The Flamingo test: A New Diagnostic Instrument for Dyslexia in Dutch Higher Education Students

Liset Rouweler[a*], Nelleke Varkevisser[a], Marc Brysbaert[b], Ben Maassen[a], Wim Tops[a]

[a] Department of Neurolinguistics, University of Groningen, Groningen, The Netherlands;

[b] Department of Experimental Psychology, Ghent University, Ghent, Belgium

* Corresponding Author. Tel.: +31 618 939 616

E-mail address: l.l.m.rouweler@rug.nl

Abstract

In this study, we present a new diagnostic test for dyslexia, called the Flamingo Test, which is a Dutch adaptation of the French Alouette Test. The purpose of the test is to measure students' word decoding skills and reading fluency by means of a grammatically correct but meaningless text. Two experiments were conducted to test the predictive validity of the Flamingo Test. In the first experiment, we compared reading time, error rate and, sensitivity and specificity of the Flamingo Test. In the second experiment, we compared the reading performance on the Flamingo Test of students with dyslexia (N = 51) and non-dyslexic matched control students (N = 51) to the reading performance on two Dutch standard word reading tests, the Leestest Een Minuut voor Studenten (LEMs; 'one-minute word reading test for students'), and the Klepel, a one-minute pseudo-word reading test. Our results show that sensitivity and specificity, as well as the positive predictive value (PPV), of the Flamingo Test are high, with even slightly higher PPVs for the Flamingo Test than for the LEMs and Klepel. Together with the fact that the Flamingo Test is short and easy to administer, we believe that the Flamingo Test is a valuable new diagnostic instrument for reading skills.

Keywords

Dyslexia, Higher Education, Text Reading, The Flamingo Test, LEMs, Klepel

Introduction

Good literacy skills are important for academic success and future vocation. Most adults can read and write without effort. However, about five to ten percent of the population fail to attain automatized reading and writing skills (Boulanger, 2013). The term for these specific reading and writing difficulties is dyslexia, which is broadly defined by inaccurate and slow reading, and/or by poor spelling skills (American Psychiatric Association, 2013; Stichting Dyslexie Nederland, 2016). Dyslexia is a lifelong impairment and many symptoms persist into adulthood. The profile of adults with dyslexia differs somewhat from that of children, for whom poor accuracy, slow reading and phonological deficits are among the core deficits. Adults, however, mainly face problems with the latter two: slow reading and phonological deficits (Callens, Tops & Brysbaert, 2012; Milne, Nicholson, & Corballis, 2003; Swanson & Hsieh, 2009).

Over the last few years, an increasing number of Dutch students with dyslexia entered higher education. It is difficult to give an exact number of the prevalence of dyslexia. Dyslexia International (2017), for example, suggests that 5-10 % of the people have dyslexia. Yet, some research suggests that it could even be as high as around 17 % (Sprenger-Charolles & Siegel, 2016). Students with dyslexia enrolled in higher education can apply for special facilities and resources. These students have to submit proof of their learning disability, e.g. a former dyslexia certificate, or the student needs to be tested in case the report is not accepted. An assessment of the spelling and reading skills is needed to obtain access to the resources and special arrangements. There are some methods available in Dutch for screening of dyslexia, such as word reading tests and questionnaires (Tamboer, Vorst & De Jong, 2017; Tops, Callens, Lammertyn, Van Hees, & Brysbaert, 2012). However, there is a lack of validated, relatively short reading instruments for adults with dyslexia. The goal of our study is therefore to present a newly designed diagnostic reading test: The Flamingo Test.

Screening for Reading Problems at University Level

Identifying and diagnosing students with dyslexia at university entrance has been hampered, because the availability of standardized screening tests and questionnaires for adults in Dutch is more limited compared to the relatively large battery of tests available for children (Tamboer, Vorst & De Jong, 2017; Tops et al., 2012). Tamboer et al. (2017) showed that a self-report questionnaire had the highest predictive validity in screening for dyslexia, but this lacks the objective comparison universities typically require. An extensive test battery also showed a rather high predictive validity in the study of Tamboer et al. (2017), but it may be questioned whether this is an efficient use of resources. Tops et al. (2012) administered an extensive test battery as well and showed that a short protocol, consisting of a word reading test, a word spelling test and a phonological awareness test, is sufficient to distinguish between higher education students with and without dyslexia.

Currently, the most used tests to detect reading problems in Dutch are word reading tests, because the Stichting Dyslexie Nederland (2016), a Dutch organization that promotes the transition from scientific knowledge about dyslexia to clinical practice and education, defines dyslexia as a persistent reading and/or spelling problem *at the word level* (a similar definition is used internationally; e.g., Gough, Hoover & Peterson, 1996). The standard tests are a word reading test (Een-Minuut-Test [One-minute Test]; Brus & Voeten, 1994) and a pseudo-word reading test (De Klepel; Van den Bos, Lutje Spelberg, Scheepstra, & De Vries, 1999). Although these tests were initially developed for children, Tops et al., (2012) and Tamboer et al. (2017) showed the validation of these tests for Dutch/Flemish[1] young adults and provided norms. Moreover, Van den Broeck

[1] Dutch is the official language of The Netherland; Flemish is the Dutch variant used in the Northern half of Belgium (Flanders). Although both languages are very similar, there are differences in pronunciation and word use, comparable to the differences between American and British English.

and Geudens (2012) reported that tests of word naming outperform tests of pseudo-word reading in Dutch.

Tops, Nouwels and Brysbaert (2019) created a new version of the Dutch One-Minute-Test (EMT), specifically designed for adults. Their Leestest Een Minuut voor Studenten (LEMs; ‘One Minute Reading Test for Students’), was designed to (1) avoid ceiling effects that often occur with the original EMT, (2) to include more up-to-date words, and (3) to be freely available for research purposes. The LEMs contains 132 words with an increasing level of difficulty, whereas the EMT only holds 116 words. The test has been normed on 200 first-year students in higher education and correlates .9 with EMT.

There are two extensive test batteries available for Dutch adults with a suspicion of dyslexia (Depessemier & Andries, 2009; 2017; Van der Leij, Bekebrede, Geudens, Schraeyen, Schijf et al., 2012). Depessemier and Andries (2009) developed the GL&SCHR, a test battery to identify reading and writing problems in Flemish speaking adolescents and young adults (from 16 to 24 years). The GL&SCHR consists of three main tests (word spelling, spelling rules and text reading) for diagnostic purposes and nine other subtests to test skills that are often associated with reading and writing problems, like phonological awareness, short term memory, rapid naming, and vocabulary. Even though the GL&SCHR is a validated instrument and the differences between Flemish and Dutch are expected to be small, normative data are lacking for Dutch adolescents and young adults in general. A second test, the Interactive Dyslexia test Amsterdam-Antwerpen (IDAA; Van der Leij et al., 2012), was developed as an online screening instrument and normed for adolescents (16 years and older) and adults from the Netherlands and the Flemish-speaking half of Belgium. It works with flashed presentation, which results in a distinct factor when compared to other tasks given to students with dyslexia (Callens, Tops, Stevens, & Brysbaert, 2014).

