

# ■ Word Decoding Development during Phonics Instruction in Children at Risk for Dyslexia

Moniek M.H. Schaars\* , Eliane Segers and Ludo Verhoeven

*Behavioural Science Institute, Radboud University, Nijmegen, The Netherlands*

In the present study, we examined the early word decoding development of 73 children at genetic risk of dyslexia and 73 matched controls. We conducted monthly curriculum-embedded word decoding measures during the first 5 months of phonics-based reading instruction followed by standardized word decoding measures halfway and by the end of first grade. In kindergarten, vocabulary, phonological awareness, lexical retrieval, and verbal and visual short-term memory were assessed. The results showed that the children at risk were less skilled in phonemic awareness in kindergarten. During the first 5 months of reading instruction, children at risk were less efficient in word decoding and the discrepancy increased over the months. In subsequent months, the discrepancy prevailed for simple words but increased for more complex words. Phonemic awareness and lexical retrieval predicted the reading development in children at risk and controls to the same extent. It is concluded that children at risk are behind their typical peers in word decoding development starting from the very beginning. Furthermore, it is concluded that the disadvantage increased during phonics instruction and that the same predictors underlie the development of word decoding in the two groups of children. © 2017 The Authors  
Dyslexia Published by John Wiley & Sons Ltd.

**Keywords:** genetic risk; early reading development; kindergarten precursors; phonics-based instruction; curriculum-embedded word decoding assessments

## Key Messages

- Children at risk are less skilled in kindergarten measures of phonemic awareness but not in lexical retrieval, grapheme-phoneme knowledge, vocabulary and short-term memory.
- Genetic disadvantage emerges already during the first few months of phonics-based reading instruction.
- Differences between typical and at-risk readers during the first year of instruction are more pronounced in complex word decoding tasks.
- Phonemic awareness and lexical retrieval predict word decoding development in children at risk and controls to the same extent.

Mastering phonological recoding skills is assumed to be the first and most fundamental step in learning to read (e.g. Ehri, 2005; Share, 1995, 2004). After learning to phonologically recode in Grade 1, children incrementally proceed towards a more efficient and advanced way of orthographic processing during reading. Reading development can be predicted from kindergarten precursor

---

\*Correspondence to: Moniek M.H. Schaars, Behavioural Science Institute, Radboud University, Montessorilaan 3, 6525 HR Nijmegen, The Netherlands. E-mail [m.schaars@pwo.ru.nl](mailto:m.schaars@pwo.ru.nl)  
The copyright line for this article was changed on 26 July 2018 after original online publication.

measures (e.g. Kirby, Desrochers, Roth, & Lai, 2008; Landerl et al., 2013; Melby-Lervåg, Lyster, & Hulme, 2012). Children at genetic risk for dyslexia already show problems in these precursor measures, prior to their later reading problems. Up till now, prospective heritability studies devoted attention to the precursors of learning to read, on the one hand, and manifestations of later persistent reading problems, on the other hand (e.g. Eklund et al., 2015; Lyytinen et al., 2006; Moll et al., 2014; Scarborough, 1989; Snowling & Melby-Lervåg, 2016; Van Bergen et al., 2011). However, the development of word decoding in children at risk for dyslexia in the early phase of phonics-based reading instruction has not yet been studied. Therefore, in the present longitudinal study, we focused on the incremental early word decoding development in Dutch children at risk for dyslexia during phonics-based reading instruction in first grade.

There are several cognitive and linguistic characteristics, typically measured prior to formal reading instruction, that have been established as precursors of actual reading (e.g. Al Otaiba et al., 2011; Kirby et al., 2008; Landerl et al., 2013; McCardle, Scarborough, & Catts, 2001; Melby-Lervåg et al., 2012; Moll et al., 2014; Nelson, Benner, & Gonzalez, 2003; Puolakanaho et al., 2007): visual and verbal short-term memory, lexical retrieval, phonemic awareness, grapheme-phoneme knowledge and vocabulary skills. These characteristics are already developing prior to formal reading instruction. During the first year of systematic phonics-based reading instruction, children learn explicitly that written words consist of graphemes and that graphemes systematically correspond to phonemes. By blending these phonemes, children learn to accurately recode written words into their auditory counterparts (Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001). Via this sublexical route, children learn to phonologically recode simple words. After they have learned how to phonologically recode words, they incrementally proceed with learning to read orthographically complex words (Ehri, 2005, 2014; Verhoeven & Perfetti, 2011). As a consequence of ongoing reading experience, orthographic representations in the mental lexicon become better specified and redundant, which makes word reading more and more efficient (Perfetti, 1992; Perfetti, 2007).

