Dynamic Testing of Reading and Writing: Analyzing the effect of training in relation to potential (gender) differences between children with and without dyslexia

Timmers, Hester

Universiteit Leiden
Psychologie
Faculteit der Sociale Wetenschappen

Dynamic Testing of Reading and Writing

Analyzing the effect of training in relation to potential (gender) differences between children with and without dyslexia

Hester Timmers

Master Thesis Child & Adolescent Psychology
Date: 25th of June 2021
Student number: s2011514
Supervisor: Dr. B. Vogelaar
Second reader: Prof.dr. W.C.M. Resing

Index

Abstract

This study investigated the effect of a dynamic test in reading and writing compared to a static test for 87 children in elementary schools in the Netherlands (aged 7-9 years old). Static tests measure the knowledge a child already has at the moment of testing, while on the other hand in a dynamic test children are provided with feedback, prompts or training in order to demonstrate more of their learning potential. Dynamic tests are especially useful for children with intellectual disabilities or learning impairments. Children with and without the diagnosis of dyslexia were allocated to either the experimental or the control condition. Children in the experimental condition of this study received a training in between pretest and posttest, whereas those in the control condition received the training after the posttest. Results showed a positive effect of training on the subtests prosodic awareness, spelling sounds and verbs and context dependent spelling, meaning that the group who received the dynamic test showed more progress from pretest to posttest in these subtests than the group who received the static test. This difference in progress did not apply to the subtest phonemic awareness. No significant gender differences and differences in performance between the dyslexic and the non-dyslexic children were found.

1. Introduction

Testing is a crucial process in any educational setting. In addition to providing an indication of the current academic level, it can also be used as a measure to decide whether previous instruction has had the intended effect (Wiliam, 2010).

In educational settings, there are several ways to conduct psychoeducational testing. Most often a static test is being used, where the manual only uses prescribed instructions and no further feedback is given to the child (Resing, Elliott & Vogelaar, 2020). These kinds of tests measure the knowledge a child already has at the moment of testing. Static testing does not provide information on what interventions will be most suitable to facilitate future learning and transfer (Jeltova, 2007). Most of the standard testing instruments used in educational settings have a static nature, which means that a child is only told what to do and is not receiving any feedback (Resing et al., 2020). The main focus of static tests is on the performance outcomes.

Another way of conducting psychoeducational testing is through dynamic testing. This approach to testing is considered to be more about the learning process, where children are explicitly provided with feedback, prompts, or training intended to enable them to demonstrate their capacity for progress in task solving (Resing et al., 2020). In a dynamic test, an individual gets the best possible guidance in order to demonstrate more of their learning potential (Dörfler et al., 2009).

Most research into dynamic testing has been conducted utilizing inductive reasoning tasks. Dynamic testing can, however, also be applied to scholastic domains. Therefore in the current study the effect of dynamic testing in the reading and writing domain was investigated in primary education in the Netherlands. This may provide starting points for the development of didactic interventions focused on the improvement of reading and writing skills.

1.1 Dynamic testing

One of the founders of dynamic testing procedures is Lev Vygotsky. He developed the social constructivism theory (Vygotsky, 1978). Children acquire knowledge and skills that they have not mastered yet with the assistance of an adult or a more knowledgeable other person, in the form of scaffolding. This learning process takes place in the zone of proximal development in which two levels of development in the child can be conceptualized: the child's actual level of development and the child's potential level of cognitive development, which is the level of problem-solving a child can reach with the assistance of someone else (Resing et al., 2020). Dynamic testing takes place in the zone of proximal development (Resing et al., 2020). The test leader guides the child through the testing process, helps to solve tasks and teaches the child how to solve these tasks independently in the future. Dynamic testing can be defined as an umbrella concept, using a variety of testing forms that incorporate feedback, hints or training into the process, aims to measure progress in tasks and in doing so provide an indication for learning (Resing et al., 2020).

Swanson and Lussier (2001) concluded in their meta-analysis on dynamic testing that dynamic testing procedures substantially improve testing performance compared to static testing conditions. Campione and Brown (1987) indicated that predictions based on initial performance significantly underestimated what children could achieve with minimal assistance. Thus, dynamic measures seem to provide more insights into the competence of a child than static tests.

Many dynamic testing procedures have been developed within the context of students with intellectual disabilities, learning difficulties, or socio-cultural disadvantages, and these are precisely the students for whom the level of effectiveness in the implementation of dynamic assessment seems to be the greatest (Navarro & Lara, 2017).

1.2 Testing reading and writing in an educational setting

In primary school children further develop their language skills and learn the processes of reading and writing. The first step in children's language learning is the phonological development: the mastery of the sound system of their language (Siegler et al., 2017). Phonological awareness refers to a child's ability to detect and manipulate the component sounds that comprise words at different grain sizes (Goswami, 2008). Phonological awareness was found to be a significant predictor of reading achievement even after controlling for other factors such as age, IQ and mother's educational level (Goswami, 2008). While reading, words can be identified in two main ways: phonological recording and visually based retrieval (Siegler et al., 2017). Correct use of phonological recording increases the associations between words' visual forms and their sounds, which in turn allows greater use of visual based retrieval. Consistent with this view, children who are better at phonological recording stop using that approach earlier because their past success enables them to shift more rapidly to visually based retrieval (Siegler et al., 2017). Part of phonological awareness is phonemic awareness, which refers to the ability to distinguish one phoneme from another, which means grouping nonidentical but similar physical sounds as being the same phoneme (Goswami, 2008). The beginning of word reading consists of step by step decoding of written words by combining letters into phonemes, which in turn are combined into words (Coltheart et al., 2001).

Writing supports the integration of important language and emergent literacy skills that lay the foundation for children's reading skills (Gerde et al., 2012). It consists of different components, such as forming letters, spelling words and using correct capitalization and punctuation. There are several ways to write down words: non-lexical or lexical spelling. Lexical spelling is applied when sounds are translated into letters (Kohnen et al., 2015). Non-lexical spelling, on the other hand, involve sounds that can be spelled in more than one way.

It is therefore often impossible to spell words correctly by just relying on the translation of sounds into letters. Accurate spelling of these kinds of words relies on lexical knowledge, such as orthographic long-term memories for words (Kohnen et al., 2015). Prosodic awareness indicates the intonation, stress and timing of words (Gutiérrez-Palma et al., 2019). Especially the lexical stress in words, meaning the accent in pronunciation, is a relevant aspect of learning word spelling (Gutiérrez-Palma et al., 2019). Prosodic cues, for instance changes in duration and stress, also carry important information for the reading process about how sounds are ordered into words when the words are multisyllabic (Goswami, 2008). Phonological awareness is also a critical element in the writing process (Vernon & Ferreiro, 1999). The development of reading and writing is best predicted by kindergarten measures of early literacy, for example phonemic awareness and grapheme-to-phoneme knowledge (Schaars et al., 2017).

