)
3. To what extent can a self-report questionnaire on symptoms of dyslexia contribute to the prediction of dyslexia?

Regarding construct validity, we aimed to create a set of tests and questions covering as many cognitive deficits of dyslexia as possible. In the Netherlands, diagnostic test batteries usually consist of measures of spelling, reading, and some presumed underlying constructs: phonological awareness, rapid naming, and short-term memory. The STB used in this study consisted of all these measures including a measure of visual attention span. In one of three previous studies (Tamboer et al., 2016), we investigated the factor structure of an extended version of the tests and questions of the MDDDT-A and found five latent variables explaining 60% of the variance. The factors were spelling, phonology, short-term memory, rhyme/confusion and whole-word-processing/complexity. It remained unclear, however, how the last two factors are related to well-known deficits of for example reading, rapid naming and visual attention span. A problematic issue is that deficits such as rapid naming and reading cannot easily be measured directly in a digitalised test battery without an assistant, because computerised speech recognition is not accurate enough. Thus, in the tests of the MDDDT-A battery some features of dyslexia cannot be included, although they may be included in the self-report questionnaire.

Regarding predictive validity, the question is whether it is necessary to measure all deficits that accompany dyslexia. The existence of many theories and symptoms does not necessarily require that all symptoms are measured to obtain a high predictive validity. For instance, Tops et al. (2012) found that the combination of only three tests (word reading, word spelling and phonological awareness) sufficed for the identification of dyslexic and non-dyslexic students in higher education with 91% accuracy. However, consistent with their recommendation to use many tests for a reliable diagnosis, it is yet unclear whether these results can be generalised to a general population sample. In one of our studies (Tamboer et al., 2014), we found that predictive validity largely depends on which criterion of dyslexia is used. For instance, predictive validity increased in this study when less stringent inclusion and exclusion criteria were used for groups of students with and without dyslexia. Therefore, an optimal criterion will be investigated in the present study as well.

In a third study (Tamboer & Vorst, 2015), we investigated predictive and construct validity of self-report questions. We found that less than 20 self-report items are sufficient to predict dyslexic and non-dyslexic students with an accuracy of at least 94%. Construct validity was supported by a five-factor structure of various dyslexia-related difficulties. We suggested that self-report questions have a higher predictive validity than standard tests of dyslexia, because self-report questions may suffer less from certain types of bias than tests. For instance, in contrast to tests, it can be assumed that there is hardly any influence of general intelligence or schooling on questions because participants compare their own performances with their own social environment.

In the present study, we selected a new sample of Dutch students with and without dyslexia. They completed the MDDDT-A, the STB (administered by a test assistant), and a digitised self-report questionnaire. We assessed construct and predictive validity of the two test batteries and questionnaire in several ways. We investigated construct validity of the test battery and the questionnaire with correlational and principal components analyses. Predictive validity was determined by comparing numerous predictions using sum scores and item scores of the two test batteries and of the self-report questionnaire. Thereby, we used an independent biographical criterion of dyslexia. Finally, we examined the factor structure of the best predictors of our biographical criterion of dyslexia.

2. Methods

2.1. Participants

A sample of 154 participants (mean age 20.9 years, 143 females) completed all tests and the questionnaire. All participants were students at the University of Amsterdam, were raised in the Netherlands, and none of them had serious health problems or a history of serious neurological disorders.

2.2. Assessment of dyslexia

For an assessment of dyslexia, a biographical criterion was used. This criterion was determined with six biographical indicators of dyslexia, which were assessed in the questionnaire:

1. An official certificate of dyslexia, acquired by an educational psychologist (dyslexia = 2, doubt = 1, no dyslexia = 0),
2. Test results at school or at an institution (dyslexia = 2, doubt = 1, no dyslexia = 0)
3. Extra lessons or remedial training during school days (yes = 2, doubt = 1, no = 0)
4. Information about dyslexic family members (biological mother, father, sisters or brothers) (at least one with dyslexia = 2,

doubt = 1, no = 0),

5. A self-report of dyslexia: “Are you dyslexic?” (yes = 2, doubt = 1, no = 0)
6. A self-report of language difficulties during school years and at present: five questions about reading, spelling, and writing (yes = 2, doubt = 1, no = 0). The score was the mean score over the five questions.

Because an official certificate can be based on tests that resemble the tests in this study, we separated the official certificate from the other biographical indicators. We summarized the scores of the remaining five biographical indicators of dyslexia. Consequently, scores could range from 0 to 10. All students on the first biographical criterion, that is with an official certificate of dyslexia (28), had a total score of 6 or higher on the sum of the other indicators of dyslexia. Only 6 participants without an official certificate had scores of 6 or higher as well. Based on this relatively high consistency, we decided to use only the participants with an official certificate as a criterion group for dyslexia. Participants without an official certificate of dyslexia and a total score of 0 or 1 (72) were categorised into the criterion group of not having dyslexia. With this procedure, 54 participants remained of whom we could not determine with high certainty whether they had dyslexia or not.

An advantage of this procedure is that the selection of the dyslexic and the non-dyslexic group was independent of the tests that we aimed to validate. A disadvantage, however, is that the number of students that did not qualify for either group is large. In most studies, tests are used for a selection of participants with and without dyslexia, although usually with strict inclusion and exclusion criteria. Participants with only moderate symptoms of dyslexia are usually excluded. Thus, also in these studies the results of studies are based on the performance of extreme groups. For analyses of reliability of predictions, this is a severe problem. The use of strict criterion groups might result in misclassifications. People with moderate symptoms of dyslexia may only be identified when their performances on tests are compared with performances of other people with moderate symptoms. Therefore, we also extended the two criterion groups to two larger groups based on consistency between the biographical information and the first prediction analyses. This resulted in groups of 37 students with dyslexia, 98 students without dyslexia and 19 students who could not be identified. For the details, see the results.

2.3. Tasks

2.3.1. MDDDT-A (Tests)

The MDDDT-A consists of 12 tests that are related to cognitive aspects of language processing, such as phonological, visual, attentional, auditory, and spelling difficulties.

Cronbach’s Alpha was calculated for each test in the whole group of 154 students, except for *Letter Order* (limited testing time resulted in not all students finishing the test), and *Counting Letters* (only two items).

In the MDDDT-A, conventional oral answers were replaced by answers typed into the computer. Most of the tests required a choice between two or three answers. In test 1, 2, 3, 6, and 9 (see below), real words or sentences had to be typed into the computer. This might be problematic for two groups of people: elderly with little or no experience in typing and people with dyslexia who might experience some sort of confusion when typing the keys on the keyboard. We took four measures to limit typing bias as much as possible. First, there was no time limit in these tests. Second, typing was kept to a minimum, mostly requiring the typing of single words. Only in test 1 and 9 whole sentences had to be typed. Third, in these typing tests (1 and 9), only a few words were scored, so that typing errors in irrelevant words were ignored. Fourth, we registered how much time it took to complete the 12 subtests, so that very slow and very fast response times could be analysed further for possible sloppy typing. In the present research, none of these issues came up, because students are usually experienced in typing into a computer. For admitting the MDDDT-A in general population samples, we also included a control task measuring typing speed, so that serious issues with correct typing can be identified in people with little experience with typing.

For people with dyslexia, it is often difficult to read task instructions from a computer screen. Therefore, most instructions were read out aloud through headphones at the same time they were displayed on screen. All tests were also preceded by an example item. In the first six tests, all separate items were presented auditorily through the headphone set. Recordings were made of a well-trained female voice reading out aloud all items. For each item, a separate recording was made. These recordings were combined with the visual presentation on the computer screen.

All tests were scored automatically by the computer. Raw scores were registered for all single test items. Total raw scores were automatically computed. These scores were used for consecutive analyses. For future use of the MDDDT-A these analyses are used for automatic reports by the computer.

1. *Dutch Dictation* consists of 10 sentences in the Dutch language. Each sentence is presented twice through the headphones, first in normal reading speed, then word by word with small breaks between them. The complete sentences must be typed into the computer. There is no time limit. Each sentence consists two words with each two spelling difficulties, thus four spelling difficulties per sentence. Only the mistakes in these two words are counted. This results in a score of 0 (4 errors) − 4 (no errors) for each sentence and a maximum score of 40 for the whole test (Cronbach’s $\alpha = 0.76$).
2. *English Dictation* consists of 10 sentences in the English language. Each sentence is presented once through the headphones. At the same time, this sentence can be read on the computer screen except for two omitted words. These words are repeated once through the headphones. The participants only had to enter these words into the computer. There was no time limit. Although Dutch students are familiar with the English language, some English words are well known for their vulnerability to spelling errors for Dutch people. Each English word represented one spelling difficulty. For instance, the word ‘noise’ can be misspelled as

'noice'. Each word is scored with 0 (incorrect) or 1 (correct). The maximum score is 20 for the whole test (Cronbach's α = 0.64).

3. *Missing Letters* consists of 10 sentences in the Dutch language. Each sentence is played out once through the headphones. At the same time, this sentence is displayed on the computer screen with the difference that in two difficult words a few letters have been omitted. These words are repeated once through the headphones. The participants only must enter the missing letters of these words into the computer. There is no time limit. Each word is scored with 0 (incorrect) or 1 (correct). The maximum score is 20 for the whole test (Cronbach's α = 0.52).
4. *Pseudowords* consists of 30 pseudowords, which are nonwords that sound like real words. Each pseudoword is played out once through the headphones. At the same time this pseudoword is displayed on the computer screen. The participants must decide whether the visually presented pseudoword is spelled correctly, which is the case for half of all pseudowords (participants should click on either 'correct' or 'incorrect', which is displayed on the computer screen). There is a time limit of about six seconds per word. The usual approach for pseudowords is to have the participants read the words aloud themselves. We decided on a different approach however, because it would have been practically impossible to get all students in individual sessions for this test. Each word is scored with 0 (incorrect) or 1 (correct). The maximum score is 30 for the whole test (Cronbach's α = 0.51).
5. *Sound Deletion* consists of 20 difficult Dutch words. Each word is played once through the headphones. Some of these words are pronounced correctly and some words incorrectly by leaving out or adding one sound. On the computer screen, each word is presented three times, each time with a slightly different spelling with one of them being spelled accordingly to what is pronounced. Participants must decide which of the visually presented words they heard through the headphone set. For example, the existing word 'fietsenstalling' − which means bicycle shed − is read out as non-existent 'fiestenstalling'. The possible answers are: 'fietsentalling', 'fiestensalling' and 'fiestenstalling'. There is a time limit of about fifteen seconds per word. Each word is scored with 0 (incorrect) or 1 (correct). The maximum score is 20 for the whole test (Cronbach's α = 0.84).
6. *Spoonerisms* consists of 20 words. A Spoonerism is a compound of two existing smaller words, that also allows for a compound of two other existing words when the first letters of the smaller words are interchanged. For example, the word 'kolen-schop' becomes 'scholen-kop'. Each original word is read out once through the headphones and the participants must type the novel word into the computer. There is a time limit of about fifteen seconds per word. Each word is scored with 0 (incorrect) or 1 (correct). The maximum score is 20 for the whole test (Cronbach's α = 0.87).
7. *Incorrect Spelling* consists of 40 Dutch words. All words are displayed on a computer screen for 50 ms each. Half of the presented words are spelled correctly and half are spelled incorrectly. Participants must decide whether the words are spelled correctly or not (they must click on 'correct' or 'incorrect', which is displayed on the computer screen). There is no time limit for the answers. Each word is scored with 0 (incorrect) or 1 (correct). The maximum score is 40 for the whole test (Cronbach's α = 0.64).
8. *Dutch–English Rhyme Words* consists of 40 Dutch-English word pairs. They are displayed on a computer screen for two seconds with the Dutch words on the right. For half of the word pairs, the nouns of the Dutch and the English word resemble each other visually (English 'deep' and Dutch 'reep'). For the other half of the word pairs, the nouns of the Dutch and the English word resemble each other aurally (English 'room' and Dutch 'bloem'). There are 20 rhyming and 20 non-rhyming pairs. Participants must decide whether the two words of a pair rhyme or not, which is the case in half of all word pairs. There is no time limit. Each response is scored with 0 (incorrect) or 1 (correct). The maximum score is 40 for the whole test (Cronbach's α = 0.83).
9. *Letter Order* consists of 20 sentences in the Dutch language. Words consisting of four letters or more are displayed with all letters in random order except the initial and final letters. Each sentence contains two long words in which the letter order is mixed up in this way. The words in the sentences become more difficult towards the end of the test. The participants must enter the complete sentences into the computer with all words spelled correctly. For example, the word "Aoccdrnig" must be typed as "According". There is a time limit of five minutes for the whole test. Each word is scored with 0 (incorrect) or 1 (correct). The maximum score is 40 for the whole test.
10. *Counting Letters* aims to measure the effects of global reading and consists of two sentences. The idea for this test is based on a well-known language puzzle aimed at counting the number of times the letter [f] appears in the following sentence: 'Finished files are the result of years of scientific study combined with the experience of years'. Many people only see the [f] three times. It has been suggested that the [f] in [of] is overlooked because it sounds more like a [v]. Another suggestion is that [of] is overlooked completely. We created two Dutch sentences based on the same principle. In the sentence 'Het deftige hondje van de man en de vrouw drinkt water uit de kraan', participants must count the number of times the letter [d] appears in the sentence (6 times). In the sentence 'Met de neus en de mond is het niet moeilijk en zelfs gemakkelijk een liedje te neuriën', participants must count the number of times the letter [n] appears in the sentence (8 times). There is a time limit of about 12 s per sentence. The total score is determined by adding the number of correctly counted letters [d] and [n]. The maximum score is 14.
11. *Short-term Memory Test Forward* is a digit span test: the number of digits a person can retain and recall. There are 24 series: 6 with 4, 6 with 5, 6 with 6, and 6 with 7 digits. The digits are presented one by one, for one second each. After displaying the last one of a series of 6, the 6 digits must be typed into the computer in the correct order. There is no time limit for the answers. Each series of digits is scored with 0 (incorrect) or 1 (correct). The maximum score is 24 for the whole test (Cronbach's α = 0.80).
12. *Short-term Memory Test Backward* is also a digit span test: the number of digits a person can retain and recall, but this time in reversed order. There are 24 series: 6 with 3, 6 with 4, 6 with 5, and 6 with 6 digits. The digits are presented one by one, for one second each on a computer screen. After displaying the last one of a series of 6 the 6 digits must be typed into the computer in reversed order. The maximum score is 24 for the whole test (Cronbach's α = 0.83).