Text reading as a diagnostic instrument

In other countries, reading aloud short texts is also used for screening. Compared to word reading tests, i.e. simply reading out loud a list of unrelated words, text reading provides a more natural way of reading, because words are almost never read in isolation. For instance, text reading presents words next to each other in lines of text rather underneath each other. So, text reading may measure the same skills as word reading tests but in a more natural way. It assesses reading time, accuracy, and yields information about the type of errors (Levafrais, 1967; 2005). Researchers in English-speaking countries have not used text reading very often in dyslexia assessment, because they fear that the text contents may obscure the measurement of word decoding skills. Words in context are indeed read faster than words out of context, because the context can be used as a top-down predictor (Jenkins et al., 2003). This means that a reader with poor word decoding skills can use contextual cues as a compensatory mechanism to mask problems with word reading. This will be particularly the case for participants familiar with the topic described in the text.

An interesting solution to the issue of discourse support has been presented in the French Alouette Test (Levafrais, 1967; 2005). The Alouette Test evaluates lexical decoding under normal reading conditions by using a text that is grammatically and syntactically correct, yet carries no meaning, which makes the predictability of content words very low. As a result, the Alouette Test does not provide contextual cues that the reader can use to compensate for decoding difficulties (Torgesen, Rashotte & Alexander, 2001). Interestingly, the French Alouette Test is much preferred to isolated word reading among French-speaking dyslexia practitioners and researchers (Sprenger-Charolles, Colé, Béchennec, & Kipffer-Piquard, 2005). Bertrand, Fluss, Billard and Ziegler (2010) called it the gold standard against which other tests are to be evaluated.

The Alouette Test was developed for the assessment of dyslexia in children, but Cavalli, Colé, Leloup, Poracchia-George, Sprenger-Charolles, and El Ahmadi (2018) showed its usefulness for adult assessment by administering the test to a large normative sample of French university students with and without dyslexia. The results showed that the test was good at predicting the diagnosis based on the outcomes of accuracy, speed and efficiency, making the Alouette Test a valid screening tool for adults, possibly superior to isolated word reading, although the authors did not test the latter possibility.

Present Study

For this study, we designed a Dutch adaptation of the Alouette Test, called the Flamingo Test. The Flamingo Test is not a direct translation but an adapted version of the Alouette Test, using the same principles for constructing the test but applied to Dutch. For instance, the name Alouette [Lark] is quite difficult to pronounce in Dutch [Leeuwerik] and was replaced by Flamingo.

The purpose of our study was threefold. The first goal was to get standardized scores from a reasonable sample of Dutch higher education students with and without dyslexia. We provide data from a normative group, unimpaired readers, and a validation sample, impaired readers. Our second goal was to examine the test's predictive validity as a diagnostic tool for students with dyslexia and to examine sensitivity and specificity for the different outcome measures (i.e. reading accuracy, reading time and reading efficiency). These results refer directly to the test's discriminatory power. Finally, our third goal is to compare the Flamingo Test to the commonly used tests in the Netherlands, the LEMs and Klepel, to check whether the Flamingo Test is as suitable as a diagnostic instrument as those tests. For this, we compare the discriminatory power, correlations, and the sensitivity and specificity scores of all three tests. We designed a first experiment to investigate our first two goals. A second experiment was designed to address our third goal.

Experiment 1

Method

Participants

A sample of 103 students participated in this study, 40 with students with dyslexia and 63 without dyslexia. The students with dyslexia were required to have an official dyslexia certificate. The criterion to validate the dyslexia diagnosis was (sub)clinical score (< pc 10) on the word reading test (LEMs) and the pseudo word reading test (Klepel) and/or the word spelling test of the GL&SCHR (Depessemier & Andries, 2009; Stichting Dyslexie Nederland, 2016). One extra dyslexic student did not meet this criterion and was excluded from the study.

Students were recruited from bachelor and master programs from university and applied science programs. All participants attended higher education in Groningen, a province in the northern part of the Netherlands. Of the participants, 13 students were master students, the majority of the participants were bachelor students. Students were between the ages of 18 and 31 years (average age for both groups: 21;8). All students had normal or corrected-to-normal vision and were native speakers of Dutch. None of the students reported neurobiological deficiencies except for dyslexia in the dyslexia group. The study followed the ethical protocol of the Faculty of Arts of the University of Groningen.

The Flamingo Test

The Flamingo Test is the Dutch adaptation of the Alouette Test (Levafrais, 1967; 2005) and evaluates word decoding and reading fluency. The set-up of the Flamingo Test is similar to the set-up of the Alouette Test, but it is not a one-to-one translation. The test contains 285 words, which should be read aloud within a time limit of 180 seconds. The Flamingo test has no meaningful content, since the text is grammatically correct but the combination of content words is meaningless. Thus, the Flamingo test prevents readers from relying on contextual clues and knowledge of the world (Rack, Snowling, & Olson, 1992).

An English translation of the first two sentences reads as follows: “Under the moss or on the roof, in living hedges or in a cleft oak, spring makes its nests. Spring with nests in the woods.”

The text is divided in five sections and is accompanied by drawings that can provoke contextual errors (e.g. a drawing of a squirrel [*eekhoorn*] close to the word *eenhoorn* [unicorn]. The text also includes rare words like *kreupelhout* [thicket] and *capriolen* [caprices] as well as some confusing words that are orthographically or phonologically similar (e.g., *Vredeleen, mijn vriendin* [Vredeleen, my friend]). Furthermore, it contains a few words that are (phonologically) similar to the word suggested by the context (e.g. *blozen* [blushing] instead of *blinken* [blinking] after *zon* [sun].

Scoring and Analyses

Performance on the test was expressed in three different scores: (1) an accuracy score, (2) a reading time score and (3) a reading efficiency score. The accuracy score shows the number of words correctly read by the participant, including words that were read correctly after a self-correction. The maximum score is 285. Secondly, the score for reading time indicates the required reading time in seconds. Since participants only have three minutes to read the text, the maximum score is 180 seconds. Thirdly, the reading efficiency score is the number of words correctly read per minute, calculated by the following formula: accuracy score / reading time in minutes. Average scores and standard deviations were calculated for each of the main scores (accuracy, time and reading efficiency). An additional error analysis was also conducted, including substantial errors, e.g. omissions, substitutions, and time-consuming errors, e.g. self-corrections, hesitations, repetitions.

Procedure

All participants with dyslexia were recruited via the Student Service Centre by e-mail or through advertising online and at various departments of the University of Groningen. Students without dyslexia were asked directly in various departments at the University of Groningen. Participants were tested in a quiet room with one experiment leader.