In primary education in the Netherlands schools are required to track the academic development of their students in order to evaluate and to adapt the teaching method (Ministerie van Onderwijs, Cultuur en Wetenschap, 2019). Schools are allowed to determine for themselves how exactly they track the development of students, however, it is compulsory to take a final test and to use a student tracking system. The most used testing method in the Netherlands is called Cito, which is a teaching method independent, standardized student tracking system (Cito, 2020). The performance and development of reading and writing skills are measured through tests focusing on technical reading, reading comprehension and word dictation (Cito, 2020). Reading comprehension monitoring is the process of keeping track of someone's understanding of a verbal description or a text (Siegler et al., 2017). This can be tested by giving children a text to read and have them answer questions about it. Technical reading is measured by evaluating the speed and accuracy with which words or written texts are decoded. Children have to read words or short stories as fast and accurately as possible.

Writing skills are tested by means of a dictation of certain words which the children have to write down accurately.

1.3 Domain specific dynamic testing of reading and writing

Most research into dynamic testing has been conducted through inductive reasoning tasks, however there has recently been a growing interest into dynamic testing of scholastic domains. Caffrey and colleagues (2008) investigated the predictive validity of dynamic testing in a review and they found the following results for dynamic testing in the reading and writing domain: Over and above both traditional achievement and cognitive tests, dynamic tests uniquely predict future achievement in phonemic and phonological awareness, reading achievement and verbal achievement. Therefore, a dynamic test of reading and writing appears to be an interesting asset in this domain. Dörfler and colleagues (2009) suggest that the various processes involved in reading provide many opportunities for feedback-based interventions which may potentially enhance learning and understanding. An important aspect in the practice of writing is the use of scaffolding by the teacher, through providing hints and prompts to support children to write independently (Gerde et al., 2012). In dynamic testing prompts and hints are involved in the testing process, thus a dynamic measure of writing seems to be a great fit.

The research on dynamic testing in the reading and writing domain has mostly been conducted in kindergarten utilizing dynamic measures as a screening tool or as a predictor for future achievement (Cho, Compton, & Josol, 2020; Gellert & Elbro, 2018; Petersen, Allen, & Spencer, 2014; Sittner Bridges & Catts, 2011). As mentioned before, phonological awareness is a crucial part of both learning reading and writing (Vernon & Ferreiro, 1999). Phonological awareness skills begin to develop at preschool ages and support reading skills during school ages (Turan & Gul, 2008). A dynamic screening measure of phonological awareness in

kindergarten was able to predict end of year reading achievement (Sittner Bridges & Catts, 2011). It appeared that the dynamic measures were more accurate than static measures in identifying children who would later develop reading difficulties (Sittner Bridges & Catts, 2011; Cho et al., 2020; Swanson, 2011). Also, a dynamic test of reading difficulties has a higher predictive and incremental validity in relation to the use of strategies for students with intellectual disabilities compared to students with socio-cultural disadvantages (Navarro & Lara, 2017).

1.4 Dyslexia and Dynamic Testing

Learning difficulties related to reading and writing are in some cases diagnosed as dyslexia. The literal translation of the Greek words 'dys' and 'lexia' is difficulty with words (Payne & Turner, 1998). This can be seen in children who have difficulties in reading, writing or spelling, which are not typical of their general level of performance (Payne & Turner, 1998). A definition of dyslexia is as follows: *"persistent literacy learning difficulties in otherwise typically developing children, despite exposure to high quality, evidence-based literacy instruction and intervention, due to an impairment in the phonological processing skills required to learn to read and write"* (Tunmer & Greaney, 2009, p. 232). Dyslexia stems primarily from a weak ability to discriminate between phonemes, which leads to great difficulty mastering the letter-sound correspondences used in phonological recording (Schaadt & Männel, 2019; Siegler et al., 2017). Teaching children with dyslexia to use strategies that enhance phonological recording appears to be helpful (Siegler et al., 2017).

The prevalence of dyslexia is about 7% in the population (Peterson & Pennington, 2012). There is a relatively small but significant male predominance (1.5-3:1). It remains an inconclusive discussion whether boys are more likely than girls to have reading disabilities and why that would be. In the writing domain, gender differences have been found in

children and adults with dyslexia (Berninger et al., 2008). They were found in measures of automatic letter naming and written expression in children with dyslexia and in those measures plus two measures of spelling in adults with dyslexia. In the dyslexic group, there appeared to be no gender differences in the reading domain. In the typically developing population however, gender differences have been found in the writing as well as in the reading domain (Adams & Simmons, 2018; Price-Mohr & Price, 2016). Even in second grade of primary school, gender ratios provide unbiased estimates of male vulnerability for reading impairment (Quinn & Wagner, 2013). In international studies including 40 different countries, girls outperformed boys in reading comprehension in every participating country (Mullis et al., 2007).

Considering the fact that many dynamic testing procedures have been developed within the context of students with intellectual disabilities and learning difficulties, it might be interesting to see the effect of a dynamic test in reading and writing in children with developmental dyslexia. Cho and colleagues (2020) have already proved that a dynamic assessment can help identify reading disabilities such as dyslexia. They found an improvement of word reading but not arithmetic skills, supporting the characteristics of dyslexia: difficulties in reading and writing, but no deficits in broader domain general cognitive abilities (Tunmer & Greaney, 2009). Reading disabled children that received a dynamic testing intervention made significantly more progress than their typically developing peers who did not receive an intervention during the same period (Aravena et al., 2016).

1.5 The current study

This study aimed to investigate the effect of a dynamic test of reading and writing in primary education for both typically developing children as well as for children with dyslexia. The main research question in this study was: Is there a difference in progress in

reading and writing between children who receive training (a dynamic test) compared to children who do not receive any training (a static test)? The expectation was that a dynamic test would lead to more improvement in both the reading and writing domain, since the child is able to learn from the training process through the feedback and prompts that are given (Resing et al., 2020). In this study, the various aspects of reading and writing are divided into the following components: phonemic awareness, prosodic awareness, spelling sound and verbs and context dependent spelling.

The second research question was: Are there any differences between dyslexic and non-dyslexic children with regard to their performance on a dynamic test of reading and writing? It was expected a) that both groups receiving the training in the test would show more progress from pretest to posttest compared to the groups who did not receive the training, b) that the group of non-dyslexic children would gain the highest scores overall, since children with dyslexia have difficulties in reading and writing (Payne & Turner, 1998; Siegler et al., 2017; Lyon et al., 2003) and c) that despite the expected lower scores for the dyslexic group, these children might show more progress on the posttest after receiving the training compared to the typically developing children in the training condition. In the study of Donovan (2019) children with dyslexia appeared to have a narrower choice of strategies available to them. He suggested that they may benefit more than typically developing peers from a dynamic spelling strategy training. Additionally, Navarro and Lara (2017) suggested that the effect of implementation of dynamic assessment seems to be the greatest for, among others, students with learning difficulties.