2.3.2. Standard test battery (STB)

The STB consists of 10 tests, which must be administered by a test assistant who scores all the answers. All tests are related to cognitive aspects of language processing.

1. The *One-Minute-Test* (Brus & Voeten, 1979) aims to measure word reading ability. This test is widely used in primary schools to assess progress in reading. The test consists of lists of words of increasing difficulty. Participants must read the words as fast as they can. The score is the number of words read aloud correctly in one minute.

2. The *Klepel Test* (Van den Bos, Lutje Spelberg, Scheepstra, & De Vries, 1994) aims to measure nonword reading ability. The test is widely used in dyslexia centres in diagnosing dyslexia. The test consists of lists of pseudowords of increasing difficulty. The score on the test is the number of pseudowords read aloud correctly in two minutes.

3. *Visual Attention Span* (van den Boer, van Bergen & de Jong, 2015) aims to measure visual attention span, the number of orthographic units (e.g., letters, letter clusters or syllables) that can be processed simultaneously at a glance (Bosse et al., 2007). First, a central fixation point is shown for 1000 ms. Second, a white screen is presented for 500 ms. Third, a group of letters is presented for 200 ms. There are 25 items: 5 items consisting of 6 letters, 5 items consisting of 7 letters and 5 items consisting of 8 letters and 10 items consisting of 10 letters. Participants must read aloud all letters. The score is determined by summarizing all letters that were read aloud correctly (maximum = 205).

4. *Amsterdam Sound Deletion Test* (van Bergen, Bishop, van Zuijen, & de Jong, 2015) aims to measure phonological awareness. Nonwords are presented auditory. The words are played from a tape recording by a professional speaker. On all items the participants must first repeat the present word and then repeat the word while leaving out one specific sound. The test consists of 12 items. In the first 8 items (2 groups of 4 items), one sound must be left out (for instance: SKOOM without K = SOOM). In the next 4 items, one sound must be left out twice in the same word (for instance: PORSVUST without S = PORVUT). For each item, accuracy and response time are registered. The score is determined as follows: For the each of the three groups of four items the median response time is determined. Next, the mean of these three values is determined. Finally, a fluency score is computed as the number of correct answers per minute.

5. *Digit Span Forward* is a subtest of the WAIS (Wechsler, 2008). An observer reads aloud a series of digits and the participant must repeat them in the correct order. The test starts with two series of two digits, then two series of three digits, and so on. If two series with the same number of digits are repeated incorrectly, the test is stopped. The score is then determined by the number of correctly repeated series.

6. *Digit Span Backward* is a subtest of the WAIS (Wechsler, 2008). An observer reads aloud a series of digits and the participant must repeat them in reversed order. The test starts with two series of two digits, then two series of three digits, and so on. If two series with the same number of digits are repeated incorrectly, the test is stopped. The score is then determined by the number of correctly repeated series.

7-10. *Rapid Automatized Naming* (RAN) is a subtest of the GL & SCHR (De Pessemier & Andries, 2009) and aims to measure speed of automatized naming. Several items are presented on a card to the participant, who must read aloud the items as fast and as precisely as possible. There are four subtests in which the items are either five different digits, letters, colours or pictures. Thus, one card consists of 50 digits and there are five different digits. Per card, the number of errors the time needed to name all items are registered. For each card the score was transformed to the number of correctly pronounced items per second.

2.3.3. MDDDT-A (Self-report questions)

In one of our previous studies (Tamboer & Vorst, 2015), we investigated 60 general language statements and 140 specific language statements. We found that these statements differentiated with 97% accuracy between persons with and without dyslexia (using a biographical criterion). The general language statements assessed general abilities of reading, speaking, writing, making mental representations, memory and knowledge of foreign languages. The 140 specific statements were designed by a cross classification of three dimensions (7 × 5 × 4). One dimension distinguished between seven distinct aspects of how language is used in daily life or at school and universities: *reading, writing, speaking, listening, copying, taking a dictation* and *reading aloud*. A second dimension distinguished between five distinct levels of how language can be represented: by *sounds, letters, words, sentences* or by *text*. A third dimension distinguished between four different difficulties that accompany dyslexic adults: *skipping (forgetting), adding, changing* and *exchanging*. For instance, dyslexics may skip parts of sentences when reading, exchange letters when writing, change words when speaking, or forget parts of a text when making a dictation. Below one example is given for each subscale of the first dimension. *Reading*: 'Sometimes I skip a letter, which results in reading a different word'; *Writing*: 'Sometimes I forget to write down a syllable'; *Speaking*: 'While speaking, I sometimes exchange similar words'; *Listening*: 'I hear a story exactly like someone tells it'; *Copying*: 'When I copy out a text, I sometimes exchange letters with similar sounds'; *Dictating*: 'I make mistakes in dictation, because I don't hear the correct sounds'; *Reading aloud*: 'When reading aloud, I sometimes repeat a part of the text'.

In the present study, we further analysed two reduced sets of 20 general language statements and 56 specific language statements, which were selected based on our previous study. We selected the best statements based on group differences and predictive power, keeping the design of the specific statements intact as much as possible. All statements were scored on a 7-point Likert scale.

2.4. Procedure

All participants were recruited by E-mail or through advertising at various departments of the University of Amsterdam. They were informed about the general nature of the tests and the questionnaires in advance in accordance with a standard protocol. All tests were assessed in one session in a special testing room where noise from outside the room could not be heard. Only one participant and one assistant were present in the room, and during the digital testing the participant was alone. The order of the STB, the MDDDT-A and the questionnaire was counterbalanced. The total testing time of all tests and questions was between two and three

Table 1
Correlations between 22 tests (Pearson) (N = 154).

	DD	ED	ML	PW	SD	SP	IS	DER	LO	CL	STMF	STMB	OMT	KT	VAS	ASDT	DSF	DSB	RANL	RANN	RANC	RANP
MDDDT-A:																						
Dutch Dictation (DD)	1																					
English Dictation (ED)	.49**	1																				
Missing Letters (ML)	.40**	.33**	1																			
Pseudowords (PW)	.36**	.30**	.41**	1																		
Sound Deletion (SD)	.13	.04	.09	.27**	1																	
Spoonerisms (SP)	.41**	.28**	.29**	.36**	.23**	1																
Incorrect Spelling (IS)	**.51****	.40**	.35**	.38**	.24**	.43**	1															
Dutch-English Rhyme Words (DER)	.27**	.16*	.09	.34**	.35**	.23**	.24**	1														
Letter Order (LO)	.28**	.19*	.14	.33**	.17*	.16	.24**	.37**	1													
Counting Letters (CL)	.19*	.17*	.21**	.20*	.18*	.02	.33**	.19*	.24**	1												
Short-term Memory Test Forward (STMB)	.36**	.24**	.36**	.42**	.20*	.24**	.27**	.22**	.17*	.07	1											
Short-term Memory Test Backward (STMF)	.27**	.15	.36**	.41**	.25**	.31**	.33**	.23**	.36**	.21*	**.67****	1										
STB:																						
One-Minute-Test (OMT)	.39**	.36**	.31**	.28**	.08	.38**	.38**	.18*	.10	.18*	.21*	.23**	1									
Klepel Test (KT)	**.53****	.43**	.36**	.38**	.09	.39**	.43**	.26**	.07	.14	.38**	.32**	**.69****	1								
Visual Attention Span (VAS)	.35**	.23**	.24**	.18*	.20*	.27**	.38**	.16*	.18*	.21**	.28**	.23**	.30**	.40**	1							
Amsterdam Sound Deletion Test (ASDT)	.43**	.27**	.27**	.27**	.18*	.38**	.32**	.24**	.04	.13	.30**	.27**	.38**	**.51****	.33**	1						
Digit Span Forward (DSF)	.16	.10	.10	.23**	.17*	.13	.17*	.15	.17*	.06	.42**	.38**	.20*	.26**	.16	.12	1					
Digit Span Backward (DSB)	.24**	.07	.38**	.28**	.10	.10	.17*	.10	.14	.10	.45**	.46**	.17*	.35**	.33**	.17*	.34**	1				
RAN Letters (RANL)	.26	.22**	.30**	.23**	.19*	.27**	.23**	.25**	.09	.06	.28**	.31**	.72**	**.57****	.30**	.33**	.09	.20*	1			
RAN Numbers (RANN)	.28	.23**	.36**	.29**	.15	.35**	.27**	.28**	.08	.05	.31**	.36**	.72**	**.67****	.30**	.38**	.17*	.27**	**.89****	1		
RAN Colours (RANC)	.29	.16	.30**	.27**	.13	.23**	.23**	.33**	.20*	.18*	.21**	.30**	.42**	.40**	.17*	.31**	.08	.28**	**.58****	**.54****	1	
RAN Pictures (RANP)	.26	.08	.26**	.23**	.09	.24**	.20*	.28**	.20*	.10	.27**	.30**	.43**	.42**	.19*	.23**	.21**	.34**	**.53****	**.57****	**.65****	1

Bold: Correlations > 0.5
* Correlation is significant at the 0.05 level (two-tailed).
** Correlation is significant at the 0.01 level (two-tailed).

hours. Afterwards, the students received a more detailed debriefing. Anonymity was guaranteed by the standard protocol of the University of Amsterdam.