Not all children learn to read effortlessly; there is large individual variation in reading performances among children. Children who experience difficulties in learning to read are likely to continue to experience reading problems throughout the years (e.g. Verhoeven & Van Leeuwe, 2008) and to develop dyslexia (Snowling & Melby-Lervåg, 2016). Several studies evidenced an elevated chance for dyslexia of 30–50% for children with parents or siblings with dyslexia, as compared with a prevalence of 5–10% in the general population (Boets et al., 2010; Olson, Keenan, Byrne, & Samuelsson, 2014). Children at genetic risk not only show poorer reading performances than children without known genetic risk; the ‘gap’ between the performances becomes larger over the school years (e.g. Van Bergen et al., 2011; but see Pfost, Hattie, Dörfler, & Artelt, 2013). It is assumed that this elevated prevalence is the result of genetic rather than environmental factors (e.g. Swagerman et al., 2015; Van Bergen et al., 2011); see Pennington and Olson (2005) for a review on the genetics of dyslexia). Therefore, manifestations of dyslexia in parents or older siblings are typically used to uncover children at genetic risk for later reading problems (see for review Snowling & Melby-Lervåg, 2016).

Children at risk already show poorer performances on precursor measures like vocabulary, lexical retrieval, phonological awareness and short-term memory

(e.g. Carroll, Mundy, & Cunningham, 2014; Dandache, Wouters, & Ghesquière, 2014; Moll, Loff, & Snowling, 2013; Snowling & Melby-Lervåg, 2016; Van Bergen, De Jong, Maassen, & Van der Leij, 2014). Although these precursors have been established as good predictors of later reading development (e.g. Kirby et al., 2008; Landerl et al., 2013; Melby-Lervåg et al., 2012), their relative contribution might differ throughout the phases of reading development (Al Otaiba & Fuchs, 2002; Moll et al., 2014; Nelson et al., 2003). Ziegler, Perry, and Zorzi (2014) found that phonemic awareness deficits, as found in children with dyslexia, may have a negative effect on the development of stable orthographic representations during formal reading instruction. In line with this, Tilanus, Segers, and Verhoeven (Tilanus, Segers, & Verhoeven, 2013) found that children with dyslexia were less able to profit from repeated word exposure. Therefore, they might be less efficient in both the initial sublexical processes of phonological recoding and the later lexically driven processes of word decoding.

So far, longitudinal reading studies in children at genetic risk have mostly focused on precursors and on their relation to the reading development of children by the end of first grade and beyond. By that time, however, children have already mastered the basic principles of word decoding. Only very few studies have focused on the early development during Grade 1 (e.g. Snowling, Muter, & Carroll, 2007; Van Bergen et al., 2011). However, the large intervals being used in these few studies do not capture the children's earliest developmental trajectory as also explained by Fuchs, Fuchs, and Compton (2004) and by Oslund et al. (2015). The actual emergence of the reading problems of children at risk has thus not been extensively documented before, and the relation between the precursor measures and the early word decoding development in children at risk still remains unclear.

It is clear that children at risk are disadvantaged in their reading development over the years. However, it is yet unknown how genetically based variation in word decoding development unfolds during the initial phases of learning to read in first grade, and how kindergarten precursors contribute to the early development of word decoding in children at risk as compared to controls. Therefore, it remains an open question how the information that is available before formal literacy development initiates (viz. family risk information and cognitive profile in kindergarten) might be used by educators for early identification and intervention of children with later reading problems. The present study aimed to document the development of early word decoding as a function of kindergarten precursors in first grade in 73 Dutch children at risk for dyslexia and 73 controls. It is important to note that Dutch has a fairly transparent orthography (e.g. Seymour, Aro, & Erskine, 2003). During the first 5 months of formal reading instruction in Grade 1, children learn the grapheme–phoneme correspondences in five incremental steps along with word decoding exercises of simple words with a consonant–vowel–consonant (CVC) structure. In the second half of the first year, words of greater orthographic complexity are introduced in explicit instruction, such as words with consonant clusters and polysyllabic words.

In the present study, we assessed vocabulary, grapheme-phoneme knowledge, phonemic awareness, lexical retrieval and verbal and visual short-term memory in kindergarten. Incremental steps in word decoding efficiency were assessed by monthly curriculum-embedded word decoding measures during the first 5 months of formal reading instruction in first grade, and standardized word decoding

measures halfway and by the end of first grade. In the present study, we aimed to answer the following research questions:

- 1 What are differences between children at genetic risk and controls in precursor measures and in the early phases of word decoding development?
- 2 How are precursor measures related to the early development of word decoding in children at genetic risk as compared to controls?