The final research question of this study was: Is there a gender difference in performance on a dynamic test of reading and writing? Even though there is a male predominance in the prevalence of developmental dyslexia and girls tend to outperform boys in reading and writing skills, the explanation of this phenomenon remains unclear (Adams &

Simmons, 2018; Quinn & Wagner, 2013; Mullis et al., 2007). It was expected that there would be a gender difference in the performance on the reading and writing assignments in a dynamic test: Girls would probably outperform boys in all of the different subtests of the dynamic test of reading and writing.

2. Methods

2.1 Participants

This study included 87 children selected from elementary schools in the Netherlands. A total of 34 boys and 53 girls with a mean age in months of 99.4 (8.283 years) (range 7 to 9 years, SD 9.461 months) participated. They were in grade 2, 3 and 4 according to the international grading system. The language in which the research was conducted is Dutch. Written informed consent for the participation of children in this study was obtained from the parents through information letters.

2.2 Design

The design in this study was a pretest-training-posttest control group design with randomized blocking. Participants were blocked by school, classroom, gender, age, as well as the results of the IDS-2 IQ screening test. The dynamic test of reading and writing used in the current study is called the EPALE, which is derived from the Spanish title “Test de evaluación del potencial de aprendizaje para la lecto-escritura”, which translates into learning potential assessment test for literacy (Mata & Serrano, 2019).

Table 1

Experimental Design (EPALE is the dynamic test of reading and writing)

	IDS-2	Pretest EPALE	Training EPALE	Posttest EPALE	Training EPALE
Experimental	X	X	X	X	
Control	X	X		X	X

2.3 Materials

Materials for this study consisted of the IDS-2 IQ screening test and The EPALE-NL, (Mata & Serrano, 2019; Grob & Hagmann-von Arx, 2018). Based on the IDS-2 IQ screening, participants were divided over the experimental and control condition.

IDS-2 IQ screening

The IDS-2 IQ screening is a global cognitive assessment for children designed to identify their knowledge, strengths and relevant areas for development (Grob & Hagmann-von Arx, 2018). It gives an indication of the intelligence quotient of the child. For the intelligence screening merely two subtests were being used: matrices reasoning and category naming. The duration was about ten minutes. The IDS-2 IQ screening is a reliable instrument with a Cronbach's alpha α of .90 (Egberink, Leng, & Vermeulen, 2020).

EPALE-NL

The EPALE-NL is a Dutch translation of the Spanish EPALE (Mata & Serrano, 2019; de Vreeze-Westgeest, 2020). It is a dynamic test of reading and writing, consisting of four different subtests. First there is a pretest, then follows a training/mediation, concluded by a posttest. Mediation is started when a child makes one error or more on the pretest. After mediation a posttest is conducted with the aim of possibly seeing progress in relation to the pretest. The duration of EPALE-NL is around 1.5 hours. The reliability of the Spanish EPALE is good, except for the subtest context dependent spelling with a Cronbach's alpha α of .62. The reliability of the other subtests was good with phonemic awareness up to α .89, prosodic awareness up to α .87 and spelling sounds and verbs up to α .83. A description of the subtests follows:

Phonemic awareness

This subtest measures if a child is capable of identifying and synthesizing the presented phonemes into words. Mediation consisted of creating sentences with words

presented on cards, dividing words into syllables, throwing a dice with vowels and consonants and merging them into non-existing words.

Prosodic awareness

This subtest measures if a child is capable of telling the emphasis in pseudowords. Each word is presented twice, with an interval of seven seconds. The child puts a cross in one of the three columns indicating where the emphasis of the word was. Mediation consisted of tapping on the table, dividing words into syllables and putting a red plug on the emphasis of the existing- and pseudowords.

Spelling of sounds and verbs

This subtest measures the sound-symbol knowledge and correspondent spelling rules in words and sentences. The words are presented in a dictation. Mediation consisted of finding similarities between words, discussing spelling rules and finishing sentences.

Context dependent spelling

This subtest measures whether a child can determine how to write down a homophone from the context of the sentence. The child is presented with a sentence with a missing word and has to decide which of three words fits. One of the options is the right answer, the other is a phoneme and the last one is a distraction. This is an exercise measuring performance in non-lexical spelling. Mediation consisted of words couples that sound the same, only with a different meaning and a different kind of spelling. These words were presented accompanied by drawings to help the child associate the words and make it easier to remember the spelling.

2.4 Procedure

This study consisted of two individual sessions per child. In the first session a pretest was taken in about 30 minutes. This consisted of the IDS-2, among some other static reading

and writing tests. Based on the results of the IQ-screening in the IDS-2, the children were divided among the experimental and control conditions of the study. The division was made per class in order to avoid a classroom effect. In the second session EPALE-NL was carried out. The participants in the experimental condition followed the original format of the dynamic test, which means a pretest, followed by a training, and concluded with the posttest. Participants in the control condition received the training as well, however they first completed both the pretest and the posttest. In this way it was possible to compare the effects of a dynamic test with the effects of a static test, without this being detrimental to children in the control condition.

3. Results

3.1 Preliminary analyses

In order to explore potential differences between the participants in the two conditions in relation to IQ-score, age, and pretest score on the several subtests, six separate one way ANOVAs were conducted prior to further analyzing the data. In these analyses, condition (training vs no training) was used as the factor variable and the dependent variables were (raw) IQ-score, age and pretest score on phonemic awareness, prosodic awareness, spelling sounds and verbs and context dependent spelling. No significant differences were found for age ($F(1,85) = .213$, $p = .645$), IQ ($F(1,85) = .397$, $p = .530$), pretest phonemic awareness ($F(1,85) = .002$, $p = .962$), pretest prosodic awareness ($F(1,85) = .004$, $p = .949$), pretest spelling sounds and verbs ($F(1,85) = 1.464$, $p = .230$), and pretest context dependent spelling ($F(1,85) = .000$, $p = .991$). The descriptive statistics are presented in Table 2.