2.5. Design and analyses

Construct validity was investigated by analysing convergent validity between the MDDDT-A with the STB, a diagnostic battery that consists of tests that are commonly used for diagnosing purposes in the Netherlands. When convergent validity between the MDDDT-A and the STB is high, this means that the same underlying constructs are measured. This means that the MDDDT-A is indeed measuring difficulties in dyslexia and not something else. Convergent validity will be investigated with correlational analyses and with a comparison between factor structures of the two test batteries. Both batteries were factor analysed with principal components analysis, and correlations between the resulting components are presented. First, convergent validity of only the tests of the MDDDT-A was investigated. Second, convergent validity of the self-report questions was investigated.

Criterion validity was investigated with discriminant analyses on biographical criterion groups. Two sets of analyses were done. First with biographical criterion groups with strict inclusion and exclusion criteria and then in adjusted criterion groups with less stringent inclusion and exclusion criteria. In each set of analyses, we used six sets of predictors: 1) all test scores of the MDDDT-A, 2) all test scores of the STB, 3) all test scores together, 4) single item scores of the questionnaire, 5) single item scores of the tests of the MDDDT-A, and 6) these item scores together.

Based on predictions with the best classification accuracy, a final classification of students with and without dyslexia was determined. These groups were used to investigate validity of factors of dyslexia. Three analyses were performed: group differences, discriminant analysis with eight factor scores as predictors, and analyses of factor structures in various discriminant functions.

3. Results

3.1. Construct validity

3.1.1. Correlations (Table 1)

To get a first impression of construct validity of the MDDDT-A, we performed correlational analyses with all tests from the MDDDT-A and the STB. High correlations between two tests or test batteries of which one is validated and another is not, indicate high convergent validity, and thus high construct validity. Relatively high and significant correlations were found between the spelling tests of the MDDDT-A and the reading tests *One-Minute-Test* and *Klepel Test*. High and significant correlations were also found between all *short-term memory tests*. Apart from these correlations, which were expected, many more moderate and significant correlations were found.

3.1.2. Construct validity: MDDDT-A (tests) (Table 2)

Table 2 shows the rotated component matrix of the twelve tests of the MDDDT-A. Four components with eigenvalues larger than 1 explained 63% of all variance. Four tests showed high factor loadings on the first component: *Dutch Dictation, English Dictation, Spoonerisms,* and *Incorrect Spelling*. Except for *Spoonerisms*, these tests largely depended on spelling abilities. Thus, we interpret this component as a spelling factor. The second component showed high factor loadings for the memory tests, thus pointing to a memory factor. The third component showed high loadings for *Sound Deletion* and *Dutch-English Rhyme Words*. Seeing that there were some smaller loadings of other tests on this component, a phonological factor seems logical. The interpretation of the fourth component is

Table 2
Rotated Component Matrix: 4 factors explaining 63% of all variance of 12 tests of MDDDT-A (N = 154).

Explained variance	Factor 1 (Spelling) 21.0%	Factor 2 (Short-term memory) 17.5%	Factor 3 (Phonology) 14.6%	Factor 4 (Attention) 10.2%
Dutch Dictation	**0.77**	0.19	0.14	0.08
English Dictation	**0.77**	0.05	-0.04	0.14
Missing Letters	**0.52**	**0.50**	-0.15	0.16
Pseudowords	**0.39**	**0.47**	**0.36**	0.10
Sound Deletion	0.01	0.14	**0.74**	0.01
Spoonerisms	**0.60**	0.18	**0.37**	**-0.36**
Incorrect Spelling	**0.68**	0.16	0.24	0.21
Dutch-English rhyme words	0.16	0.06	**0.76**	0.15
Letter Order	0.14	0.20	**0.45**	**0.47**
Counting Letters	0.17	0.06	0.10	**0.85**
Short-term Memory Test Forward	0.17	**0.86**	0.11	-0.05
Short-term Memory Test Backward	0.09	**0.85**	0.23	0.14

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.
Bold: Factor loadings > 0.3

Table 3
Rotated Component Matrix: 3 factors explaining 69% of all variance of 10 tests of STB (N = 154).

Explained variance:	Factor 1 (Rapid naming) 31.8%	Factor 2 (Productive Phonol gy) 22.7%	Factor 3 (Short-term memory) 14.7%
One-Minute-Test	**0.60**	**0.59**	-0.02
Klepel Test	**0.46**	**0.70**	0.20
Visual Attention Span	-0.01	**0.69**	**0.32**
Amsterdam Sound Deletion Test	0.16	**0.72**	0.04
Digit Span Forward	0.04	0.13	**0.75**
Digit Span Backward	0.21	0.14	**0.79**
RAN Letters	**0.80**	**0.43**	-0.06
RAN Numbers	**0.78**	**0.48**	0.04
RAN Colours	**0.82**	0.05	0.15
RAN Pictures	**0.80**	0.00	**0.32**

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.
Bold: Factor loadings > 0.3

difficult. This factor could represent exchanging, confusion or attention. Only *Counting Letters* showed a high loading. Thus, most likely this component represents some form of attention. As an alternative, we repeated this analysis with oblique rotation instead of varimax rotation, which allowed factors to correlate. However, the patterns of loadings were the same and the correlations between factors low. In summary, we will consider these four factors:

1. Spelling
2. Short-term memory
3. Phonology
4. Attention (exchanging, confusion)

3.1.3. Construct validity: STB (Table 3)

Table 3 shows the rotated component matrix of the ten tests of the STB. Three components with eigenvalues larger than 1 explained 69% of all variance. The first component can be interpreted as a factor rapid naming. It comes as no surprise that the *One-Minute-Test* and *Klepel Test* also load on this factor, because in these tests the words had to be read aloud. The third component can also be interpreted easily as a factor memory because only the memory tests show high factor loadings. The second component is difficult to interpret. High loadings are found for the *Klepel Test*, *Visual Attention Span*, and *Amsterdam Sound Deletion Test*. We decided to interpret this as a productive phonology factor. We will see what happens in the following analyses. As an alternative, we repeated this analysis with oblique rotation instead of varimax rotation, which allows factors to correlate. However, the patterns of loadings were the same and the correlations between factors low. In summary, we will consider these three factors:

1. Rapid naming
2. Productive phonology
3. Short-term memory

Table 4
Correlations between 7 factor scores of MDDDT-A and STB (Pearson) (N = 154)

		MDDDT-A Spelling	MDDDT-A Short-term memory	MDDDT-A Phonology	MDDDT-A Attention	STB Rapid naming	STB Productive phonology	STB Short-term memory
MDDDT-A	Spelling	1						
MDDDT-A	Short-term memory	0	1					
MDDDT-A	Phonology	0	0	1				
MDDDT-A	Attention	0	0	0	1			
STB	Rapid naming	.18*	.23**	.15	.02	1		
STB	Productive phonology	**.52****	.14	.11	-.03	0	1	
STB	Short-term memory	.01	**.49****	.09	.10	0	0	1

Bold: Correlations > 0.4
* Correlation is significant at the 0.05 level (two-talied).
** Correlation is significant at the 0.01 level (two-talied).

Table 5
Rotated Component Matrix: 6 factors explaining 65% of all variance of 22 tests of MDDDT-A and STB (N = 154).

Explained variance:	Factor 1 (Rapid naming) 17.6%	Factor 2 (productive phonology) 15.4%	Factor 3 (Short-term memory) 12.3%	Factor 4 (Confusion?) 7.4%	Factor 5 (Receptive phonology) 6.3%	Factor 6 (Attention) 6.1%
MDDDT-A:						
Dutch Dictation	.16	**.73**	.16	.17	-.01	.15
English Dictation	.07	**.75**	-.01	.15	-.16	.07
Missing Letters	.23	**.45**	**.35**	.14	-.29	.21
Pseudowords	.12	**.45**	**.36**	**.43**	.12	-.02
Sound Deletion	.04	.04	.17	.17	**.78**	.17
Spoonerisms	.19	**.61**	.12	.07	**.33**	-.18
Incorrect Spelling	.10	**.63**	.12	.14	.18	**.35**
Dutch-English Rhyme Words	.25	.14	.04	**.55**	.49	.01
Letter Order	.01	.14	.15	**.73**	.08	.18
Counting Letters	.04	.10	-.01	.29	.04	**.79**
Short-term Memory Test Forward	.13	.27	**.77**	.11	.07	-.08
Short-term Memory Test Backward	.19	.16	**.73**	.27	.12	.07
STB:						
One-Minute-Test	**.70**	**.45**	.03	-.11	.05	.10
Klepel Test	**.57**	**.57**	.24	-.14	.05	.09
Visual Attention Span	.19	**.32**	.23	-.23	.26	**.56**
Amsterdam Sound Deletion Test	**.30**	**.51**	.15	-.16	.29	.08
Digit Span Forward	.04	.05	**.66**	-.01	.20	-.03
Digit Span Backward	.22	.01	**.73**	-.02	-.15	.27
RAN Letters	**.88**	.18	.09	-.03	.13	.02
RAN Numbers	**.87**	.24	.17	-.05	.11	-.02
RAN Colours	**.74**	.06	.09	**.34**	-.03	.11
RAN Pictures	**.74**	-.01	.23	.26	-.04	.05

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.
Bold: Factor loadings > 0.3.

3.1.4. Correlations between test factors of MDDDT-A (tests) and STB (Table 4)

Table 4 shows the correlations between the factor scores of the two analyses. One quite high and significant correlation was found between the two short-term memory factors ($r = 0.49$). Another quite high and significant correlation ($r = 0.52$) was found between the spelling factor of MDDDT-A and the productive phonology factor of STB. No correlations were found between phonology and attention of the MDDDT-A and the factors of the VTB. Thus, the construct validity of the MDDDT-A is only partly supported by the STB.

3.1.5. Construct validity: MDDDT-A and STB (Table 5)

Another way to investigate construct validity of the MDDDT-A is to find latent variables of the full set of tests of both the MDDDT-A and the STB, and then to investigate relations between the separate tests and the latent variables. We performed principal components analyses on the 22 tests. We found 6 components with eigenvalues larger than 1 which explained 65% of variance, see Table 5 for the rotated component matrix. Alternatively, we repeated this analysis with oblique rotation instead of varimax rotation which allows factors to correlate. However, the patterns of loadings were the same and the correlations between factors low.

Comparable to the previous analysis of the STB, here the first component also showed high loadings for the subtests of *Rapid Naming* and *the One-Minute-Test*, and it can therefore be interpreted as a rapid naming factor. The second component is determined by the spelling factor in the MDDDT-A analysis and the productive phonology factor of the STB analysis. We interpreted this as a productive phonology factor. The loadings of the STB tests are lower than in the previous analysis. The reason for this remains somewhat unclear. The third component can clearly be considered a short-term memory factor with high loadings from the memory tests of both batteries. The fourth component seems impossible to interpret. The loadings of *Letter Order* and *Dutch-English Rhyme Words* may point to some sort of confusion. This is our preliminary interpretation. The fifth component has only one high factor loading from *Sound Deletion*. Because the tests *Sound Deletion* and *Dutch-English Rhyme Words* required responses instead of phonological production, we interpreted this factor as a receptive phonology factor. The sixth component is the attention factor of the MDDDT-A analysis with a high loading for *Counting Letters*. A moderately high loading of *Visual Attention Span* seems to confirm this. In summary, we will consider these six factors:

1. Rapid naming
2. Productive phonology

3. Short-term memory
4. Confusion (?)
5. Receptive phonology
6. Attention

Comparing these results with the previous analyses, it becomes clear that the six components are not merely the summation of the components found in the separate test batteries. The rotated component matrix does show that latent variables reveal relations between the two test batteries. Although some interpretations of components are preliminary, it is apparent that both batteries represent most components or latent factors. The MDDDT-A does not support the rapid naming factor, which makes sense, because rapid naming could not be assessed digitally. The confusion factor is not well represented by the tests of the STB.