Results

Accuracy, Time and Reading Efficiency

The mean scores of the students with dyslexia and control group can be found in Table 1.

[Table 1 near here]

All participants from the control group (NonDys) read the text within the time limit of 180 seconds. Almost every student of the dyslexia group (Dys) was also capable of reading the text within the time limit. Only three participants did not. The largest difference between the two groups in terms of effect size was found for reading time (NonDys = 111 s, SD = 18.6; Dys = 150 s, SD = 16.8): $t(76.6) = 10.7$, $p < .001$, $d = 2.23$, which we evaluated with a Welch's t-test. Because of the ceiling effect for accuracy, the difference between the NonDys group (283 words, SD = 2.0) and the Dys group (274 words, SD = 17.4) was smaller in terms of effect size, but still significant: $t(39.7) = 3.35$, $p = .002$, $d = .85$. Against expectation, the effect size was not larger for reading efficiency than for reading time (NonDys = 156 words per minute, SD = 24.8; Dys = 112 words per minut, SD = 18.3): $t(96.0) = 10.8$, $p < .001$, $d = 2.07$. Norm scores for the three dependent variables can be found in Table 2.

[Table 2 near here]

Error Analysis and Comparison

The total number of errors made was calculated. Norm scores for errors can be found in Table 2. Errors were divided in two subcategories: substantial errors and time-consuming. The total number of errors was the sum of the number of substantial errors and time-consuming errors. Only the substantial errors were of influence on the accuracy score. Time-consuming errors were of influence on the reading time and efficiency. The mean number of errors and SD per group and per category are shown in Table 3.

[Table 3 near here]

The NonDys group made fewer errors than the Dys group both in total and in the two subcategories. A significant difference between the two groups was found for total errors: $t(50.0) = 5.94$, $p < .001$, $d = 1.39$. The differences between the groups in terms of time-consuming errors ($t(63.2) = 4.80$, $p < .001$, $d = 1.05$) and substantial errors ($t(44.9) = 4.38$, $p < .001$, $d = 1.06$) were also significant.

Sensitivity and Specificity

Cut-off scores, false positives and negatives can be found in Table 4. Cut-off scores were based on the lowest 10 % scores of this population.

For accuracy, the cut-off score was 280 or lower. Out of the group with 63 control students, five students were false positives, i.e. were marked as being dyslexic. In the dyslexia group, 11 students out of 40 were false negatives and thus marked as non-dyslexic for the accuracy score.

The cut-off point for reading time was 132 seconds or higher. Within the control group, five students were false positives based on their reading time score. In the group of students with dyslexia, five students were false negatives. The cut-off score for reading efficiency was 127 words per minute or lower. Within the group of control students, six students were marked as being dyslexic based on their efficiency score. In the group of students with dyslexia, six students were marked as non-dyslexic based on their efficiency score. Lastly, we checked whether the false positives and false negatives concern the same students for each category (accuracy, time and efficiency) or whether they concern different students for each category. For this we combined the three categories, so criterion 1 (accuracy) + criterion 2 (reading speed) + criterion 3 (efficiency), in which we took all three cut-off points, i.e. the score that represents the 10^{th} percentile, into account. A student was only identified as a false positive or false negative when (1) a dyslexic student received a score on all three criteria above the 10^{th} percentile or (2) a non-dyslexic student received a score on all three criteria below the 10^{th} percentile. This resulted in three truly false negatives, i.e. students with dyslexia being marked as non-dyslexic. No control students were marked as being dyslexic when combining all three scores.

[Table 4 near here]

Based on the cut-off scores, sensitivity and specificity, and positive and negative predictor values (PPV and NPV), were determined for each measure: accuracy, reading time and reading efficiency. These scores can be found in Table 5.

After calculating the PPV and NPV for our sample group, the PPV and NPV were also calculated when taking the prevalence of dyslexia into account, which we estimated around 10 % based on the numbers of Dyslexia International (2017) and Sprenger-Charolles and Siegel (2016).

For accuracy, the PPV was 51 % and the NPV was 97 %. For reading time the PPV was calculated at 55 % and the NPV at 99 %. The PPV for reading efficiency was 50 % and the NPV was 98 %. Lastly, the PPV for the combined score was calculated at 100 % and the NPV at 99 %.

[Table 5 near here]

Comparison with French Data

Although the tests are not entirely identical (e.g. language and the number of words differ), we compared our raw scores with those of Cavalli et al. (2018). Both the data by Cavalli et al. (2018) and the data we collected are presented in Table 6. Since Cavalli et al. (2018) calculated the efficiency score over 180 seconds instead of 60 seconds, we transformed their scores to the efficiency score as calculated in our study for comparison.

[Table 6 near here]

The Alouette Test has a total number of 265 words, compared to 285 words in the Flamingo test. Therefore, we calculated percentages in order to compare the accuracy scores. The percentage of words that were read accurately was identical for the NonDys groups; both groups read 99-100% of the text correctly. For the Dys groups the accuracy scores were almost identical as well, 96.8% for the Dutch group and 94.7% for the French group. The effect size for accuracy was larger for the Dutch Flamingo Test. Reading time is difficult to compare, since the numbers of words in the texts differ, but we found large standardized effect sizes in both languages ($d = 2.68$ for French compared to $d = 2.14$ for Dutch). Finally, we compared the reading efficiency between the two tests. For the dyslexia groups the efficiency scores were very similar, 113 words read correctly

(French) vs. 112 words read correctly (Dutch) per minute. The efficiency scores between the NonDys groups did differ, however.

Discussion

The Flamingo Test is able to discriminate between students with dyslexia and students without dyslexia on all three measures, which is also the case for the Alouette Test (Cavalli et al., 2018; Lefavrais, 1967; 2005). Furthermore, effect sizes are comparable for both languages, with slightly higher effect sizes for Dutch on accuracy and reading efficiency.

For accuracy, the NonDys group attained ceiling level scores, with the Dys group scoring somewhat below that. This pattern was also found for the French test (Cavalli et al., 2018). Therefore, the predictive validity of accuracy as a separate measure is not as high as for reading time or efficiency. This again shows that accuracy is not the most sensitive marker for dyslexia in adults (Swanson & Hsieh, 2009). Reading time was a more sensitive marker for dyslexia: students with dyslexia were more impaired on reading speed than on accuracy, which was also shown by Cavalli et al. (2018) for French and by Swanson and Hsieh (2009) in a meta-analysis. Sensitivity and specificity were the highest for this individual measure.