Subsequently, potential differences between the participants with or without the diagnosis of dyslexia were investigated in relation to these variables by mean of six additional one way ANOVAs. They were repeated with diagnosis as the factor variable. Likewise, no significant differences were found for IQ ($F(1,85) = 2.117$, $p = .149$), pretest phonemic awareness ($F(1,85) = 1.\ 074$, $p = .303$), pretest prosodic awareness ($F(1,85) = .310$, $p = .579$), pretest spelling sound and verbs ($F(1,85) = .953$, $p = .332$), and pretest context dependent spelling ($F(1,85) = 1.825$, $p = .180$). For age, on the other hand, a significant group difference was found ($F\ (1,85) = 92,283$, $p < .001$). The non-dyslexic group (M= 109.06) was older than the dyslexic group (M= 94.89). Therefore, an extra analysis was conducted to investigate the correlation between age and pretest scores. Only minor correlations were found for the subtests spelling sounds and verbs and context dependent spelling. The results of this analysis are presented in Table 3.

Another point of interest was whether the distribution of boys and girls was equal across condition and diagnosis group. A Chi square analysis was used to investigate this. The results demonstrated no significant differences between the gender distribution across the conditions ($\chi^2(86) = .276$, $p = .599$) and across the diagnosis group ($\chi^2(86) = .748$, $p = .387$).

Table 2

Descriptive Statistics for Age, IQ and all of the separate Pretest Scores

	Training (Experimental)	No training (Control)	Dyslexia (Diagnosis)	No dyslexia (No diagnosis)
Age in months				
M	100.42	99.48	94.89	109.06
SD	10.09	8.90	7.05	5.65
N	43	44	31	56
IQ-score IDS2				
M	24.70	24.09	23.88	25.32
SD	4.49	4.49	4.53	4.29
N	43	44	31	56
Pretest phonemic awareness				
M	12.12	12.14	11.84	12.29
SD	1.92	1.96	1.77	2.01
N	43	44	31	56
Pretest prosodic awareness				
M	9.28	9.20	9.00	9.68
SD	5.32	5.56	5.39	5.50
N	43	44	31	56
Pretest sounds and verbs				
M	13.88	12.52	12.79	13.94
SD	5.60	4.87	5.44	4.91
N	43	44	31	56
Pretest context dependent				
M	11.42	11.41	10.98	12.19
SD	3.91	4.18	4.09	3.84
N	43	44	31	56

Table 3

Correlations between Age in months and the various Pretest Scores.

	Pretest phonemic awareness	Pretest prosodic awareness	Pretest spelling sounds and verbs	Pretest context dependent spelling
Age in months				
Pearson	-.010	.075	.311	.235
P	.924	.489	.003	.028

3.2 Analyzing potential differences in progress from pretest to posttest

Considering that the preliminary analyses revealed no major differences between the distribution across age, IQ, gender and pretest scores, a repeated measures MANOVA was conducted to analyze the differences in progression from pretest to posttest across the different conditions, diagnosis group and gender. The within subjects factor was session (pretest/posttest) and the between subject factors were condition (training/no training), diagnosis (dyslexia/no dyslexia) and gender. The scores on the subtests phonemic awareness, prosodic awareness, spelling sounds and verbs and context dependent spelling were the dependent variables.

First, the multivariate data were analyzed. As can be seen in Table 4, the results showed significant multivariate effects for session and session * condition. The multivariate effect for session * gender was almost significant, with a trend toward significance ($F(4,76) = 2.306$, $p = .066$). Session * dyslexia showed a nonsignificant effect. Due to insignificant multivariate effects, the three-way interaction effects were not further interpreted. The multivariate results are presented in Table 4.

Table 4
Multivariate within subjects data from the Repeated Measures MANOVA

	F	df	df(error)	*p*	η^2_p
Session	13.526	4	76	< .001	.416
Session * Condition	8.289	4	76	< .001	.304
Session * Dyslexia	1.482	4	76	.216	.072
Session * Gender	2.306	4	76	.066	.108
Session * Condition * Dyslexia	2.892	4	76	.028	.132
Session * Condition * Gender	.415	4	76	.798	.021
Session * Dyslexia * Gender	.803	4	76	.527	.041
Session * Condition * Dyslexia * Gender	1.048	4	76	.388	.052

Additionally, the univariate effects were analyzed. These results will be discussed below per research question. The univariate results are presented in Table 5.

Table 5

Univariate data from the Repeated Measures MANOVA

	F	df	df(error)	*p*	η^2_p
Session					
Phonemic	38.633	1	86	< .001	.328
Prosodic	9.179	1	86	.003	.104
Sounds and Verbs	6.208	1	86	.015	.073
Context dependent	8.139	1	86	.006	.093
Session * Condition					
Phonemic	.378	1	86	.541	.005
Prosodic	25.875	1	86	< .001	.247
Sounds and Verbs	3.768	1	86	.056	.046
Context dependent	5.713	1	86	.019	.067
Session * Dyslexia					
Phonemic	2.145	1	86	.147	.026
Prosodic	2.828	1	86	.097	.035
Sounds and Verbs	1.150	1	86	.287	.014
Context dependent	.424	1	86	.517	.005
Session * Gender					
Phonemic	.777	1	86	.381	.010
Prosodic	4.913	1	86	.030	.059
Sounds and Verbs	2.875	1	86	.094	.035
Context dependent	1.434	1	86	.235	.018
Session * Condition * Dyslexia					
Phonemic	.102	1	86	.751	.001
Prosodic	3.007	1	86	.087	.037
Sounds and Verbs	4.294	1	86	.042	.052
Context dependent	2.858	1	86	.095	.035
Session * Condition * Gender					
Phonemic	.070	1	86	.792	.001
Prosodic	.768	1	86	.383	.010
Sounds and Verbs	.073	1	86	.788	.001
Context dependent	.816	1	86	.369	.010
Session * Dyslexia * Gender					
Phonemic	.049	1	86	.825	.001
Prosodic	1.235	1	86	.270	.015
Sounds and Verbs	1.434	1	86	.235	.018
Context dependent	.659	1	86	.407	.009
Session * Condition * Dyslexia * Gender					
Phonemic	.082	1	86	.775	.000
Prosodic	.002	1	86	.968	.000
Sounds and Verbs	1.737	1	86	.191	.022
Context dependent	2.052	1	86	.156	.025

3.2.1 Effect of training on potential differences in progress from pretest to posttest

In order to investigate whether there was a difference in reading and writing progress between children who received training compared to children who did not receive any training, results were analyzed per subtest. All univariate effects of session were significant as can be seen in Table 5. This means that every group showed improvement from pretest to posttest, regardless of the condition. Univariate results indicated that the interaction effects for session * condition were significant for the subtests prosodic awareness and context dependent spelling, indicating a difference in progress on these specific subtests between the group who received training and the group who did not receive any training. The mean scores revealed that higher scores on the posttests for prosodic awareness and context dependent spelling were obtained by the group who received training, suggesting a positive effect of training on the results. The interaction effect for the subtest phonemic awareness was not significant, indicating no significant difference in progress for this subtest between the group who received training and the group who did not. The interaction effect for the subtest spelling sounds and verbs was also not significant, however there was a trend towards significance.