3.1.6. Construct validity: questionnaire (No table)

With the aim of finding confirmation and possibly better interpretations of the six latent variables which were found in the previous paragraph, we investigated the existence of latent variables in the questionnaire. We performed a principal components analysis on the 76 self-report statements. We found 8 components with eigenvalues larger than 1.6 which explained 62% of variance (varimax rotation). Many components with eigenvalues lower than 1.3 (and higher than 1) were found as well, but they hardly added interpretable variance. We did not include the table here because it would have taken up too much space. The eight components can be interpreted as follows:

1. Spelling/Grammatical rules
2. Confusion/Exchanging
3. Rapid naming
4. Complexity in language
5. Learning English
6. Reading
7. Phonology
8. Attention/Short-term memory

Comparing these factors with six of the tests, we conclude that there are two extra factors and that interpretation is not very straight-forward for all components. There is no clear memory component. The component complexity, which was found in our previous study (Tamboer et al., 2016), was not found in the tests of the present study. The component learning English is new. The fact that more components were found in the present study is the result of two differences between the set of questions of the two studies. In our previous study, we analysed 140 specific language statements whereas in the present study, we analysed a reduced set of 56 specific language statements. In the present study, we also factor analysed general language statements.

3.1.7. Construct validity: MDDDT-A (tests), STB and questionnaire (Table 6)

A principal components analysis on a combination of tests and questions is complicated by method variance. In our previous study (Tamboer et al., 2016), we avoided method variance by transforming sum scores to factor scores. Here, we tried to replicate this.

For the questionnaire, we categorised all self-report statements into the eight components, depending on their highest factor loading. Then we summarized and standardized them to eight Z-scores. We entered these scores in a factor analysis and requested eight components, thus transforming the sum scores to standardized factor scores and preserving all variance.

The set of 22 tests was reduced to 13 standardized sum scores. Tests that showed factor loadings higher than 0.6 on one component but lower than 0.4 on all other components were combined. The *One-Minute-Test* and *the Klepel Test* were combined because they had a high and significant correlation (0.7). This resulted in five sum scores (rapid naming: four subtests; short-term memory: four subtests; spelling: *Dutch dictation* and *English dictation*; reading: *One-Minute-Test* and *Klepel Test*; attention: *Counting letters* and *Visual Attention Span*). Together with eight single test scores, these five sum scores were standardized to 13 Z-scores. We entered these scores in a factor analysis and requested 13 components, thus transforming the sum scores to standardized factor scores and preserving all variance.

The structure within the set of all 21 (i.e. 13 + 8) transformed factor scores was analysed in a principal components analysis. This resulted in eight components with eigenvalues larger than 1, explaining 52% of variance. Table 6 shows the rotated component matrix. As an alternative, we repeated this analysis with oblique rotation instead of varimax rotation which allows factors to correlate. However, the patterns of loadings were the same and the correlations between factors low. We named the components based on their highest loading. In summary, we will consider these eight factors:

1. Spelling
2. Rapid naming
3. Reading
4. Attention
5. Short-term memory/Confusion
6. English (?)
7. Phonology

Table 6
Rotated Component Matrix: 8 factors explaining 52% of 13 factor scores of tests and 8 factor scores of a questionnaire (N = 154).

	Factor 1 (Spelling)	Factor 2 (Rapid naming)	Factor 3 (Reading)	Factor 4 (Attention)	Factor 5 (Confusion/ Short-term memory)	Factor 6 (English?)	Factor 7 (Phonology)	Factor 8 (Complexity?)
Explained variance:	7.4%	7.0%	6.7%	6.4%	6.4%	6.2%	6.0%	5.9%
Questionnaire:								
Spelling / Grammatical rules	**.88**	-.01	.01	.04	.07	.14	.00	-.01
Rapid Naming	-.01	**.80**	-.09	-.23	.07	.03	.09	-.04
Reading	.10	.15	**.80**	.09	-.11	-.12	-.07	.02
Attention / Short-term memory	.02	.23	-.10	**.75**	-.13	.12	.00	.05
Phonology	.00	-.05	.07	.02	.03	.03	**.78**	.03
Confusion / Exchanging	.02	.05	.09	.19	**.79**	.01	-.04	.12
Complexity in language	.01	.06	-.02	-.11	-.03	.04	-.02	**.78**
Learning English	-.04	-.06	.17	-.09	.07	**.77**	-.03	-.03
MDDDT-A:								
Spelling (Dutch + English Dictation)	**.74**	-.05	.16	-.03	.01	-.06	-.01	.05
Sound Deletion	.10	.06	-.15	.24	.02	-.05	**.43**	**-.40**
Spoonerisms	-.04	-.24	-.02	.09	**.34**	**.37**	**-.40**	-.08
Incorrect Spelling	**.32**	.08	-.20	.16	-.18	**.58**	.06	.04
Missing Letters	.22	.27	.18	.01	.22	.02	-.27	**-.32**
Pseudowords	-.07	.17	.02	.05	.19	.27	.23	.27
Letter Order	.13	-.28	-.03	**.38**	.04	-.18	-.11	.05
Dutch-English Rhyme Words	-.08	-.01	.14	-.15	.24	.09	.27	.01
STB:								
Rapid Naming (4 subtests)	-.04	**.67**	.14	.28	.04	-.09	-.13	.12
Reading (OMT + Klepel)	.10	-.08	**.65**	-.06	.07	.19	.16	-.01
Amsterdam Sound Deletion Test	.12	-.01	.01	.16	.21	-.09	.02	**.50**
MDDDT-A/STB:								
Short-term Memory (4 subtests)	.13	.10	**-.32**	-.16	**.56**	-.03	.13	.04
Attention (Counting Letters + VAS)	-.06	-.14	.15	**.53**	.16	.01	.09	-.10

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.
Bold: Factor loadings > 0.3.

8. Complexity (?)

The most important finding is that many factors of the tests and questionnaire with the same interpretation load on the same component in this analysis. The first four components (spelling, rapid naming, reading and attention) show one high factor loading of a test and one high factor loading of the questionnaire. The same is true for phonology, only with moderate loadings on the tests. The fifth factor, short-term memory/confusion is a combination of two factors in our previous study, in which we concluded that these are separate factors. It is also remarkable that the sum score of the *One-Minute-Test* and the *Klepel Test* (reading) is only loading on a separate factor reading, while these tests loaded on rapid naming and attention in the factor-structure of the tests. The sixth and eighth factor are difficult to interpret. In summary, the relations between subtests and the latent variables of this analysis are not very clear. Therefore, conclusions can only be given with the greatest caution. Nevertheless, the MDDDT-A does seem to represent many aspects of dyslexia, although reading and rapid naming was only represented by the self-report questionnaire and not by the tests of the MDDDT-A.

3.2. Criterion validity

3.2.1. Details of analyses

All classifications were derived from discriminant analysis carried out in SPSS. Alternatively, logistic regression analyses could have been used. Pohar, Blas and Turk (2004) evaluated both analyses in various situations, and we found that their recommendations provided a reliable guideline for evaluating various predictions in our previous study (Tamboer et al., 2014). As the above-mentioned

Table 7
Discriminant analyses (cross-validated) on two biographical criterion groups: 28 participants with dyslexia (D) 72 participants without dyslexia (ND).

Discriminant analysis 1:	Sum scores of MDDDT-A	
Potential predictors:	12 sum scores of 12 tests	
Selected predictors in discriminant function:		
1. Dutch Dictation		
2. Missing Letters		
3. Sound Deletion		
4. Incorrect Spelling		
5. Counting Letters		
Classification accuracy (100 participants):	88%	
Sensitivity (28 participants with dyslexia):	82%	
Specificity (72 participants without dyslexia):	90%	
	D	ND
28 D	23	5
72 ND	7	65
Discriminant analysis 2:	Sum scores of STB	
Potential predictors:	10 sum scores of 10 tests	
Selected predictors in discriminant function:		
1. Klepel Test		
2. Amsterdam Sound Deletion Test		
Classification accuracy (100 participants):	84%	
Sensitivity (28 participants with dyslexia):	82%	
Specificity (72 participants without dyslexia):	85%	
	D	ND
28 D	23	5
72 ND	11	61
Discriminant analysis 3:	Sum scores of all tests of both batteries	
Potential predictors:	22 sum scores of 22 tests	
Selected predictors in discriminant function:		
1. Dutch Dictation		
2. Missing Letters		
3. Klepel Test		
4. Amsterdam Sound Deletion Test		
Classification accuracy (100 participants):	90%	
Sensitivity (28 participants with dyslexia):	89%	
Specificity (72 participants without dyslexia):	90%	
	D	ND
28 D	25	3
72 ND	7	65
Discriminant analysis 4:	Items of questionnaire	
Potential predictors:	17 statements	
Selected predictors in discriminant function:		
1. When I'm writing, I often swap letters.		
2. Reading difficult Dutch words does not cause a lot of problems for me.		
3. When making a dictation I almost automatically write down the words without mistakes.		
Classification accuracy (100 participants):	90%	
Sensitivity (28 participants with dyslexia):	86%	
Specificity (72 participants without dyslexia):	92%	
	D	ND
28 D	24	4
72 ND	6	66
Discriminant analysis 5:	Items of MDDDT-A	
Potential predictors:	22 items	

(*continued on next page*)

Table 7 (*continued*)

Selected predictors in discriminant function:		
1. Dutch Dictation: De hachelijke onderneming leidde tot twijfel.		
2. English Dictation: That woman is my teacher, she is very strict.		
3. Short-term Memory Forward: 382986		
Classification accuracy (100 participants):	78%	
Sensitivity (28 participants with dyslexia):	61%	
Specificity (72 participants without dyslexia):	85%	
	D	ND
28 D	17	11
72 ND	11	61
Discriminant analysis 6:	Items of MDDDT-A and questionnaire	
Potential predictors:	39 items	
Selected predictors in discriminant function:		
1. English Dictation: That woman is my teacher, she is very strict.		
2. Statement: When I'm writing, I often swap letters.		
3. Statement: When making a dictation I almost automatically write down the words without mistakes.		
Classification accuracy (100 participants):	92%	
Sensitivity (28 participants with dyslexia):	89%	
Specificity (72 participants without dyslexia):	93%	
	D	ND
28 D	25	3
72 ND	5	67

researchers indicated, discriminant analysis should be preferred when violations of normality of predictor variables are not too bad (skewness within the interval [−0.2, 0.2]) and when the independent variables of predictors have four or more answer categories. Based on these recommendations we decided to use discriminant analysis. This analysis assumes normal score distributions of predictors which was the case for all variables used in our analysis.

We used the *stepwise method* of discriminant analysis to predict group membership (dyslexic or non-dyslexic), because we wanted to acquire a reduced set of predictors while we did not want to assign some predictors higher priority than others. We set prior probabilities to *all groups equal*. A cross-validation procedure was chosen by selecting the option *leave-one-out classification*, to prevent that a sole case would be classified based partly on itself, and to minimize the effects of outliers.

We used only a selection of items from our previous study (Tamboer, et al., 2016) to prevent overfitting. It is generally recommended that the number of predictors used in a prediction analysis does not exceed the number of participants in the smallest group because this could lead to overfitting, meaning that the results become sample-specific and do not generalise to other samples (Tabachnik & Fidell, 2007).

We present total classification accuracy, sensitivity (correctly identified students with dyslexia) and specificity (correctly identified students without dyslexia). It is generally recommended that the latter should be above 90%. Also for the MDDDT-A, high specificity is more important than high sensitivity. For instance, a sensitivity of 90% would imply that 10% of the students with dyslexia are not identified. For them, there is always the possibility to take further steps, although these might be expensive. On the other hand, a specificity of 90% would imply that 10% of students without dyslexia incorrectly obtain a diagnosis of dyslexia. These students may start to worry whereas in fact, they have nothing to worry about. This situation should be avoided as much as possible.