Similar results were found for the Dutch and French students with dyslexia on reading accuracy. Interestingly, reading efficiency was different in the non-dyslexic groups: French students read the text more efficiently. At first glance this result is surprising because students in higher education without a reading disability are generally good readers. A possible explanation might be that the Dutch adaptation is a bit more difficult than The Alouette Test. In this context, analyses of the errors are of interest. Obviously, a significant difference was found in the amount of errors: students with dyslexia made more substantial and time-consuming errors than students without dyslexia, which was also true in Cavalli et al. (2018). However, when taking a more in-depth look, some words in the Dutch version seemed particularly challenging for both dyslexic and non-dyslexic students, such as *krabbelt* (error) – *kabbelt* (target). Maybe word frequency plays a role, as *krabbelt* [scribbles] is more frequent in Dutch than *kabbelt* [ripples]. Furthermore, some names derived from another language, such as *Pierrot*, were challenging for the Dutch participants, which could have influenced our results as well.

When studying dyslexia, it is helpful to keep the prevalence numbers of dyslexia in mind. We estimated the prevalence number of dyslexia at 10 % based on international organizations and previous literature (Dyslexia International, 2017; Sprenger-Charolles & Siegel, 2016). Based on our sensitivity and specificity measures we were able to calculate the PPV and NPV for the 10 % prevalence criterion. PPV's for the individual measures, i.e. accuracy, reading time and reading efficiency, varied between 50 and 55 %. NPV's however varied between 97 and 99 %. This indicates that the Flamingo Test is able to classify between the 50 and 55 % of the population with dyslexia correctly based on the individual measures. This number however increases enormously when combining all three scores, resulting in a PPV of 100 % and a NPV of 98 %. This indicates that the Flamingo shows the highest predictive validity when combining the cut-off points of all three scores, i.e. criterion 1 (accuracy) + criterion 2 (reading speed) + criterion 3 (efficiency).

At this point we have reason to believe that the Flamingo Test can be used as a screening instrument for dyslexia in Dutch adults. The test discriminates well between students with dyslexia and students without dyslexia, and sensitivity and specificity scores are high. Specificity ranged between 90.5% to 92.1% and sensitivity ranged between 72.5% and 87.5%, with reading time displaying the highest sensitivity. However, when we included the prevalence of dyslexia, PPV's dropped to 50 - 55 % on the individual measures. We thus believe that more validation is necessary. In particular, we felt it necessary to compare the Flamingo Test to LEMs and Klepel, tests that are currently used to diagnose dyslexia. This is done in Experiment 2.

Experiment 2

Method

Participants

For the second study, 51 students with dyslexia and 51 matched controls completed the tests. Of these students, 21 control students and 39 students with dyslexia also participated in Experiment 1. The control participants were matched to the students with dyslexia on age, gender and field of study. An official dyslexia certificate sufficed as diagnosis dyslexia. This was validated in the same way as in Experiment 1. Therefore, one extra student with dyslexia and the matched control student were excluded from the study.

Mean age for the control group was 21;4 and the mean age for the dyslexic group was 21;5 years old. All students had normal or corrected-to-normal vision and were native speakers of Dutch. None of the students had neurobiological deficiencies which could possibly influence the results. The study followed the ethical protocol of the Faculty of Arts of the University of Groningen.

The Flamingo Test

The design of the Flamingo test is explained in the method section of the first study (see Experiment 1 for more details).

LEMs

The Leestest Een Minuut voor Studenten (LEMs; Tops et al., 2018; 'One Minute Word reading Test for Students) is a Dutch word reading task specifically designed for students in higher education and is based on the original EMT. Participants were instructed to read as many words as accurately and quickly as possible within one minute.

The Klepel

The Klepel (Van den Bos et al., 1994) is a Dutch pseudo-word reading test consisting of 116 pseudo-words (non-existing words that correspond to the Dutch grapheme-phoneme correspondence rules). To avoid ceiling effects, the test was administered in one minute instead of two minutes. Participants were instructed to read as many pseudo-words as accurately and as quickly as possible within one minute.

Scoring

Scoring for the Flamingo Test can be found in the method section of Study 1. For LEMs (Tops et al., 2018) and Klepel (Van den Bos et al., 1994), the number of words read, the number of errors, and the number of words that were read correctly were scored.

Procedure

All participants were recruited via the Student Service Centre by e-mail or through advertising online and at various departments of the University of Groningen. This study was part of a larger test protocol and all participants were informed about this protocol before testing. All tests in this 2,5 hour protocol were assessed in a quiet room with one experimental leader.

Results

LEMs, Klepel and Flamingo test

Scores for the LEMs, Klepel and Flamingo Test can be found in Table 7. Scores for the LEMs are divided in a raw score (number of words read in one minute) and a reading efficiency score (number of words read correctly in one minute). There was a significant difference between the groups on the raw score: *t*(100) = 14.8, *p* < .001, *d* = 3.0, as well as on the reading efficiency score: *t*(100.0) = 15.2, p < .001, *d* = 3.01.

For the Klepel, similar to the LEMs, the scores are divided in a raw score and a reading efficiency score. For the Klepel we also found a significant difference on both the raw score (*t*(100) = 14.1, *p* < .001, *d* = 2.7) and the reading efficiency score (*t*(100.0) = 15.4, p < .001, *d* = 2.9). For the Flamingo Test there were significant differences between the Dys and NonDys groups in terms of accuracy (*t*(100.0) = 3.54, p < .001, *d* = .7, reading time (*t*(100.0) = 12.5, p < .001, *d* = 2.5), and reading efficiency (t(t(100.0) = 12.0, p < .001, *d* = 2.4).

[Table 7 near here]

Correlations LEMs, the Klepel and the Flamingo Test

Correlations were calculated between the reading efficiency scores of each task. Figure 1 presents scatterplots of the scores of each pair of tasks. The results indicated significant and comparable correlations between the Flamingo Test and the LEMs (*r* = .82, *p* < .01), the Flamingo Test and the Klepel (*r* = .85, *p* < .01) and between the LEMs and the Klepel (*r* = .89, *p* < .01).

[Figure 1 near here]

Sensitivity and Specificity

Cut-off scores and, false positives and negatives based on the efficiency scores can be found in Table 8. Cut-off scores of the Flamingo Test were based on the lowest 10% scores of our population. For the Flamingo Test, the cut-off score was 127 or lower. Out of the group with 51 control students, 1 student was marked as being dyslexic on this criterion. In the dyslexia group, 7 students out of 51 were false negatives and therefore marked as non-dyslexic. The cut-off point for the LEMs was <89 (based on Tops et al., 2019). Within the control group, two students were false positives based on their score. In the group of students with dyslexia, seven students were false negatives. The cut-off score for the Klepel was < 50 (based on Tops et al., 2012). Within the group of control students, four students were marked as being dyslexic based on their efficiency score. In the group of students with dyslexia, four students were marked as non-dyslexic.

[Table 8 near here]

Sensitivity and specificity, and PPV and NPV were determined for the reading efficiency score of each test. These scores can be found in Table 9.

[Table 9 near here]

After calculating the PPV and NPV for our sample group, the PPV and NPV were also calculated when including the estimated prevalence number for dyslexia (10 %). For the Flamingo Test, the PPV was 82 % and the NPV was 98 %. For the LEMs, the PPV was calculated at 71 % and the NPV on 98 %. The PPV for the Klepel was 56 % and the NPV was 99 %.