3.2.2 Effect of training on potential differences in progress from pretest to posttest between children with and without dyslexia

In addition, potential differences between dyslexic and non-dyslexic children with regard to their performance on a dynamic test of reading and writing were investigated. The univariate results for session and session * condition were already discussed above. The multivariate within subjects effect for session * dyslexia was not significant, therefore no differences in progress were found between the group of children with dyslexia and the group without the diagnosis of dyslexia. The multivariate between subjects effect for dyslexia revealed no significant differences in the height of scores between children with and without dyslexia ($F(4,76) = 1.725$, $p = .153$). Due to the insignificant multivariate interaction effect for session * dyslexia, the three-way multivariate interaction effect of session * condition * dyslexia was not interpreted and discussed anymore (see Table 5). A visual representation of the scores on the various subtests can be found in Figure 1.

3.2.3 Effect of training on potential gender differences in progress from pretest to posttest

After detecting a trend towards significance in the multivariate analyses, the mean scores were analyzed and these suggested that girls obtained higher scores on the subtests spelling sounds and verbs and context dependent spelling regardless of the condition they were in. On the subtest prosodic awareness boys showed more progress from pretest to posttest after receiving the training compared to girls, even though in the control condition girls obtained higher scores. The between subjects effect for gender revealed no significant differences in the height of scores between boys and girls ($F(4,76) = 1.359$, $p = .256$). Because of the insignificant multivariate effects, the univariate results were not discussed and interpreted. See Table 5 for the results of the univariate analysis. A visual representation of the different scores for boys and girls on the subtests is shown in Figure 2.

Figure 1

Line graph demonstrating the effect of training on potential differences in progress between the children with and without dyslexia

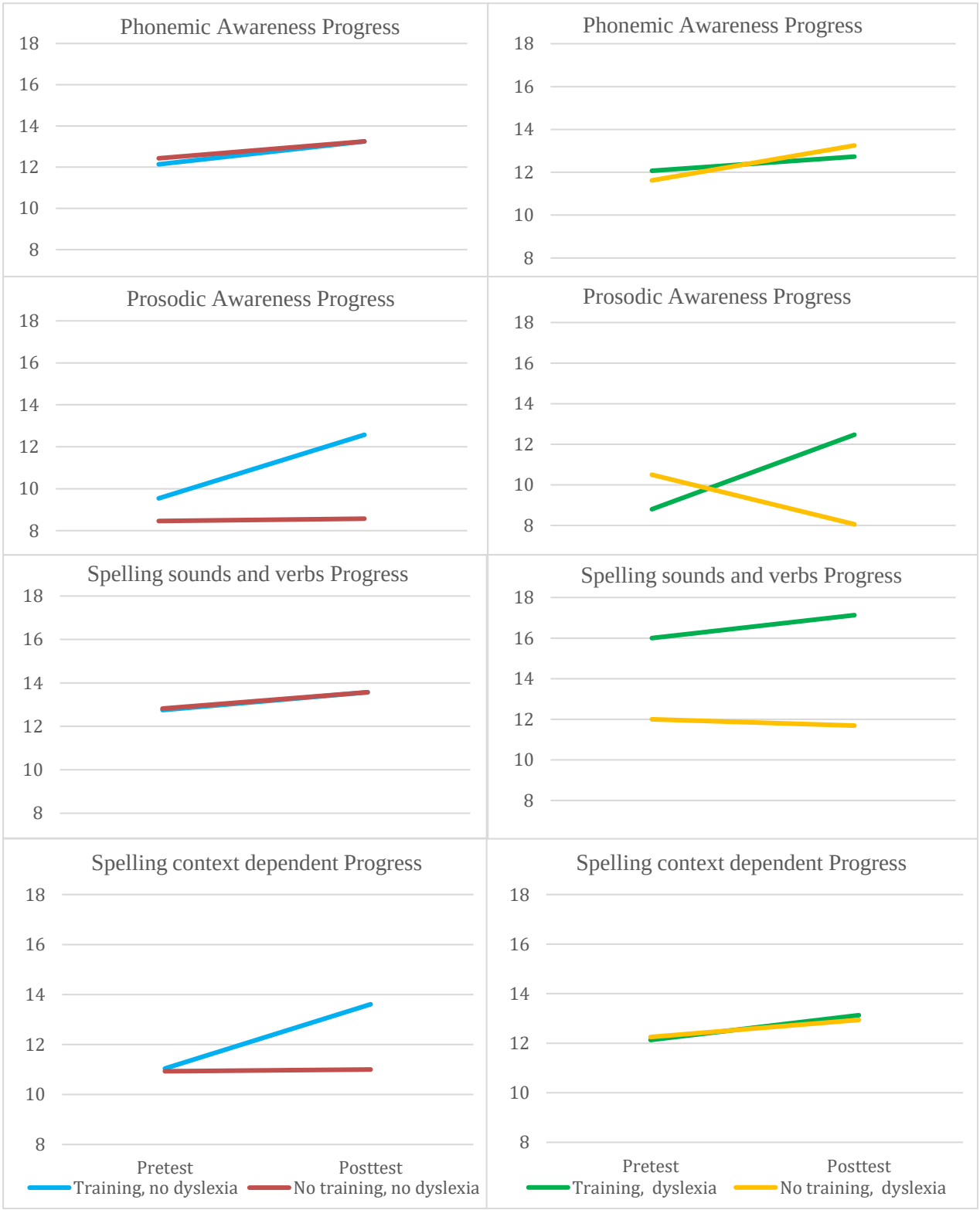
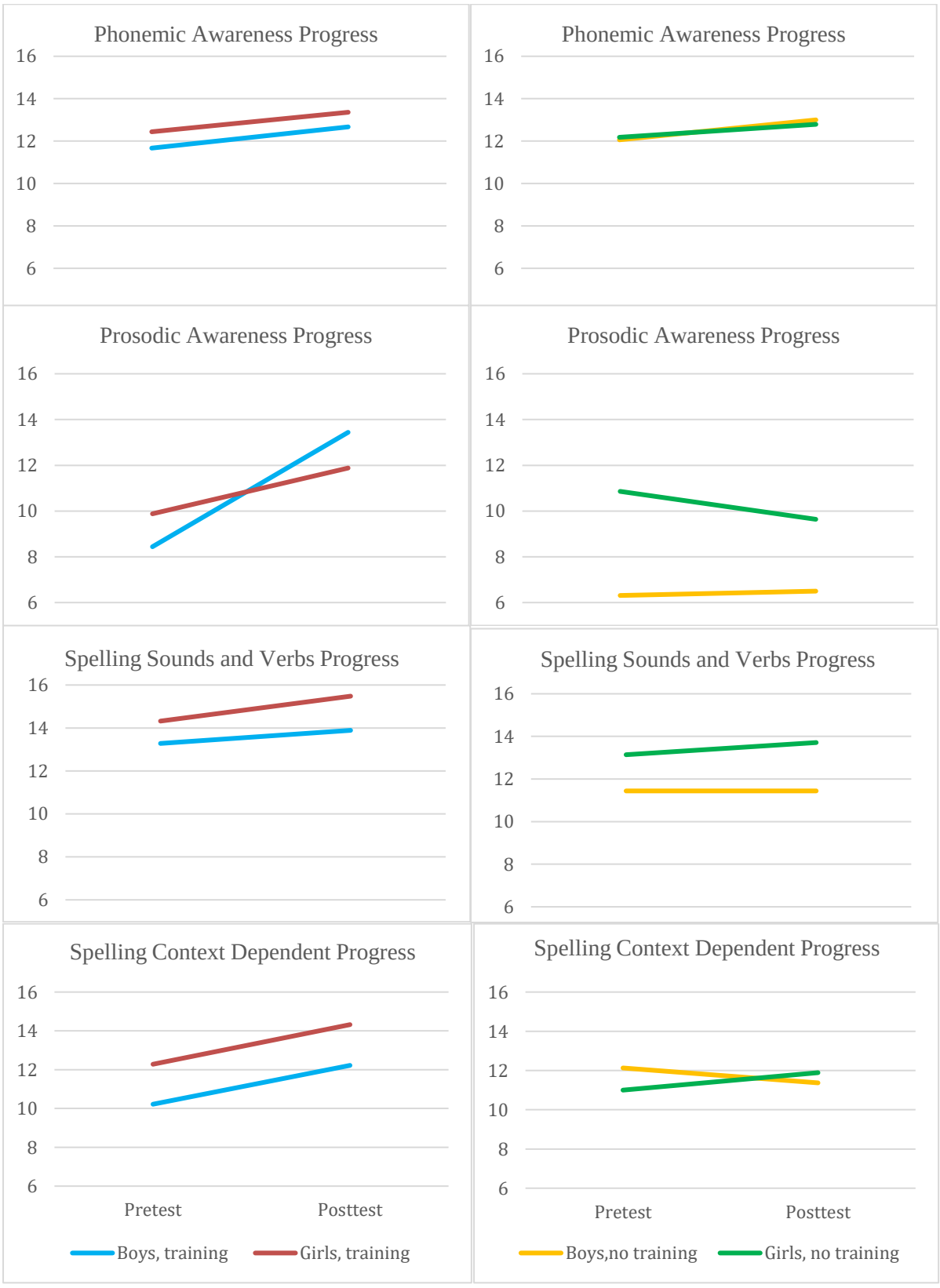