3.2.2. Discriminant analyses on strict criterion groups (Table 7)

Two biographical criterion groups were determined with the questionnaire: 28 students with dyslexia and 72 students without dyslexia as a criterion. The remaining 54 students could not be identified as either having dyslexia or not having dyslexia with reasonable reliability. The two criterion groups were entered in a discriminant analysis. Next, all 154 students were classified. Thus, also the students that were not included in the criterion groups were classified. These were not included in the tables.

We analysed six different sets of potential predictors: 12 sum scores of the tests of the MDDDT-A, 10 sum scores of the tests of the STB, 22 sum scores of both test batteries together, a selection of 17 items of the questionnaire, a selection of 22 test items of the MDDDT-A, and these 39 items together. Six discriminant functions were calculated that maximally distinguished participants with and without dyslexia.

Table 7 shows six classification tables. Also presented are: overall classification accuracy of the two criterion groups combined (100 participants), sensitivity (percentage of 28 participants with dyslexia that is correctly classified), and specificity (percentage of 72 participants without dyslexia that is correctly classified).

In Discriminant Analysis 1, five tests of the MDDDT-A classified 88% of 100 participants correctly. In Discriminant Analysis 2, two tests of the validated battery classified 84% of 100 participants correctly. In Discriminant Analysis 3, four tests of all tests together classified 90% of 100 participants correctly. In Discriminant Analysis 4, three statements of the self-report questionnaire classified

90% of 100 participants correctly. In Discriminant Analysis 5, three items of all tests of the MDDDT-A classified 78% of 100 participants correctly. In Discriminant Analysis 6, one item of the MDDDT-A and two statements of the self-report questionnaire classified 92% of 100 participants correctly. Comparing the classification results, we can draw several conclusions:

1. A comparison of the sets of sum scores of the two test batteries shows that highest classification accuracy is reached when a combination of spelling/reading and phonological abilities is selected in the discriminant function.
2. Highest classification accuracy based on sum scores is reached when all 22 tests are used as potential predictors.
3. When we compare the predictions with single items, we see that self-report statements are better predictors than test items. It is remarkable that only three self-report statements suffice for a good classification.
4. All items together resulted in highest classification accuracy, but overfitting may be at play here.

3.2.3. Adjusting the criterion groups (No table)

Based on the results of the six discriminant analyses, we reordered the two criterion groups and included the remaining group of 54 students that could not be assigned earlier to the dyslexic and non-dyslexic group. Participants could be assigned to a group if the following demands were met. Participants are considered to have dyslexia when they are categorised as such in at least four out of six classifications. They are considered to have no dyslexia when they are categorised as such in at least five out of six classifications. The remaining participants are not categorised as having dyslexia or not.

Of the 28 participants with dyslexia according to the biographical criterion, 26 were categorised as having dyslexia and 2 could not be categorised. Of the 72 participants with no dyslexia on the biographical criterion, 64 were categorised as having no dyslexia and 8 could not be categorised. Within the group of 54 remaining participants, 11 were now categorised as having dyslexia and 34 as having no dyslexia, while 9 participants still could not be categorised. As such, this procedure resulted in 37 participants with dyslexia, 98 participants without dyslexia, and 19 with no clear classification. This resulted in larger groups of students with and without dyslexia then before, while consistency with the criterion groups based on only biographical information remained high.

3.2.4. Discriminant analyses on adjusted criterion groups (Table 8)

Next, we repeated all analyses with the new criterion groups: 37 students with dyslexia and 98 students without dyslexia. Based on the previous analyses, the remaining 19 students could not be reliably assigned to a criterion group. Therefore, the main difference with the previous analyses is that a less stringent diagnostic criterion for dyslexia and no dyslexia is used. The two new criterion groups were entered in a discriminant analysis and then all 154 students were classified. Thus, also the students that were not included in the criterion groups were classified. These were not included in the tables.

We analysed again six different sets of potential predictors: 12 sum scores of the tests of the MDDDT-A, 10 sum scores of the tests of the STB, 22 sum scores of both test batteries together, a selection of 17 items of the questionnaire, a selection of 22 test items of the MDDDT-A, and these 39 items together. Six discriminant functions were calculated that maximally distinguished between participants with and without dyslexia.

Table 8 shows six classification tables indicating overall classification accuracy of the two criterion groups together (135 participants), sensitivity (percentage of 37 participants with dyslexia that is correctly classified), and specificity (percentage of 98 participants without dyslexia that is correctly classified).

In Discriminant Analysis 1, five tests of the MDDDT-A classified 90% of 135 participants correctly. In Discriminant Analysis 2, two tests of the validated battery classified 90% of 135 participants correctly. In Discriminant Analysis 3, four tests of all tests together classified 96% of 135 participants correctly. In Discriminant Analysis 4, three statements of the self-report questionnaire classified 97% of 135 participants correctly. In Discriminant Analysis 5, three items of all tests of the MDDDT-A classified 84% of 135 participants correctly. In Discriminant Analysis 6, one item of the MDDDT-A and two statements of the self-report questionnaire classified 98% of 135 participants correctly. Comparing the classification results with each other and with the six classification results on the biographical criterion, we can draw several conclusions:

1. A comparison of these six analyses with the previous six ones shows that these second series resulted in higher classification accuracy for all sets of potential predictors.
2. Compared to the previous analyses, more predictors were selected in the discriminant functions. This is remarkable, but it also makes sense because the classification of students with less severe dyslexia probably required more predictors in the prediction formulas.
3. All conclusions of the previous analyses can be drawn here as well. It is again remarkable that small sets of single items result in almost perfect classification accuracy in analysis 4 and analysis 6.

3.2.5. Final classification of people with and without dyslexia (No table)

In the next section, we will describe the results of analyses of group differences and predictive validity of factors of dyslexia. Therefore, we made a final classification groups of students by categorizing as many students as possible into groups with and without dyslexia.

We reordered the groups with the same demands as before. Participants are considered to have dyslexia when they are categorised as such in at least four out of six predictions. Participants are considered not to have dyslexia when they are categorised as such in at least five out of six predictions. The remaining participants are not categorised in either group. This resulted in 40 participants with dyslexia, 96 participants without dyslexia, and 18 without categorisation.

Table 8
Discriminant analyses (cross-validated) on two adjusted criterion groups: 37 participants with dyslexia (D) 98 participants without dyslexia (ND).

Discriminant analysis 1:	Sum scores of the MDDDT-A	
Potential predictors:	12 sum scores of 12 tests	
Selected predictors in discriminant function:		
1. Dutch Dictation		
2. English Dictation		
3. Missing Letters		
4. Sound Deletion		
5. Incorrect Spelling		
Classification accuracy (135 participants):	90%	
Sensitivity (37 participants with dyslexia):	76%	
Specificity (98 participants without dyslexia):	95%	
	D	ND
37 D	28	9
98 ND	5	93
Discriminant analysis 2:	Sum scores of the STB	
Potential predictors:	10 sum scores of 10 tests	
Selected predictors in discriminant function:		
1. Klepel Test		
2. Amsterdam Sound Deletion Test		
Classification accuracy (135 participants):	90%	
Sensitivity (37 participants with dyslexia):	86%	
Specificity (98 participants without dyslexia):	91%	
	D	ND
37 D	32	5
98 ND	9	89
Discriminant analysis 3:	Sum scores of all tests of both batteries	
Potential predictors:	22 sum scores of 22 tests	
Selected predictors in discriminant function:		
1. Dutch Dictation		
2. Missing Letters		
3. Sound Deletion		
4. Klepel Test		
5. Digit Span Backward		
6. Amsterdam Sound Deletion Test		
Classification accuracy (135 participants):	96%	
Sensitivity (37 participants with dyslexia):	89%	
Specificity (98 participants without dyslexia):	98%	
	D	ND
37 D	33	4
98 ND	2	96
Discriminant analysis 4:	Items of questionnaire	
Potential predictors:	17 statements	
Selected predictors in discriminant function:		
1. When I'm writing, I often swap letters.		
2. I like to play word games.		
3. When I have to transcribe a word repeatedly, it happens that I write it down differently every time.		
4. When making a dictation I almost automatically write down the words without mistakes.		
Classification accuracy (135 participants):	97%	
Sensitivity (37 participants with dyslexia):	92%	
Specificity (98 participants without dyslexia):	99%	
	D	ND
37 D	34	3
98 ND	1	97

(*continued on next page*)

Table 8 (*continued*)

Discriminant analysis 5:	Items of MDDDT-A
Potential predictors:	22 items
Selected predictors in discriminant function:	
1. Dutch Dictation: Het nieuwe apparaat werd ijverig bestudeerd.	
2. Dutch Dictation: De hachelijke onderneming leidde tot twijfel.	
3. English Dictation: That woman is my teacher, she is very strict.	
4. English Dictation: I write my homework in my notebook.	
5. Missing Letters: De goo … aar gaf een mooie voorste … . (goochelaar, voorstelling)	
6. Sound Deletion: Schermusteling	
7. Short-term Memory Backward: 3952	
Classification accuracy (135 participants):	84%
Sensitivity (37 participants with dyslexia):	65%
Specificity (98 participants without dyslexia):	92%

	D	ND
37 D	24	13
98 ND	8	90

Discriminant analysis 6:	Items of MDDDT-A and questionnaire
Potential predictors:	39 items
Selected predictors in discriminant function:	
1. When I'm writing, I often swap letters.	
2. I like to play word games.	
3. When I have to transcribe a word repeatedly, it happens that I write it down differently every time.	
4. When making a dictation I almost automatically write down the words without mistakes.	
5. English Dictation: That woman is my teacher, she is very strict.	
6. English Dictation: I write my homework in my notebook.	
7. Missing Letters: De goo … aar gaf een mooie voorste … . (goochelaar, voorstelling)	
8. Sound Deletion: Schermusteling	
Classification accuracy (135 participants):	98%
Sensitivity (37 participants with dyslexia):	95%
Specificity (98 participants without dyslexia):	99%

	D	ND
37 D	35	2
98 ND	1	97

For a final classification, we required consistency between the two classifications (Tables 7 and 8). This resulted in 34 participants with dyslexia, 95 participants without dyslexia and 25 participants who could not be classified with high reliability.

Next, we compared this final classification with the classification based only on biographical information. Of the 28 participants with dyslexia according the biographical information, 25 were categorised as having dyslexia and 3 could not be categorised. Of the 72 participants not having dyslexia according biographical information, 62 were categorised as not having dyslexia and 10 could not be categorised. Of the group of 54 participants that could not be categorised with biographical information, 9 were now categorised as having dyslexia and 33 as not having dyslexia, while 12 participants still could not be categorised. We conclude that the classification based on two sets of six predictions was highly consistent with the biographical information.

3.3. Validity of eight dyslexia factors: three analyses

3.3.1. Introduction

We conducted three additional analyses to find support for the legitimacy of the factors of the different analyses described above: group differences, predictive accuracy, and factor analyses of the predictors of dyslexia in prediction formulas.

3.3.2. Group differences (Tables 9 and 10)

We distinguished eight latent variables in the data set of this study. However, if people with and without dyslexia do not differ on these variables, these variables may not be involved in dyslexia. Therefore, group differences on both tests and factors were investigated with the final classification from the predictive analyses (34 people with dyslexia and 95 people without dyslexia). Effect sizes (Cohen's *d*) were calculated by dividing the mean difference by the mean standard deviations over the two groups.

On all tests (see Table 9), students without dyslexia performed significantly better than students with dyslexia ($p < 0.05$). Mean effect size for the tests of the MDDDT-A was 1.07. Mean effect size for the tests of the STB was 1.16. Such effect sizes are generally considered as very high. We also note that most standard deviations are higher for the group of people with dyslexia as compared to the group of people without dyslexia.

Table 9
Differences between people with and without dyslexia on 22 tests (Z-scores).