Discussion

The second study compared the Flamingo Test with two diagnostic tests for adults: the LEMs for word reading, and the Klepel for pseudo-word reading. Our results showed that the LEMs, the Klepel and the Flamingo Test were all able to distinguish between students with dyslexia and students without dyslexia on all different measures with large effect sizes. The reading efficiency measure, which is the combined time-accuracy score in all tests, was used to compare the Flamingo Test to the LEMs and the Klepel. Correlational analyses revealed highly significant, positive correlations between all three tests. Moreover, sensitivity, specificity, and PPV for the reading efficiency measure were high for all three tests.

In terms of standardized effect size, the difference between the dyslexic and the control group was largest for LEMs, closely followed by the Klepel, and then the Flamingo Test. In contrast, PPV was the highest for the Flamingo Test, also when considering the 10 % prevalence number of dyslexia, being 82.1 % compared to 70,6 % of the LEMS and 56,2 % of the Klepel.

The LEMs and the Klepel test are already widely used as diagnostic instruments for dyslexia in the Netherlands. In combination with the fact that they are easy and quick to administer, this makes them appealing tests to use. The Flamingo Test is as simple and quick to administer, and appears to be equally valid to the LEMs and the Klepel. Interestingly, the correlation between the Flamingo Test and the other two tests is slightly lower than the inter-correlation of LEMs and the Klepel, suggesting that the Flamingo Test may be tapping into a process not assessed by the other two tests, arguably due to the fact that the words are presented in lines of text and form syntactically coherent sentences. As a result, the main message of our findings is that the Flamingo is a nice addition to the two existing tests, rather than a replacement of one of them. Indeed, when the results of the three tests are combined, we might find optimal assessment.

General Discussion

Dyslexia is the most prevalent learning disability and there is need for more practical assessment instruments specifically designed for adults. For that reason, we present the Flamingo Test, inspired by the French Alouette Test, which is considered the gold standard in French-language research (Bertrand et al., 2010; Cavalli et al., 2018; Sprengers-Charolles et al., 2005).

Our study supports the expectation that the Flamingo Test can be useful for both research purposes and clinical practice. At a coarse level, it measures the same skills as word list reading and pseudo-word list reading, resulting in high correlations with these tests (as shown in Figure 1). This is in line with the claim that tests like the Flamingo Test and the Alouette Test measure word decoding skills and not higher-level text comprehension. The interesting addition of the new test form is that words are presented in coherent lines of text like in normal reading. This allows researchers to examine the extent to which reading difficulties are due to factors such as visual discomfort or to eye movement problems, as investigated for instance by Wilkins (2002), Jones, Obregón, Kelly, and Branigan (2008), Zorzi et al. (2012), and Pan, Yan, Laubrock, Shu, and Kliegl (2014). The similarity in results for Dutch and French suggests that the test can easily be adapted for other languages.

As for clinical practice, the test can be used as a short, hands-on screening test for dyslexia in adults, as indicated by the high predictive validity of the Flamingo Test in Experiments 1 and 2. As a single component, reading time was the strongest marker of dyslexia, followed by reading efficiency. This supports the finding of Swanson and Hsieh (2009), Callens et al. (2012), and Cavalli et al. (2018) that a speed deficit, rather than accuracy, is the core impairment in adult dyslexia. The difference between dyslexics and controls in terms of standardized effect size is larger for reading time (number of seconds needed to read the text) than for reading efficiency (number of words read per minute). This was true in our studies and in Cavalli et al. (2018). Apparently, the use of a variable denominator introduces noise in the measure (see Bruyer & Brysbaert, 2011, for a related observation). A better approach may be to ask all participants to read the full text and simply measure the total reading time (with no maximum of 3 minutes). Alternatively, our data show that a joint combination of accuracy, reading time, and reading efficiency with their own criteria result in the best assessment.

Comparisons between the Flamingo Test, the LEMs and the Klepel indicate that all three tests largely measure the same construct (visual word decoding) and can be used together to improve assessment. Given that each variable involves some measurement error (reliability of the tests is about .9) and unique processes, combining the measures increases accuracy, and deviations between tests can point to specific issues (e.g., processing words in lines of text rather than in columns). Combining the measures is feasible given that each test only takes a few minutes.

The Flamingo Test has some advantages compared to LEMs and Klepel. First, it measures reading in a more natural way than word list reading: reading aloud a list of words may not feel natural to students. Second, the Flamingo Test contains function words and syntactic information, making it possible to see how well participants can make use of this information. Indeed, the paradigm introduced by the Alouette Test exploits a clever way of diminishing top-down text information, while keeping much of the bottom-up processes intact. Finally, more extensive error analysis can be done with the Flamingo Test than with word list reading. The test is somewhat longer and more words can be read in one minute. Such analysis can provide insights into the difficulties of dyslexic readers.

As a suggestion for future research we would recommend comparing the Flamingo Test to tests measuring meaningful text reading. For instance, Brysbaert (2019) reported an average reading aloud rate in healthy readers of 183 words per minute, which is close to the reading rate reported for the Alouette Test (Cavalli et al., 2018; see Table 6), but one standard deviation above the reading rate we observed with the Flamingo test (Table 6). A possible factor may be the length of the words in the text. More in general, by comparing different types of reading materials, we can assess the contribution of various processes to reading speed and how these are affected in specific groups and/or in specific individuals.

A limitation of our study is that we could not control for the potential influence of pronunciation difficulties. We have no reasons to believe that pronunciation was a problem for our students, but at some point it may be interesting to address this possibility. Alternatively, it may be that there is less difference between reading aloud and silent reading in students with dyslexia than in control students, so that a reading aloud test underestimates the difficulties students with dyslexia are confronted with, as argued by Gagliano, Ciuffo, Ingrassia, Ghidoni, Angelini, Benedetto, and Stella (2015). Notice that differences in pronunciation rate may also account for some of the differences between students with and without dyslexia in word list reading and pseudo-word list reading. A second possible limitation is that we do not know what effect the pictures have on performance. Remember that the test contains nine pictures, some of which are related to words in the text, and some of which are confusing given the words in the text. This is a feature of the Alouette Test we kept to maximise the similarity with the test. However, to our knowledge nobody has yet systematically investigated the impact of these pictures, by comparing performance in conditions with and without them.

Funding

This research was supported by the University of Groningen, Centre for Language and Cognition.

Acknowledgements

We would like to thank all participants for their contribution in this study. We are also very grateful to our student assistant Tineke Quartel for her help with the data collection.

References

Bertrand, D., Fluss, J., Billard, C., & Ziegler, J. C. (2010). Efficacité, sensibilité, spécificité: Comparaison de différents tests de lecture. *L'Année Psychologique, 110(2)*, 299-320.