Figure 2

Line graph demonstrating the effect of training on potential differences in progress between boys and girls



4. Discussion

The purpose of the current study was to investigate the effect of a dynamic test in the reading and writing domain, compared to the effect of a static test in this specific area.

Results indicated a positive effect of training on progress from pretest to posttest. The group who received training (dynamic test) demonstrated a larger amount of growth compared to the group who did not receive training (static test) in the subtests prosodic awareness, spelling sounds and verbs and context dependent spelling. This is consistent with the existing literature about the effect of dynamic testing (Resing et al., 2020; Swanson & Lussier, 2001). Unlike expectations, for the subtest phonemic awareness there was no difference in progress between the group who received training and the group who did not. Next, the effect of a dynamic test in specific groups, who usually tend to gain lower scores in this scholastic domain, was addressed. Contrary to expectations, no differences in performance were found between the dyslexic and the non-dyslexic children. Children with dyslexia have learning difficulties related to reading and writing, generally manifesting itself in underperforming in this domain compared to their general level of performance (Payne & Turner, 1998). In addition, gender differences in performance were also not found. Girls usually tend to outperform boys in the domain of reading and writing (Adams & Simmons, 2018; Price-Mohr & Price, 2016; Mullis et al., 2007).

A possible explanation for the unsupported hypothesis about dyslexia is that the participants in this study might have been high-functioning dyslexics, explaining that their scores were higher than expected for dyslexics. In the current study, the dyslexic children were all following dyslexia treatment. In adults with high-functioning dyslexia, a potential factor for compensating reading comprehension is morphological processing (Law et al., 2017). The children with dyslexia in this study might have been good at morphological processing, due to the therapy they were already following. The performance level of

dyslexics can be measured through various tests of reading and writing, for instance the One-minute reading test (Brus & Voeten, 1997), Klepel-R (van den Bos et al., 2019), CELF Phonological Awareness test (Kort & Compaan, 2008) and Two Minutes Spelling (Kort et al., 2005). Moreover, some dynamic testing procedures have been specifically developed for students with learning difficulties, possibly explaining the absence of a group difference for dyslexics vs non-dyslexics in all the different subtests of the dynamic test of reading and writing. Dynamic tests give children with a learning problem the opportunity to show their capabilities (Navarro & Lara, 2017). Another possibility is that the amount of growth on the dynamic test was moderated by working memory performance, since working memory plays a central role in language comprehension and writing (Swanson, 2011). The dynamic test of reading and writing in the current study focusses on specific linguistic aspects of reading and writing such as phonological awareness, instead of a more complete framework of the neuropsychological functioning underneath (Cainelli & Bisiacchi, 2019). Knowledge about how children use their executive functions, for instance working memory, can contribute to understanding and improving their reading and writing skills (Berninger et al., 2017). A helpful future addition to the dynamic test of reading and writing could be a task measuring working memory, for example a digit span.

The absence of the expected gender differences in performance might be explained by the exercises that were being used in this study to measure reading and writing skills. Girls usually outperform boys in specific aspects of reading and writing, such as reading comprehension and verbal fluency (Price-Mohr & Price, 2016; Reilly et al., 2019). This was not measured in the current study. In addition, the training in reading and writing is more in line with boys' learning strategies. The mediation in the dynamic test is largely a combination between aural and visual material. This multimodal information processing approach is preferred by male students, while girls have a preference for a single mode of information

presentation (Wehrwein et al., 2007). The manner in which materials were provided in the training may have served as compensation on the gender differences in performance on the reading and writing domain, because the training was more suitable for boys.

4.1 Limitations of the current study

There are several limitations to the current study, leading to recommendations for future research on this topic. A possible influence on the results of this study was the pandemic situation due to COVID-19 in 2020 and 2021, causing schools to close down for a few months twice within a year. In these two periods, most children had to stay home and follow homeschooling. In some cases, this led to an educational disadvantage in the reading and writing domain, due to a lack of materials or incompetence of parents to support and assist their children with schoolwork at home (Joosten et al., 2020).