	No dyslexia (N = 95)		Dyslexia (N = 34)			
	Mean	SD	Mean	SD	p	Effect size
MDDDT-A						
Dutch Dictation	0.49	0.60	-1.12	1.09	< 0.0001	1.91
English Dictation	0.37	0.44	-0.90	1.48	< 0.0001	1.32
Missing Letters	0.44	0.61	-0.95	1.21	< 0.0001	1.53
Pseudowords	0.33	0.80	-0.74	1.14	< 0.0001	1.10
Sound Deletion	0.16	0.80	-0.58	1.51	0.001	0.64
Spoonerisms	0.25	0.86	-0.85	1.02	< 0.0001	1.17
Incorrect Spelling	0.39	0.74	-1.03	1.08	< 0.0001	1.56
Dutch-English Rhyme Wrds	0.15	0.84	-0.41	1.23	0.004	0.54
Letter Order	0.15	0.89	-0.28	1.10	0.024	0.43
Counting Letters	0.22	0.85	-0.34	1.13	0.003	0.57
Short-term Memory Test Forward	0.31	0.94	-0.69	0.98	< 0.0001	1.04
Short-term Memory Test Backward	0.31	0.85	-0.54	0.83	< 0.0001	1.01
STD						
One-Minute-Test	0.33	0.82	-0.95	0.96	< 0.0001	1.44
Klepel Test	0.48	0.71	-1.19	0.79	< 0.0001	2.23
Visual Attention Span	0.27	0.82	-0.74	1.12	< 0.0001	1.04
Amsterdam Sound Deletion Test	0.36	0.93	-0.94	0.59	< 0.0001	1.71
Digit Span Forward	0.13	1.00	-0.32	1.02	0.027	0.45
Digit Span Backward	0.16	0.93	-0.36	0.99	0.007	0.54
Rapid Automized Naming (RAN):						
RAN Letters	0.28	0.89	-0.76	1.04	< 0.0001	1.08
RAN Numbers	0.31	0.86	-0.93	0.91	< 0.0001	1.40
RAN Colours	0.21	0.91	-0.61	1.15	< 0.0001	0.80
RAN Pictures	0.23	0.91	-0.59	0.99	< 0.0001	0.86

Furthermore, we analysed the differences in factor scores of the latent variables which were found in the full set of tests and questions (see Table 10). On six of eight factors, people with dyslexia had significantly lower factor scores than people without dyslexia (not on phonology and complexity). Mean effect size of these six factors is 0.87. This is generally considered as high. We also note that all standard deviations are higher for the group of people with dyslexia as compared to the group of people without dyslexia.

3.3.3. Predictive validity of factors (Table 11)

Predictive validity of the factors was investigated with discriminant analysis (stepwise method, cross-validated) (see Table 11). We entered the eight factor scores in a discriminant analysis as potential predictor. Six factors were used in the regression equation: spelling, rapid naming, reading, attention, short-term memory/confusion and English. Classification accuracy of the whole group was 98% (cross-validated) with a sensitivity of 94% and a specificity of 99%. Apparently, the best prediction of dyslexia is achieved by using many aspects of dyslexia. Consistent with analyses of group differences, phonology and complexity are in this study not necessary for distinguishing between students with and without dyslexia.

3.3.4. Factor structures in discriminant functions (Tables 12–14)

We performed six different predictions twice, once with biographical criterion groups and once with adjusted criterion groups. For convenience, we will call the two different sets of criterion groups Sample 1 and Sample 2. The sixth prediction (all items of the tests of the MDDDT-A and all questions) was most successful, but this was perhaps partly so because the number of potential

Table 10
Differences between people with and without dyslexia on 8 factors.

	No dyslexia (N = 95)		Dyslexia (N = 34)			
	Mean	SD	Mean	SD	p	Effect size
Factors of Dyslexia						
Spelling	0.44	0.60	-0.94	1.21	< 0.0001	1.52
Rapid naming	0.21	0.80	-0.44	1.23	0.001	0.64
Reading	0.16	0.81	-0.55	1.18	< 0.0001	0.71
Attention	0.19	0.87	-0.50	1.19	0.001	0.67
Short-term memory/Confusion	0.25	0.88	-0.65	0.92	< 0.0001	1.00
English (?)	0.15	0.82	-0.54	1.31	0.001	0.65
Phonology	0.01	0.86	-0.10	1.47	0.598	0.09
Complexity (?)	0.04	0.84	-0.02	1.37	0.763	0.05

Table 11
Discriminant analyses (cross-validated) of people with dyslexia (34) and without dyslexia (95).

Potential predictors:	8 factor scores
Selected predictors in discriminant function:	
1. Spelling	
2. Rapid naming	
3. Reading	
4. Attention	
5. Confusion / Short-term memory	
6. English (?)	
Classification accuracy (129 participants):	97.70%
Sensitivity (34 participants with dyslexia):	94.10%
Specificity (95 participants without dyslexia):	98.90%

	D	ND
34 D	32	2
95 ND	1	94

predictors in the discriminant analysis was higher than the size of the smallest group of participants. We will analyse construct validity of all discriminant functions one by one.

Discriminant function 1 (MDDDT-A) consisted of five tests in both samples. In Sample 1, spelling, phonology and attention were represented. In Sample 2, the test *Counting Letters* was replaced by *English Dictation*, eliminating attention.

Discriminant function 2 (STB) consisted of two tests in both samples. Phonology was represented by the *AKT*. The *Klepel Test* loads on two components: rapid naming and attention.

Discriminant function 3 (all tests) consisted of four tests in Sample 1, but of six tests in Sample 2. The difference between discriminant functions 1 and 2 was spelling and attention. In discriminant function 3, both were represented. In Sample 2, short-term memory was added as well. In summary, a discriminant function acquired with all tests as potential predictor resulted in highest classification accuracy in Sample 2 with high construct validity: only the factor confusion was not necessary for the best prediction.

Discriminant function 4 (items of questionnaire) consisted of three items in Sample 1 and of four items in Sample 2. Table 12 shows the correlations between the five different items and the eight factors found previously. Only spelling and short-term memory/confusion show high correlations. But the main conclusion is that variance of all statements is explained by more than one factor. This probably explains their selection in the discriminant function: these items somehow just hit the essence of dyslexia. Another main conclusion is that the first six factors − the factors which were supported by analyses of group differences and predictive analyses − are all well supported by more than one statement. Thus, both sets of statements (in both samples) represent six factors of dyslexia.

Discriminant function 5 (items of MDDDT-A) consisted of three items in Sample 1 and of seven items in Sample 2. Table 13 shows the correlations between the eight different items and the eight factors found previously. Compared to the self-report statements, the test items performed less well in representing the factors of dyslexia. This is consistent with much lower predictive validity: classification accuracy of 84% (self-report statements: classification accuracy of 97%). Apparently, relatively low construct validity in test items results in relatively low predictive validity.

Discriminant function 6 (all items) consisted of three items in Sample 1 and of eight items in Sample 2. No novel items were selected for the discriminant function as compared to the previous discriminant functions. Table 14 shows that − as in discriminant function 4–variance of all statements is explained by more than one factor and that the first six factors all are well supported by more than one statement. Test items do contribute to construct validity (and thus to predictive validity), but only in combination with self-report statements.

3.3.5. Summary

Overall, the legitimacy of at least six factors was found (spelling, rapid naming, reading, attention, short-term memory/confusion and English). These six factors showed significant group differences and were used in a discriminant function. Remarkably, phonology is not very well supported by these analyses as being a key feature of dyslexia. However, some relations between the tests and these factors or between the factors of tests and questions remain unclear. Furthermore, construct validity is best in the discriminant functions of Sample 2, which justifies repeated predictions. The most important finding is that predictive validity is higher when construct validity is high.

Table 12
Correlations between 8factors of dyslexia and self-report statements of discriminant function 4 (N = 154).

	Factor Spelling	Factor Rapid Naming	Factor Reading	Factor Attention	Factor Short-term Memory/ Confusion	Factor English	Factor Phonology	Factor Complexity
When I'm writing, I often swap letters. (Sample 1 & 2)	.32**	.24**	.13	.03	.54**	.23**	.06	.00
I like to play word games. (Sample 2)	.48**	.14	.21*	.26**	.17*	.07	−.01	.07
Reading difficult Dutch words does not cause a lot of problems for me. (Sample 1)	.51**	.18*	.19*	.06	.32**	.24**	.12	.17*
When I have to transcribe a word repeatedly, it happens that I write it down differently every time. (Sample 2)	.38**	.09	.17*	.27**	.51**	.19*	.04	.07
When making a dictation, I almost automatically write down the words without mistakes. (Sample 1 & 2)	.68**	.22**	.18*	.18*	.28**	.29**	.06	.17*

* Correlation is significant at the 0.05 level (two-talied).
** Correlation is significant at the 0.01 level (two-talied).

Table 13
Correlations between 8factors of dyslexia and test items (MDDDT-A) of discriminant function 5 (N = 154).

	Factor Spelling	Factor Rapid Naming	Factor Reading	Factor Attention	Factor Short-term Memory/ Confusion	Factor English	Factor Phonology	Factor Complexity
Dutch Dictation:								
Het nieuwe apparaat werd ijverig bestudeerd. (Sample 2)	.45**	-.04	.11	.17*	.18*	.04	-.15	.02
Dutch Dictation:								
De hachelijke onderneming leidde tot twijfel. (Sample 1 & 2)	.40**	.13	.10	.08	.05	.05	-.02	.07
English Dictation:								
That woman is my teacher, she is very strict. (Sample 1 & 2)	.42**	.09	.31**	.03	.10	.21**	-.07	.11
English Dictation:								
I write my homework in my notebook. (Sample 2)	.35**	-.17*	.13	-.03	.13	.08	-.15	-.11
Missing Letters:								
De goo … aar gaf een mooie voorste … . (goochelaar, voorstelling) (Sample 2)	.19*	.08	.15	.12	.24**	.03	-.14	-.22**
Sound Deletion:								
Schermusteling (Sample 2)	.26**	.04	-.04	.11	.11	.19*	.15	-.12
Short-term Memory Forward: 382986 (Sample 1)	.18*	.04	.02	-.08	.25**	-.02	.05	.08
Short-term Memory Backward: 3952 (Sample 2)	.26**	-.01	.02	.14	.24**	-.06	-.01	.07

* Correlation is significant at the 0.05 level (two-talied).
** Correlation is significant at the 0.01 level (two-talied).

Table 14
Correlations between 8 factors of dyslexia and self-report statements and test items (MDDDT-A) of discriminant function 6 (N = 154).

	Factor Spelling	Factor Rapid Naming	Factor Reading	Factor Attention	Factor Short-term Memory/ Confusion	Factor English	Factor Phonology	Factor Complexity
When I'm writing, I often swap letters. (Sample 1 & 2)	.32**	.24**	.13	.03	.54**	.23**	.06	.00
I like to play word games. (Sample 2)	.48**	.14	.21*	.26**	.17*	.07	-.01	.07
When I have to transcribe a word repeatedly, it happens that I write it down differently every time. (Sample 2)	.38**	.09	.17*	.27**	.51**	.19*	.04	.07
When making a dictation, I almost automatically write down the words without mistakes. (Sample 1 & 2)	.68**	.22**	.18*	.18*	.28**	.29**	.06	.17*
English Dictation:								
That woman is my teacher, she is very strict. (Sample 1 & 2)	.42**	.09	.31**	.03	.10	.21**	-.07	.11
English Dictation:								
I write my homework in my notebook. (Sample 2)	.35**	-.17*	.13	-.03	.13	.08	-.15	-.11
Missing Letters:								
De goo … aar gaf een mooie voorste … . (goochelaar, voorstelling) (Sample 2)	.19*	.08	.15	.12	.24**	.03	-.14	-.22**
Sound Deletion: Schermusteling (Sample 2)	.26**	.04	-.04	.11	.11	.19*	.15	-.12

* Correlation is significant at the 0.05 level (two-talied).
** Correlation is significant at the 0.01 level (two-talied).