Bos, K.P., van den, Lutje Spelberg, H.C., Scheepstra, A.J.M., & Vries, J.R., de (1994). "De Klepel: Een test voor de leesvaardigheid van pseudowoorden." Amsterdam: Pearson.

Boulanger, V. (2013). "Studeren met een functiebeperking" (Masterproef, KU Leuven). http://www.siho.be/files/nieuwsbrief/Thesis_VincentBoulanger.pdf

Broeninck, N., & Gorter, K. (2001). "Studeren met een handicap. Belemmeringen die studenten met een lichamelijke beperking, psychische klachten of dyslexie in het hoger onderwijs ondervinden." Verwey-Jonker Instituut.

Brus, B. T., & Voeten, M. J. M. (1994). "Een-Minuut-Test (EMT)." Harcourt Test Publishers.

Bruyer, R. & Brysbaert, M. (2011). Combining speed and accuracy in cognitive psychology: Is the Inverse Efficiency Score (IES) a better dependent variable than the mean Reaction Time (RT) and the Percentage of Errors (PE)? Psychologica Belgica, 51, 5-13.

Brysbaert, M. (2019). How many words do we read per minute? A review and meta-analysis of reading rate. Journal of Memory and Language, 109, 104047.

Callens, M., & Tops, W. (2015). "Dyslexie in het hoger onderwijs." *Tijdschrift van de Vlaamse vereniging voor Logopedie en Foniatrie*, (July-August), 80-93.

Callens, M., Tops, W., & Brysbaert, M. (2012). "Cognitive Profile of Students Who Enter Higher Education with an Indication of Dyslexia." *Plos One*, 7(6).

Callens, M., Tops, W., Stevens, M., & Brysbaert, M. (2014). "An exploratory factor analysis of the cognitive functioning of first-year bachelor students with dyslexia." *Annals of Dyslexia*, 64, 91-119.

Cavalli, E., Colé, P., Leloup, G., Poracchia-George, F., Sprenger-Charolles, L., & El Ahmadi, A.

(2018). “Screening for dyslexia in French-speaking university students: An evaluation of the detection accuracy of the *Alouette test.” Journal of Learning Disabilities,* 51(3), 268–282.

De Pessemier, P. & Andries, C. (2009). “Gl&Schr. Test voor gevorderd Lezen en Schrijven.” Antwerpen: Garant Uitgevers.

Dyslexia International (2017) - Better Training, Better Teaching. Cited on, https://www.dyslexia-international.org/wp-content/uploads/2016/04/DI-Duke-Report-final-4-29-14.pdf - accessed 23 October 2019

Gagliano, A., Ciuffo, M., Ingrassia, M., Ghidoni, E., Angelini, D., Benedetto, L., Stella, G et al. (2015). “Silent reading fluency: Implications for the assessment of adults with developmental dyslexia.” *Journal of Clinical and Experimental Neuropsychology*, 37, 972–980.

Jenkins, J.R., Fuchs, L.S., Van den Broek, P., Espin, C. & Deno, S.L. (2003). “Sources of individual differences in reading comprehension and reading fluency.” *Journal of Educational Psychology*, 4, 719-729.

Jones, M. W., Obregón, M., Kelly, M. L., & Branigan, H. P. (2008). Elucidating the component processes involved in dyslexic and non-dyslexic reading fluency: An eye-tracking study. Cognition, 109(3), 389-407.

Leij, A., van der, Bekebrede, J.I., Geudens, A., Schraeyen, K., Schijf, G.M., Garst, H., Willems, H., Janssens, V., Meersschaert, E., & Schijf, T.J. (2014). “*Interactieve Dyslexietest Amsterdam-Antwerpen: Handleiding.*” Uithoorn: Muiswerk Educatief.

Levafrais, P. (1967). “Test de l’alouette: manuel.” Paris, France: Les éditions du centre de psychologie appliquée.

Levafrais, P. (2005). “Alouette-R. Paris: Les éditions du centre de psychologie appliquée.” https://www.nuffic.nl/publicaties/vind-een-publicatie/onderwijssysteem-belgie.pdf

Milne, R. D., Nicholson, T., & Corballis, M. C. (2003). :Lexical access and phonological decoding in adults dyslexic subtypes.” *Neuropsychology*, 17, 362–368.

Pan, J., Yan, M., Laubrock, J., Shu, H., & Kliegl, R. (2014). Saccade-target selection of dyslexic children when reading Chinese. Vision Research, 97, 24-30.

Sprenger-Charolles, L., Colé, P., Béchennec, D., & Kipffer-Piquard, A. (2005). French normative data on reading and related skills from EVALEC, a new computerized battery of tests (end Grade 1, Grade 2, Grade 3, and Grade 4). *Revue Européenne de Psychologie Appliquée/European Review of Applied Psychology, 55(3)*, 157-186.

Sprenger-Charolles., L., L., S. and Siegel., B. (2016) “Prevalence and Reliability of Phonological, Surface, and Mixed Profiles in Dyslexia: A Review of Studies Conducted in Languages Varying in Orthographic Depth,” Scientific Studies of Reading, 15(6): 498-521.

Stichting Dyslexie Nederland (2016). *“Diagnose dyslexie. Brochure van de Stichting Dyslexie Nederland. Herziene versie.”* Bilthoven: Stichting Dyslexie Nederland.

Swanson, H. L., & Hsieh, C. J. (2009). “Reading disabilities in adults: A selective meta-analysis of the literature.” *Review of Educational Research*, 79, 1362–1390.

Tamboer, P., Vorst, H. & De Jong, P. (2017). “Six factors of adult dyslexia assessed by cognitive tests and self- report questions: Very high predictive validity.” *Research in Developmental Disabilities*, 73, 148-163.

Tops, W. (2012). “Dyslexia in Higher Education: Research in Assessment, Writing skills, and Metacognition”. PhD Diss. Ghent University. Available at http://crr.ugent.be/papers/Doctoraat_wim_tops.pdf.

Tops, W., Callens, M., Lammertyn, J., Van Hees, V., & Brysbaert, M. (2012). “Identifying students with dyslexia in higher education.” *Annals of Dyslexia*, 62(3), 186-203.

Tops, W., Nouwels, A., & Brysbaert, M. (2019). “Leestest 1-minuut onder Nederlandse studenten.” *Stem-, Spraak- en Taalpathologie,* 24, 1-22. Available at https://sstp.nl/article/view/26110/33076.

Torgesen, J. K., Rashotte, C. A., & Alexander, A. A. (2001). “Principles of fluency instruction in reading: Relationships with established empirical outcomes.” *Dyslexia, fluency, and the brain* (pp. 333–356). Timonium, MD: York Press.

Van den Broeck, W., & Geudens, A. (2012). Old and new ways to study characteristics of reading disability: The case of the nonword-reading deficit. *Cognitive Psychology, 65(3)*, 414-456.