Another limitation is that the children in the dyslexic group were recruited through an organization, called OnderwijsAdvies, which gives treatment to children with dyslexia. All dyslexic children had already received dyslexia treatment, possibly causing a higher level of performance compared to dyslexic children who had not yet received treatment. For future research on this topic, it might be interesting to recruit dyslexic children who have been diagnosed recently and have not already received any form of treatment.

Additionally, a group difference for age was found in the preliminary analysis. It appeared that the non-dyslexic group of children was older than the dyslexic group. A possible explanation for this group difference is a number of dropouts due to the COVID-19 restrictions.

One more possible limitation of this study is that the dynamic test of reading and writing does not differentiate enough between the various groups, such as gender and the diagnosis of dyslexia. However, this can also be interpreted as a positive aspect of the dynamic test,

considering all groups get to show their capabilities. Certain differences that were expected in advance, were not found in the results. Perhaps an explanation is that the level of performance was already high on the task, making it difficult to demonstrate full learning potential. In the future, the level of difficulty of the test could be higher.

4.2 Conclusion and Practical Implications

The results of this study indicated a positive effect of dynamic testing compared to static testing. Interestingly, there were no differences in performance found between the dyslexic and the non-dyslexic group nor between boys and girls. This illustrates and supports the importance of dynamic testing in educational settings. It gives children with a learning problem more opportunity to show their capabilities. Therefore, dynamic testing should be incorporated into the educational system in order to gain more insight into the learning potential of all children. A practical obstacle in the appliance of the EPALE in educational settings it that the individual training that was used in the current study is very time consuming. A computerized dynamic test might be a solution for this time intensive issue and it could also contribute to a more standardized dynamic testing procedure. This study provides educational practice with an insight into the effects of dynamic testing in the reading and writing domain. Implementation in educational settings may contribute to a better overview of the learning potential of students and can eventually also be used to determine which educational interventions will be most suitable to facilitate future learning and transfer.

5. Literature

Adams, A. M., & Simmons, F. R. (2018). Exploring individual and gender differences in early writing performance. *Reading and Writing, 32*(2), 235–263. https://doi.org/10.1007/s11145-018-9859-0

Aravena, S., Tijms, J., Snellings, P., & van der Molen, M. W. (2016). Predicting responsiveness to intervention in dyslexia using dynamic assessment. *Learning and Individual Differences, 49,* 209–215. https://doi.org/10.1016/j.lindif.2016.06.024

Berninger, V., Abbott, R., Cook, C. R., & Nagy, W. (2017). Relationships of Attention and Executive Functions to Oral Language, Reading, and Writing Skills and Systems in Middle Childhood and Early Adolescence. *Journal of Learning Disabilities*, *50*(4), 434–449. https://doi.org/10.1177/0022219415617167

Berninger, V., Nielsen, K. H., Abbott, R. D., Wijsman, E., & Raskind, W. (2008). Gender differences in severity of writing and reading disabilities. J*ournal of School Psychology, 46*(2), 151-172.

Brus, B. T., & Voeten, M. J. M. (1997). *Een-minuut-test.* Pearson Benelux B. V.

Caffrey, E., Fuchs, D., & Fuchs, L. S. (2008). The Predictive Validity of Dynamic Assessment. *The Journal of Special Education, 41*(4), 254–270. https://doi.org/10.1177/0022466907310366

Cainelli, E., & Bisiacchi, P. S. (2019). Diagnosis and Treatment of Developmental Dyslexia and Specific Learning Disabilities: Primum Non Nocere. *Journal of Developmental and Behavioral Pediatrics, 40*(7), 558-562.

Campione, J. C., & Brown, A. L. (1987). *Dynamic assessment: an international approach evaluating learning potential.* 82-115. Guilford Press.

Cho, E., Compton, D. L., & Josol, C. K. (2020). Dynamic assessment as a screening tool for early identification of reading disabilities: A latent change score approach. *Reading & Writing, 33*(3), 719-739.

Cito. (2020). *Zin en onzin van toetsen en scores in het primair onderwijs*. https://www.cito.nl/onderwijs/primair-onderwijs/zin-en-onzin-van-toetsen

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review, 108*(1), 204. http://dx.doi.org/10.1037//0033-295x.108.1.204.

de Vreeze-Westgeest, M. (2020). Vertaling EPALE-NL: Leerpotentiële beoordelingstest voor geletterdheid.

Donovan, J. L. (2019). Exploring the efficacy of a dynamic assessment procedure of spelling in children with dyslexia. University College London.

Dörfler, T., Golke, S., & Artelt C. (2009). Dynamic assessment and its potential for the assessment of reading competence. *Studies in educational evaluation (35),* 77-82. doi_10.1016_j.stueduc.2009.10.005

Egberink, I. J. L., Leng, W. E. de, & Vermeulen, C. S. M. (16 juni 2021). COTAN beoordeling 2020, intelligentie en ontwikkelingsschalen voor jongeren IDS2. Bekeken via www.cotandocumentatie.nl

Jeltova, I., Birney, D., Fredine, N., Jarvin, L., Sternberg, R. J., & Grigorenko, E. L. (2007). Dynamic assessment as a process-oriented assessment in educational settings. *Advances in Speech Language Pathology, 9*(4), 273–285. https://doi.org/10.1080/14417040701460390

Gerde, H. K., Bingham, G. E., & Wasik, B. A. (2012). Writing in Early Childhood Classrooms: Guidance for Best Practices. *Early Childhood Education Journal*, *40*(6), 351–359. https://doi.org/10.1007/s10643-012-0531-z

Goswami, U. (2008). *Cognitive development: The learning brain*. Routledge.

Gutiérrez-Palma, N., Valencia Naranjo, N., Justicia-Galiano, M. J., & Carpio Fernández, M. V. (2019). Beyond phonological awareness: Stress awareness and learning word spelling. *Learning and Individual Differences, 74*, 101755. https://doi.org/10.1016/j.lindif.2019.101755

Grob, A., & Hagmann-von Arx, P. (2018). *IDS-2 Intelligentie- en Ontwikkelingsschalen voor Kinderen en Jongeren.*(Nederlandse bewerking door Ruiter, S., Timmerman, M., & Visser L.) Hogrefe Uitgevers BV.

Joosten, K., Jong-van Kempen, B., van Els. A., & Leeuwenburgh-Pronk, W. (2020). De gevolgen van de COVID-pandemie op de fysieke, cognitieve, emotionele en sociale ontwikkeling van kinderen. *Dokters voor kinderen*.