4. Discussion

4.1. Construct validity

The first question of this study was how components of the newly developed MDDDT-A are related to components of the STB, a battery of standard tests covering the key features of dyslexia. We investigated construct validity of both test batteries with various principal components analyses and correlations and found that many well-known deficits of dyslexia are represented by the MDDDT-A. The tests of the MDDDT-A did not include measures of rapid naming and reading. However, the self-report questions of the MDDDT-A did identify all deficits of dyslexia that were investigated in this study. Thus, construct validity of the self-report questions was higher than both test batteries in the present study. The conclusion is that construct validity of the MDDDT-A is high.

In one of our previous studies (Tamboer et al., 2016), we found that five factors of dyslexia could be distinguished: spelling, phonology, short-term memory, rhyme/confusion and whole-word-processing/complexity. In the present study, we identified four factors in the tests of the MDDDT-A, three factors in the STB, six factors in the full set of tests, eight factors in the questionnaire, and eight factors when we combined the tests and the questionnaire. We also found three areas with inconsistencies. First, not all factors showed group differences. Second, some factors were separate factors in one analysis and taken together as one factor in another analysis. And third, some tests showed loadings on distinct factors in different analyses. For instance, the *One Minute Test* and the *Klepel Test* of the STB loaded on rapid naming and productive phonology in one analysis, but on reading in another analysis.

Our main aim was to find a single analysis that clearly shows which factors characterize dyslexia best. The fact that the number of selected components varies between analyses is not only a result of the number of latent variables that are present, but also of the choices that had to be made during the analyses. For instance, choices had to be made about the lowest eigenvalue or the amount of variance that required explanation. Additionally, we were confronted with various difficulties such as losing variance when using factor scores, or incorporating more variables in the analysis than could be justified based on a relatively small sample size. Another difference between our previous and present study is found in the samples used. Previously, we performed principal components analyses on a general sample of students, in which the number of students with dyslexia was in accordance with the general prevalence rate, whereas in the present sample the percentage of students with dyslexia was relatively high.

The best way to interpret all analyses is to have an overview and detect the latent variables that seem to be at play in more than one analysis. Then, we can distinguish nine factors in total: Spelling, reading, rapid naming, attention, short-term memory, confusion, phonology, complexity, and learning English. Recent literature showed that adult dyslexia can be characterized by various symptoms (e.g. Cavalli et al., 2016; Kemp et al., 2009; Nergård-Nilssen & Hulme, 2014; Smith-Spark et al., 2016; Tops et al., 2012; Vellutio et al., 2004). Some of the factors of the present study can easily be related to one or more of these symptoms, but others are more difficult to interpret. We believe that specific conclusions can only be made with the greatest caution.

Regarding spelling, we found that spelling difficulties are well supported by the MDDDT-A. Spelling difficulties were found to be a separate factor in most analyses and apparently constitute a key symptom of dyslexia in adults. This is consistent with previous findings that spelling impairments are relatively resistant to improvement as compared to other impairments (Nergård-Nilssen & Hulme, 2014). However, the exact nature of the relations between spelling and some other symptoms (reading, rapid naming, and attention) remain unresolved in the present study.

A remarkable high correlation was found between the spelling factor of the MDDDT-A and the productive phonology factor of the STB. This productive phonology factor was determined by various tests in the STB: two reading tasks (*One Minute Test* and *Klepel Test*), one phonological task (*Amsterdam Sound Deletion Task*), two rapid naming tasks (*RAN Letters and Ran Numbers*) and one visual attention task (*Visual Attention Span*). The loading of visual attention span on this phonological factor seems strange, but the result is in line with findings by van den Boer et al. (2015) suggesting that phonology is involved in this task. The composition of this productive phonology factor clearly points to a complex nature of both spelling abilities and phonological abilities. Note that in the combined analysis of the tests and self-report questions, these factors were separated again.

Regarding short-term memory, the main conclusion from the present study is that it is problematic to find relations between this type of difficulty and other difficulties. Often, the role of working or short-term memory in dyslexia is related to phonological difficulties. However, the relations between working or short-term memory and the typical reading difficulties that accompany dyslexia are harder to understand. For a clear description of the relations between working memory and language, see Gathercole and Baddeley (2014). A remarkable finding in the present study was that short-term memory of both test batteries highly correlated with each other. In the STB, short-term memory was assessed auditory, whereas in the MDDDT-A this was assessed visually. Apparently, this difference does not interfere with what the tests in both batteries aim to measure.

Regarding phonology, we distinguished productive and receptive phonology in different analyses. A phonological factor was found in the final principal components analysis, however, without showing a group difference. The phonological task the *Amsterdam Sound Deletion Test* of the STB loaded in the final analysis on a factor named complexity. Possibly, phonology is at play in various domains of dyslexic's difficulties. This is consistent with the distinction between productive and receptive phonology in the analysis where only the tests were involved. The tests related to phonology in the MDDDT-A were not completely auditory. They had to be filled out visually on a computer screen after listening to audio through headphones. In contrast, the tests related to phonology in the STB required the production of speech. Unresolved also remains that phonological tests showed group differences while the final phonology factor did not. Clearly, future analyses are required to clarify these issues of phonology in relation to the MDDDT-A.

Three factors were hard to interpret: confusion, complexity and learning English. Confusion and complexity were also found in our previous study, but learning English was not. However, of these three, only learning English showed significant group differences. We have tried to interpret them as best as we can, but, we cannot rule out that some tests do not actually measure what they aim to measure and in fact measure something else related to language and dyslexia. For example, *Spoonerisms* is only supposed to measure phonology, but it also appeared to be related to spelling and attention.

The factor learning English first came forward from the questionnaire, but it appeared to correlate with spoonerisms and incorrect spelling. An explanation for why this factor plays a significant role in adult dyslexia in the Netherlands may be that high performing students in the Netherlands could have overcome some of their difficulties with the Dutch language, but not with foreign languages such as English. In the Netherlands, it is a well-known effect for high performing students who have been diagnosed with dyslexia in primary school to experience difficulties with foreign languages all over again when attending secondary school. This is especially true for English because of the lack of transparency between its spelling and its pronunciation and thus higher demands on phonological abilities. The factor learning English did not load on the spelling factor of the tests which included both Dutch and English spelling. Maybe the test *English Dictation* did not represent the severe difficulties that people with dyslexia experience when for instance reading a scientific English textbook. In the questionnaire items refer to these severe difficulties. Some of the spelling items of *English Dictation* may be relatively easy for high performing students. Whether this is the case in other samples remains unknown for now. Therefore, we decided to include *English Dictation* in the MDDDT-A for further analyses in other samples. In a general population, it may be expected that *English Dictation* is relatively difficult to perform as compared to *Dutch Dictation*.

Although confusion and complexity were significant factors in our previous study, they were hard to interpret in the present study. Both came forward as such in the questionnaire. Self-reported confusion correlated to some extent with short-term memory in tests. An interpretation is that short-term memory difficulties are not exclusively memory difficulties but arise from more general working memory deficits such as quick processing of single digits or letters. On the same factor confusion, various tests showed low loadings, such as *Missing Letters*, *Pseudowords*, and *Dutch-English Rhyme Words*. Processing pseudowords strongly depends on the ability to process the sequences of letters or the ability to retrieve from memory the right place of each letter. The factor confusion in the questionnaire is partly determined by items that measure difficulties of for instance exchanging letters or words. Maybe, during quick processing of single digits or letters there is a problem with the quick organization of sounds and letters resulting in for instance exchanging letters and words. This may point to the existence of more general cognitive deficits that accompany dyslexia which are not yet fully understood. This is consistent with the finding of deficits of various executive functions (inhibition, set shifting) in a recent study (Smith-Spark et al., 2016).

The same may be the case with complexity. In the present study, a self-report of complexity in language correlated with sound deletion in the STB. In our previous study, complexity correlated with the processing of whole words, represented by the *Spoonerisms* and *Letter Order* tests in this study. These two tests are recent developments and require that participants perform something that they never did before. Thus, for these tests something new had to be learned and executed. Although speculative, this points to a general difficulty accompanying dyslexia which would allow for dyslexia to be described as a learning disorder instead of solely a language disorder. The idea of a general learning deficit underlying dyslexia is not new. According to the procedural learning hypothesis (Nicolson & Fawcett, 2007; Nicolson, Fawcett, Brookes, & Needle, 2010) dyslexia is the result of impairments in functional networks in the brain including the cerebellum, which permits various other secondary deficits outside the literacy domain while declarative learning remains relatively intact. However, this hypothesis has also been criticized (West, Vadillo, Shanks, & Hulme, 2017).

In summary, we conclude that this study supports the idea that dyslexia is characterized by various impairments in adults, but that the exact nature and relations among these impairments remain unclear. In the present study, some relations or a lack of relations can be the result of differences between the nature of assessments: the digitised MDDDT-A, the STB requiring oral answers and the self-report of difficulties. Nevertheless, many correlations that came forward in this study point towards the existence of general cognitive aspects that accompany adult dyslexia, of which some are better understood and described in the literature than others.

4.2. Predictive validity

The second question of this study was how well the MDDDT-A can predict dyslexia in students as compared to the STB. The results of the tests of the MDDDT-A and the STB were about the same, both with a strict criterion as well as with a more flexible criterion. This came as no surprise because we found a high construct validity of the tests of the MDDDT-A as well as the STB, even though there were some differences. With a flexible criterion, the tests of the MDDDT-A provided a classification accuracy of 90% (five tests in the discriminant function), and the tests of the STB also provided a classification accuracy of 90% (but with only two tests in the discriminant function). This is slightly lower than the classification accuracy (91%) with three tests (word reading, word spelling and phonological awareness) in the study of Tops et al. (2012). The same classification accuracy of 90% was reported in a study for the Norwegian language (Nergård-Nilssen & Hulme, 2014) using six tests in logistic regression analysis (text reading fluency, nonword reading, word spelling, phoneme awareness, working memory, and rapid automatized naming). When the tests of the MDDDT-A and the STB were taken together, this resulted in an even higher classification accuracy of 96% (six tests in the regression equation). The similarity between the Norwegian study and the present study is striking. In both studies, the best combination of predictors of dyslexia were nonword reading (*Klepel Test*), phoneme awareness (*Sound Deletion*, *Amsterdam Sound Deletion Test*), working memory (*Digit Span Backward*), and word spelling (*Dutch Dictation*). A remarkable finding of the present study was that factor scores showed an even higher classification accuracy of 98% (six factor scores in the regression equation). These results point to the general hypothesis that higher classification accuracy is achieved when more separate difficulties of dyslexia are incorporated in the form of factor scores. In factor scores, task and method variance is probably eliminated, thereby resulting in a higher construct validity.

Another finding of the present study regarding predictive validity is that it is higher when a flexible criterion is used than when a strict criterion is used. This was found in all six predictions, meaning that the diagnosis of dyslexia depends on the nature and/or severity of the dyslexia as found in the adults used in the criterion groups. A strict biographical criterion probably only selects the adults with severe dyslexia and adults who are high performing in general. A less stringent biographical criterion probably also includes adults with less severe symptoms of dyslexia or adults who only suffer from a few symptoms. It therefore makes sense to assume that most people with dyslexia are to be found in the middle areas of a normal distribution of severity of dyslexia. Comparing this large group with only people on the far left and right end of the distribution may lead to false positives and false negatives. People on the extreme ends of the distribution cannot only be characterised by having dyslexia or not, but also by having or lacking the ability to compensate for dyslexia.