Wilkins, A. (2002). Coloured overlays and their effects on reading speed: a review. Ophthalmic and Physiological Optics, 22(5), 448-454.

Zorzi, M., Barbiero, C., Facoetti, A., Lonciari, I., Carrozzi, M., Montico, M., ... & Ziegler, J. C. (2012). Extra-large letter spacing improves reading in dyslexia. Proceedings of the National Academy of Sciences, 109(28), 11455-11459.

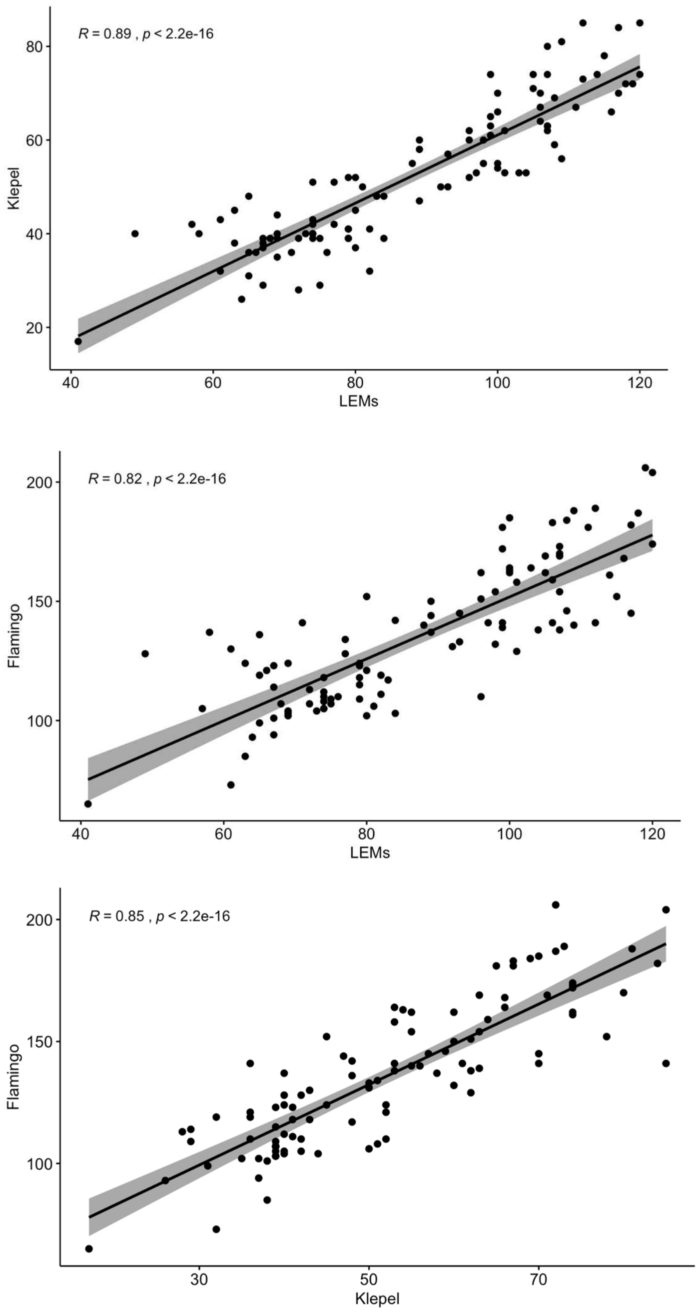


Figure 1. Correlations (and regression line with 95% confidence interval) across groups — between the reading efficiency scores of the Flamingo test and the LEMs and Klepel

Table 1. Accuracy, Time and Reading Efficiency scores on the Flamingo test

	Dys		NonDys				
	M	***SD***	**M**	***SD***	***t***	***p***	***d***
Accuracy	274	*17.1*	283	*2.0*	158.5	< 0.001*	0.94
Time	150	*16.8*	111	*18.6*	54.3	< 0.001*	2.14
Efficiency	112	*18.3*	156	*24.8*	31.3	< 0.001*	1.96

Note. Accuracy = number of words correctly read [max. = 285]; Time = reading time in seconds; Reading efficiency = number of words correctly read in one minute; *p < .001; Dys = dyslexia group; NonDys = control group; d = Cohen's d.

Table 2. Norm scores per measure based on the NonDys group

Percentiles	Accuracy	Time	Efficiency	Total errors	Subs. errors	Time-c. errors
1	≤ 278	≥ 166	≤ 103	≥ 16	≥ 9	≥ 10
5	278	139	122	15	7	9
10	280	132	127	11	6	7
15	281	126	135	10	5	6
20	282	124	137	8	4	6
25	282	123	139	8	4	5
30	282	120	141	8	3	5
35	283	118	144	6	2	4
40	283	115	145	6	2	3
45	284	111	151	5	2	3
50	284	110	155	4	1	3
55	284	108	158	4	1	3
60	284	105	162	4	1	2
65	284	104	163	3	1	1
70	284	103	164	3	1	1
75	284	100	170	2	1	1
80	285	98	172	2	0	1
85	285	92	181	2	0	1
90	285	91	188	2	0	0
95	285	89	191	1	0	0
99	≥ 285	≤ 71	237	≤ 1	0	0

Note. Accuracy = number of words correctly read [max. = 285]; Time = reading time in seconds; Reading efficiency = number of words correctly read in one minute; Substantial errors = number of errors; Time-consuming errors = number of time errors; Total errors = substantial errors + time-consuming errors

Table 3. Error analysis of the Flamingo test

	Dys		NonDys	
	M	***SD***	**M**	***SD***
Substantial errors	7.2 (52%)	*6.8*	2.3 (41%)	*2.3*
Time-consuming errors	6.5 (48%)	*3.7*	3.3 (59%)	*2.6*
Total errors	13.7	*8.1*	5.6	*3.8*

Note. Substantial errors = number of errors; Time-consuming errors = number of time errors; Total errors = substantial errors + time-consuming errors; Dys = dyslexia group; NonDys = control group; total amount of errors is 100 %

Table 4. False positives and false negatives

	Cut-off	*Correctly identified (Dys = 40)*	*Correctly identified (NonDys = 63)*	*False positives*	*False negatives*
Accuracy	< 280	29	58	5	11
Time	> 132	34	58	5	6
Efficiency	< 127	34	57	6	6
Combined score		37	63	0	3

Note. Accuracy = number of words correctly read; Time = reading time in seconds; Reading efficiency = number of words correctly read in one minute; Correctly identified = correctly identified per category and group; false positives = control students being marked as dyslexic; false negatives = dyslexic students being marked as non-dyslexic; Combined score = combining the three cut-off points: accuracy, time and efficiency