Kohnen, S., Colenbrander, D., Krajenbrink, T., & Nickels, L. (2015). Assessment of lexical and non-lexical spelling in students in Grades 1-7. *Australian Journal of Learning Difficulties, 20*(1), 15-38.

Kort, W. M., & Compaan E. (2008). *CELF-4NL: Fonologisch bewustzijn*. Pearson Assessment and Information BV.

Kort, W., van den Bos, K. P. Lutje van den Spelberg, H. C., van der Wild, S., Schittekatte, S., Vermeir, G., & Verhaeghe, P. (2005). *Dyslexie Screening Test, Nederlandse bewerking.* Harcourt Test Publishers.

Law, J. M., Veispak, A., Vanderauwera, J., & Ghesquière, P. (2017). Morphological awareness and visual processing of derivational morphology in high-functioning adults with dyslexia: An avenue to compensation? *Applied Psycholinguistics*, *39*(3), 483–506. https://doi.org/10.1017/s0142716417000467

Lyon, G. R., Shaywitz, S. E., & Shaywitz, B. A. (2003). A definition of dyslexia. *Annals of Dyslexia*, *53*(1), 1–14. https://doi.org/10.1007/s11881-003-0001-9

Mata, S., & Serrano, F. (2019). *EPALE: Test de evaluación del potencial de aprendizaje para la lecto-escritura.* Granada, Espana: La Paz Charts.

Ministerie van Onderwijs, Cultuur en Wetenschap. (2019, 10 december). *Toetsen in het basisonderwijs*. Onderwerp | Inspectie van het onderwijs. https://www.onderwijsinspectie.nl/onderwerpen/toetsen-in-po#:%7E:text=De%20eindtoets%20is%20geen%20examen,over%20de%20toelating%20van%20eindtoetsen.

Mullis, I.V.S., Martin, M.O., Kennedy A.M., & Foy P. (2007). PIRLS 2006 international report: IEA's progress in international reading literacy study in primary schools in 40 countries. Chestnut Hill, MA: Boston College.

Navarro, J. J., & Lara, L. (2017). Dynamic Assessment of Reading Difficulties: Predictive and Incremental Validity on Attitude toward Reading and the Use of Dialogue/Participation Strategies in Classroom Activities. *Frontiers in Psychology* (8). https://doi.org/10.3389/fpsyg.2017.00173

Payne, T., & Turner, E. (1998). *Dyslexia Parents' and teachers' guides* (third edition). Multilingual Matters.

Petersen, D. B., Allen, M. M., & Spencer, T. D. (2014). Predicting Reading Difficulty in First Grade Using Dynamic Assessment of Decoding in Early Kindergarten. *Journal of Learning Disabilities*, *49*(2), 200–215. https://doi.org/10.1177/0022219414538518

Peterson, R. L., & Pennington, B. F. (2012). Developmental dyslexia. *The Lancet*, *379*(9830), 1997–2007. https://doi.org/10.1016/s0140-6736(12)60198-6

Price-Mohr, R., & Price, C. (2016). Gender Differences in Early Reading Strategies: A Comparison of Synthetic Phonics Only with a Mixed Approach to Teaching Reading to 4–5 Year-Old Children. *Early Childhood Education Journal*, *45*(5), 613–620. https://doi.org/10.1007/s10643-016-0813-y

Quinn, J. M., & Wagner, R. K. (2013). Gender Differences in Reading Impairment and in the Identification of Impaired Readers. *Journal of Learning Disabilities*, *48*(4), 433–445. https://doi.org/10.1177/0022219413508323

Reilly, D., Neumann, D. L., & Andrews, G. (2019b). Gender differences in reading and writing achievement: Evidence from the National Assessment of Educational Progress (NAEP). *American Psychologist*, *74*(4), 445–458. https://doi.org/10.1037/amp0000356

Resing, W., Elliott, J., & Vogelaar, B. (2020). Assessing Potential for Learning in School Children. *Oxford Research Encyclopedia of Education.*

Schaadt, G., & Männel, C. (2019). Phonemes, words, and phrases: Tracking phonological processing in pre-schoolers developing dyslexia. *Clinical Neurophysiology*, *130*(8), 1329–1341. https://doi.org/10.1016/j.clinph.2019.05.018

Schaars, M. M. H., Segers, E., & Verhoeven, L. (2017). Predicting the integrated development of word reading and spelling in the early primary grades. *Learning and Individual Differences*, *59*, 127–140. https://doi.org/10.1016/j.lindif.2017.09.006

Siegler, R. S., Saffran, J., Eisenberg, N., DeLoache, J. S., Gershoff, E., & Leaper, C. (2017). *How Children Develop* (Fifth edition). Worth Publishers

Sittner Bridges, M., & Catts, H. W. (2011). The Use of a Dynamic Screening of Phonological Awareness to Predict Risk for Reading Disabilities in Kindergarten Children. *Journal of Learning Disabilities*, *44*(4), 330–338. https://doi.org/10.1177/0022219411407863

Swanson, H. L. (2011). Dynamic Testing, Working Memory, and Reading Comprehension Growth in Children With Reading Disabilities. *Journal of Learning Disabilities, 44*(4), 358-371.

Swanson H. L., & Lussier C. M. A. (2001). Selective Synthesis of the Experimental Literature on Dynamic Assessment. *Review of Educational Research. 71*(2), 321-363. doi:10.3102/00346543071002321

Tunmer, W., & Greaney, K. (2009). Defining Dyslexia. *Journal of Learning Disabilities*, *43*(3), 229–243. https://doi.org/10.1177/0022219409345009

Turan, F., & Gul, G. (2008). Early Precursor of Reading: Acquisition of Phonological Awareness Skills. *Educational Sciences: Theory & Practice*, *8*(1), 279–284.

van den Bos, K. P., de Groot, B. J. A., & de Vries, J. R. (2019). *Klepel-R.* Pearson Benelux B. V.

Vernon, S., & Ferreiro, E. (1999). Writing Development: A Neglected Variable in the Consideration of Phonological Awareness. *Harvard Educational Review*, *69*(4), 395–416. https://doi.org/10.17763/haer.69.4.p411667586738x0w

Vygotsky, L. S. (1978).*Mind in society: The development of higher psychological processes.* Harvard University Press.

Wehrwein, E. A., Lujan, H. L., & DiCarlo, S. E. (2007). Gender differences in learning style preferences among undergraduate physiology students. *Advances in Physiology Education*, *31*(2), 153–157. https://doi.org/10.1152/advan.00060.2006

Wiliam, D. (2010). Standardized Testing and School Accountability. *Educational Psychologist*, *45*(2), 107–122. https://doi.org/10.1080/00461521003703060