We also emphasize that higher predictive validity with a flexible criterion was based on more predictors in the regression equation than when a strict criterion was used. An interpretation is that strict criterion groups are relatively homogeneous. This might explain that in some studies only a few difficulties of people with dyslexia need to be assessed for a reliable diagnosis. People with dyslexia in one specific and homogeneous sample may resemble each other in applying compensation strategies for some difficulties while retaining other difficulties. Thus, homogeneous samples can be reliably diagnosed with only a few tests, while heterogeneous samples need more tests for a reliable prediction. Based on the results of the present study, it cannot be established how many and which difficulties should be assessed for high predictive validity in general. This depends on the characteristics of a specific sample which may vary to a large extent across countries. Moreover, also cut-off scores may depend on the characteristics of a specific sample. Therefore, it is recommended to use regression analyses such as logistic regression or discriminant analysis. These analyses determine which tests or questions discriminate best and determine weights for each test or item, which is a better solution than using cut-off scores. Taken one step further, it might even be concluded that not the quality of the different test batteries determines predictive validity, but that predictive validity depends on the method used for analysing the test or item scores.

The third question of this study was to what extent a self-report questionnaire can contribute to the prediction of dyslexia. Self-report questions can contribute to construct validity, especially in a digitised battery that cannot include the standard oral measures of reading and rapid naming. This is an advantage mainly because predictive validity seems to depend on construct validity, which was concluded above, and aligns with the conclusion of Nergård-Nilssen & Hulme (2014), and is recommended by Tops et al. (2012). Furthermore, this was confirmed by predictive validity of the questionnaire itself. Construct validity of the questionnaire was high with eight latent variables, and predictive validity was high with classification accuracy of 97% (with a flexible criterion) being even higher than classification accuracies of the tests, although factor scores provided 98% accuracy. When single test items were combined with self-report statements, classification accuracy was 98%. This result is questionable though, because more potential predictors were used than participants in the smallest group of the sample. Nevertheless, this raises the question whether tests should be used at all, if self-report questions by itself are able to predict dyslexia so well. However, it should be noted that this result came from a high achieving sample of students. We cannot yet be sure whether only a questionnaire would suffice for reliable diagnoses of dyslexia in other samples as well. Finally, seeing that the questionnaire resulted in high classification accuracy, it is remarkable that only four self-report statements were used in the discriminant function. This seems to contradict high construct validity, but the correlational analysis between these statements and the eight factors indicated that these four statements together represented almost all factors. For instance, the statement "When making a dictation, I almost automatically write down the words without mistakes" showed a significant correlation with seven factors. Apparently, some single self-report statements capture the exact nature of certain difficulties of adults with dyslexia, while single tests do not. The finding that single statements predict better than test scores can be explained by the fact that the results of test scores not only depend on the typical difficulties of adults with dyslexia, but also on intelligence and schooling.

4.3. Conclusion and limitations

We validated the MDDDT-A and found that construct validity and predictive validity are very good. We also found that highest predictive validity is achieved when construct validity is high, and when a flexible criterion is used. A remarkable result was that self-report questions have high construct validity and an even higher predictive validity than tests. These results show that the MDDDT-A can be used online as a reliable screening instrument without requiring much time. Theoretically, we found support for dyslexia as a multiple cognitive deficit, although it remains unclear how many independent impairments can be distinguished.

A limitation of this study was that sample size was too small for conducting confirmatory analyses. A second limitation was that we only investigated students in our sample. A third limitation was that reading and rapid naming could not be assessed with digitised tests, although this was compensated by self-report assessment of these abilities.

Screening instruments should account for differences between samples, but we expect that when construct validity is kept as high as possible (for instance by assessing both tests and self-report questions), predictive validity will be high in samples within and across countries. We recommend using regression analyses to determine which tests and items and which weights are most appropriate for high predictive validity in a specific sample. One symptom may be important in one sample, but less important in another sample because the severity of any symptom of dyslexia may vary between languages and between samples of any kind.

References

Bosse, M.-L., Tainturier, M. J., & Valdois, S. (2007). Developmental dyslexia: The visual attention span deficit hypothesis. *Cognition, 104*, 198–230.

Brus, B. T., & Voeten, M. J. M. (1979). *Een-minuut-test, vorm A en B. Verantwoording en HandleidingNijmegen, Berkhout Testmateriaal* (2e druk). Lisse: Swets & Zeitlinger.

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.

De Pessemier, P., & Andries, C. (2009). *Gletschr. test voor gevorderd lezen & schrijven.* Antwerpen: Garant.
Elliott, J. G., & Grigorenko, E. L. (2014). *The dyslexia debate.* Cambridge: Cambridge University Press.
Facoetti, A., Ruffino, M., Peru, A., Paganoni, P., & Chelazzi, L. (2008). Sluggish engagement and disengagement of non-spatial attention in dyslexic children. *Cortex, 44*, 1221–1233.
Facoetti, A., Trussardi, A. N., Ruffino, M., Lorusso, M. L., Cattaneo, C., Galli, R., et al. (2009). Multisensory spatial attention deficits are predictive of phonological decoding skills in developmental dyslexia. *Journal of Cognitive Neuroscience, 22*, 1011–1025.
Gathercole, S. E., & Baddeley, A. D. (2014). *Working memory and language.* Psychology Press.
Goswami, U. (2015). Visual attention deficits and assessing causality in developmental dyslexia. *Nature Reviews Neuroscience, 16*. http://dx.doi.org/10.1038/nrn3836-c2.
Harrison, A. G., & Nichols, E. (2005). A validation of the Dyslexia Adult Screening Test (DAST) in a post-secondary population. *Journal of Research in Reading, 28*(4), 423–434.
Kemp, N., Parrila, R. K., & Kirby, J. R. (2009). Phonological and orthographic spelling in high-functioning adult dyslexics. *Dyslexia, 15*, 105–128.
Lobier, M., & Valdois, S. (2015). Visual attention deficits in developmental dyslexia cannot be ascribed solely to poor reading experience. *Nature Reviews Neuroscience, 16*. http://dx.doi.org/10.1038/nrn3836-c1.
Lobier, M., Zoubrinetzky, R., & Valdois, S. (2012). The visual attention span deficit in dyslexia is visual and not verbal. *Cortex, 48* 768–737.
Lorusso, M. L., Facoetti, A., Pesenti, S., Cattaneo, C., Molteni, M., & Geiger, G. (2004). Wider recognition in peripheral vision common to different subtypes of dyslexia. *Vision Research, 44*, 2413–2424.
Martelli, M., Di Filippo, G., Spinelli, D., & Zoccolotti, P. (2009). Crowding, reading: and developmental dyslexia. *Journal of Vision, 9*, 1–18.
Nergård-Nilssen, T., & Hulme, C. (2014). Developmental dyslexia in adults: Behavioural manifestations and cognitive correlates. *Dyslexia, 20*, 191–207.
Nicolson, R. I., & Fawcett, A. J. (1997). Development of objective procedures for screening and assessment of dyslexic students in higher education. *Journal of Research in Reading, 20*(1), 77–83.
Nicolson, R. I., & Fawcett, A. J. (1998). *The dyslexia adult screening test.* London: Psychological Corporation.
Nicolson, R. I., & Fawcett, A. J. (2007). Procedural learning difficulties: Reuniting the developmental disorders? *Trends in Neuroscience, 30*(4), 135–141.
Nicolson, R. I., Fawcett, A. J., Brookes, R. L., & Needle, J. (2010). Procedural learning and dyslexia. *Dyslexia, 16*, 194–212.
Pennington, B. F. (2006). From single to multiple deficit models of developmental disorders. *Cognition, 101*, 385–413.
Peterson, R. L., & Pennington, B. F. (2012). Developmental dyslexia. *Annual Review of Clinical Psychology, 11*, 283–307.
Pohar, M., Blas, M., & Turk, S. (2004). Comparison of logistic regression and linear discriminant analysis: A simulation study. *Metodološki, 1*, 143–161.
Reynolds, A. E., & Caravolas, M. (2016). Evaluation of the bangor dyslexia test (BDT) for use with adults. *Dyslexia, 22*, 27–46.
Romani, C., Tsouknida, E., Betta, A. M., & di Olson, A. (2010). Reduced attentional capacity, but normal processing speed and shifting of attention in developmental dyslexia: Evidence from a serial task. *Cortex, 47*, 715–733.
Schraeyen, K., Janssens, V., Aerts, A., Meersschaert, E., Baeyens, D., & Geudens, A. (2009). *Diagnose van dyslexie bij jongvolwassenen: Validiteitsonderzoek van de Interactieve Dyslexietest Amsterdam-Antwerpen (IDAA)*.
Shaywitz, S. E., & Shaywitz, B. A. (2005). Dyslexia (Specific reading disability). *Biological Psychiatry, 57*, 1301–1309.
Smith-Spark, J. H., Henry, L. A., Messer, D. J., Edvardsdottir, E., & Ziecik, A. P. (2016). Executive functions in adults with developmental dyslexia. *Research in Developmental Disabilities, 53* 323–241.
Snowling, M. J., & Hulme, Ch. (Eds.). (2005). *The science of reading: A handbook*. Oxford: Blackwell Publishing.
Snowling, M., Dawes, P., Nash, H., & Hulme, C. (2012). Validity of a protocol for adult self-report of dyslexia and related difficulties. *Dyslexia, 18*, 1–15. http://dx.doi.org/10.1002/dys.1432.
Sperling, A. J., Lu, Z.-L., Manis, F. R., & Seidenberg, M. S. (2006). Deficits in perceptual noise exclusion in developmental dyslexia. *Nature Neuroscience, 9*, 862–863.
Tabachnik, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5th ed.). Boston: Pearson/Allyn & Bacon.
Tamboer, P., & Vorst, H. (2015). A new self-Report inventory of dyslexia for students: Criterion and construct validity. *Dyslexia, 21*, 1–34.
Tamboer, P., Vorst, H. C., & Oort, F. J. (2014). Identifying dyslexia in adults: An iterative method using the predictive value of item scores and self-report questions. *Annals of Dyslexia, 64*, 34–56.
Tamboer, P., Vorst, H. C., & Oort, F. J. (2016). Five describing factors of dyslexia. *Journal of Learning Disabilities, 49*, 466–483.
Tops, W., Callens, M., Lammertyn, J., Van Hees, V., & Brysbaert, M. (2012). Identifying students with dyslexia in higher education. *Annals of Dyslexia, 62*, 186–203.
van Bergen, E., Van der Leij, A., & De Jong, P. F. (2014). The intergenerational multiple deficit model and the case of dyslexia. *Frontiers in Human Neuroscience, 8*, 346. http://dx.doi.org/10.3389/fnhum.2014.00346.
van Bergen, E., Bishop, D., van Zuijen, T., & de Jong, P. F. (2015). How does parental reading influence children's reading?: A study of cognitive mediation. *Scientific Studies of Reading, 19*, 325–339.
van den Boer, M., Van Bergen, E., & De Jong, P. F. (2015). The specific relation of visual attention span with reading and spelling in Dutch. *Learning and Individual Differences, 39*, 141–149.
van den Bos, K. P., Lutje Spelberg, H. C., Scheepstra, A. J. M., & De Vries, J. R. (1994). *De Klepel. Vorm A en B. Een test voor de leesvaardigheid van pseudowoorden.* Nijmegen: Berkhout.
Vellutio, F. R., Fletcher, J. M., Snowling, M. J., & Scanlon, D. M. (2004). Specific reading disability (dyslexia): What have we learned in the past four decades? *Journal of Child Psychology and Psychiatry, 45*, 2–40.
Wechsler, D. (2008). *Wechsler adult intelligence Scale–Fourth edition (WAIS–IV)*. San Antonio, TX: NCS Pearson.
West, G., Vadillo, M. A., Shanks, D. R., & Hulme, C. (2017). The procedural learning deficit hypothesis of language learning disorders: We see some problems. *Developmental Science*.
Willcutt, E. G., Boada, R., Riddle, M. W., Chhabilda, N., De Fries, J. C., & Pennington, B. F. (2011). Colorado learning difficulties questionnaire: Validation of a parent-report screening measure. *Psychological Assessment, 23*, 778–791. http://dx.doi.org/10.1037/a0023290.
World Health Organization (2010). *International statistical classification of deseases and related health problems ?tenth revision (Version: 2010).* Geneva, Switserland.