Table 5. Sensitivity and specificity scores of the Flamingo test

	Sensitivity	*95% Confidence interval sensitivity*	*Specificity*	*95% Confidence interval specificity*	*PPV*	*NPV*
Accuracy	72.5	56.1 – 85.4	92.1	82.4 – 97.4	85.3	84.1
Time	87.5	73.2 – 95.8	92.1	82.4 – 97.4	87.2	90.6
Efficiency	85	70.2 – 94.3	90.5	80.4 – 96.4	87.2	90.5
Combined score	92.5	79.6 – 98.4	100	94.3 –100	100	95.5

Note. Sensitivity = probability that a test result is positive when the diagnosis is present; specificity = probability that a test result is negative when the diagnosis is not present; Accuracy = number of words correctly read; Time = reading time in seconds; Efficiency = number of words correctly read in one minute; Combined score = combining the three cut-off points to check whether the diagnosis was still present; Positive predictor value (PPV) = probability that dyslexia is present when the test is positive; Negative predictor value (NPV) = probability that dyslexia is not present when the test is negative N = 103 per category

Table 6. Accuracy, Time and Reading Efficiency of the Alouette test and Flamingo test

	Alouette					**Flamingo**				
	Dys (N = 83)		**NonDys (N= 164)**			**Dys (N = 40)**		**NonDys (N = 63)**		
	M	*SD*	M	*SD*	*d*	M	SD	M	SD	*d*
Accuracy	251	*13.4*	262	*2.2*	0.12	274	17.1	283	2.0	0.94
Time	138	*24.1*	87	*11.9*	2.68	150	16.8	111	18.6	2.14
Efficiency	113	*54.7*	184	*81.1*	1.03	112	18.3	156	24.8	1.96

Note. Alouette accuracy and time data taken from Cavalli et al. (2017); Accuracy = number of words correctly read [max. score Alouette = 265, max. score Flamingo test = 285]; Time = reading time in seconds; Reading efficiency = number of words correctly read in one minute

Table 7. LEMs, Klepel and Flamingo scores

Flamingo	**Dys N = 51)**		**NonDys (N = 51)**				
	M	***SD***	**M**	***SD***	***t***	***p***	***d***
Accuracy	275.1	*14.7*	282.4	*2.1*	187.3	< 0.001*	0.7
Time	147.7	*17.5*	108.5	*14.1*	66.8	< 0.001*	2.5
Efficiency	113.6	*16.8*	158.9	*21.8*	43.4	< 0.001*	2.3
LEMs							
Raw score	74.2	*10.8*	104.4	*10.0*	51.4	< 0.001*	3.0
Reading efficiency	72.4	*10.6*	103.5	*10.0*	50.7	< 0.001*	3.0
Klepel							
Raw score	44.6	*7.3*	67.6	*9.6*	37.9	< 0.001*	2.7
Reading efficiency	39.8	*7.2*	64.8	*10.0*	33.2	< 0.001*	2.9

Note. Raw score = number of words read in one minute; Reading efficiency = number of words read correctly in one minute; Accuracy = number of words read; Time = total reading time in seconds; **p* < .001; Dys = dyslexia group; NonDys = control group; *d* = Cohen's d.

Table 8. False positives and false negatives

	Cut-off	*Correctly identified Dys*	*Correctly identified NonDys*	*False positives*	*False negatives*
Flamingo	<127	44	50	1	7
LEMs	<89	44	49	2	7
Klepel	<50	47	47	4	4

Note. Correctly identified = correctly identified per category and group; false positives = control students being marked as dyslexic; false negatives = dyslexic students being marked as non-dyslexic;

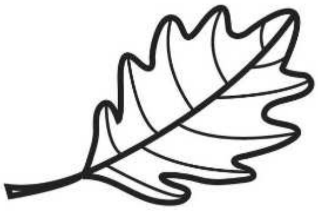
Table 9. Sensitivity and specificity scores for reading efficiency

	Sensitivity	*Specificity*	*95% Confidence interval sensitivity*	*95% Confidence interval specificity*	*PPV*	*NPV*
Flamingo	82.7	98	69.7 – 91.8	89.4 – 99.7	97.7	87.5
LEMs	86.5	96	74.2 – 94.4	86.3 – 99.5	95.7	87.5
Klepel	92.3	92	81.5 – 97.9	80.8 – 97.8	92.3	92.2

Note. Sensitivity = probability that a test result is positive when the diagnosis is present; specificity = probability that a test result is negative when the diagnosis is not present; Positive predictor value (PPV) = probability that dyslexia is present when the test is positive; Negative predictor value (NPV) = probability that dyslexia is not present when the test is negative

Appendix A – *Flamingo test*

De flamingo



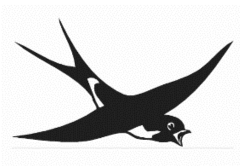
Onder het mos of op het dak,

in levende hagen of in een gespleten eik,

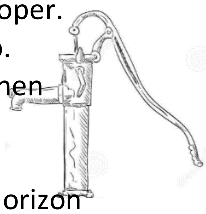
maakt de lente haar nesten.

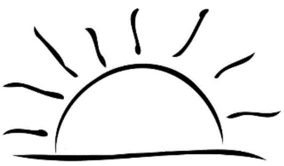
De lente met nesten in het hout.

Vredeleen, mijn vriendin, eindelijk is het zachte weer daar.
Vriendin Vredeleen, in het mooie bos speels een vink.
In het struikgewas, in het lieflijke leger van een hinde, in het zingende woud.
Vredeleen! Vredeleen! Met je mooie vingers, een wilde roos, verliest bloed:
na verloop van tijd wordt feesten verveling.



De flamingo is aan het spelen; flamingo maakt een knoop met een beetje stro.
De zwaluw tjilpt onder een luifel van kreupelhout, helder en guitig, die Vlaamse gaai,
op de ruwe schors van de zilverberk, paradeert de stengel van een uitloper.
In de boomgaard, in het ochtendgloren, druppelt een ontdooide pomp.
Men ziet een glimmende bek die hartstochtelijk trilt van de heldere tonen
en, in een gouden wijnstok, gedragen door de oude poort,
verrassen we het strijdgewoel van mussen.
Rond de moestuin spannen zich snoeren; de taxus staat triest aan de horizon
en zwaar en langzaam boven het zwerk vliegen kraaien.





Een meer kabbelt aan serene kusten en, wanneer de avond valt,
spiegelt het water zich aan het gif van perfide capriolen.
En, als de avond valt, wanneer het paarse van de zonsondergang speelt,
dan doet de lucht haar wateren blozen.
In de weerspiegeling van het water danst de schaduw van een eenhoorn.
Overal klinken schreeuwen! Overal weergalmen geluiden!

Een ligplaats wordt aangemeerd... een boot komt aan… zeilers
gooien hun ankers op de oever...
Overal klinken schreeuwen! Overal weergalmen geluiden!
In het maanlicht mijn vriend Pierrot...
In het maanlicht mijn vriendin Vredeleen...
In het maanlicht mijn vriend Pierrot, leen me een pen om poëzie te schrijven.