We expected that children at risk would score lower than the control group on the precursor measures, in particular, measures of phonemic awareness, lexical retrieval and grapheme-phoneme knowledge. We further expected that differences between children at risk and controls in early word decoding development would already be established during the first 5 months of incremental word decoding development and that the differences would subsequently increase. In addition, we expected that the early development of word decoding skills would be stable over time for both children at risk and controls and that precursor measures would be related to word decoding development to a similar extent in children at risk.

## METHOD

### Participants

The current subsample was composed out of a larger cohort of a longitudinal reading study among 1006 Dutch children in 37 primary schools (Schaars, Segers, & Verhoeven, 2017). For the current study, we included children at genetic risk for reading problems and a matched control group of children without known genetic risk (for readability purposes, in this article called children at risk vs controls). Children were considered at risk if at least one of two biological parents or a sibling was diagnosed with dyslexia. This information was provided by the school teams. Schools in the Netherlands ask this information to parents upon school entry of their child. We had a response rate of 95.2% on the questionnaire about family risk status in the larger cohort. From this group, 7.6% of the children turned out to have a parent or sibling with an official diagnoses of dyslexia. This resulted in a subsample of 73 children with a genetic risk indication distributed across 26 schools, located in both rural and urban areas. This group was matched to an age-matched and gender-matched control group of 73 children without known genetic risk. To control for the interference of environmental influences, also school and classroom as well as home language and educational level of their main caregivers were taken into account in matching the children.

At the start of the study, the mean age of the children in the at risk group was 6;2 years ( $SD = 0;4$ , 47 boys and 26 girls). The control group consisted of 48 boys and 25 girls with a mean age of 6;1 ( $SD = 0;4$ ). Four children (two matched pairs) spoke Arabic or Kurdish at home; the other children in the subsample had Dutch as their home language. School, class, educational level and home language were considered on higher priority in matching the children than gender was. On this basis, gender was not properly matched in one child pair. There were no differences between the two groups on nonverbal logical reasoning, according

to raw scores of the Raven Coloured Matrices assessment (Raven, 1958) by the end of Grade 1 ( $M_{\text{atrisk}} = 27.28$ ;  $SD = 4.05$ ;  $M_{\text{controls}} = 28.06$ ;  $SD = 4.12$ ;  $t(141) = 1.14$ ,  $p = .25$ ). The Raven scores were similar to the scores in our larger cohort ( $M_{\text{generalcohort}} = 27.65$ ). The mean educational level of their main caregivers was similar in both groups and comparable with the Dutch population (Centraal Bureau voor de Statistiek [Statistics Netherlands], 2013) and with our larger cohort ( $M_{\text{subsample}} = 2.84$ ;  $M_{\text{generalcohort}} = 2.68$  on a categorical scale ranging from 0 (low educated) to 4 (high educated)).

## MEASURES

### Curriculum-embedded Word Decoding

The monthly measurements that were assessed during the first 5 months of first grade, were embedded in the systematic phonics-based reading curriculum that was used by all participating schools. After each curriculum-based training block of 3–4 weeks, word decoding mastery level up until then was measured by embedded assessment cards *Veilig en Vlot* ('Safe and Smooth'; Mommers et al., 2003). Because the assessment cards were specifically designed to measure the short-term attainment of recently taught skills, the consecutive measurements were not equivalent but instead developed along with the content of the curriculum (see also Oslund et al., 2015) for further explanation of curriculum-embedded measurement. The words on the assessment cards were composed of the graphemes and structures that the children had trained on during the previous blocks. The used words in the first six blocks were all CVC (and CV and VC)-structured one-syllable words with consistent mappings between graphemes and phonemes. The children were asked to read out aloud the words as accurately and quickly as possible for 1 min. The number of accurately read words in 1 min was the word decoding efficiency score.

### Standardized Word Decoding

The tasks to measure absolute word decoding skills at two-time points during Grade 1 (halfway and by the end of the first grade) were standardized, and scores were comparable over time. The standardized word decoding task comprised two reading cards of the *Drie-minutentoets* ('Three minute test'; Krom, Jongen, Verhelst, Kamphuis, & Kleintjes, 2010). The first card contained simple and transparent CVC (and CV and VC)-structured one-syllable words. The one-syllable words on the second card contained at least one consonant cluster (-CC and CC-) and were therefore considered more complex and advanced. In addition, more complex orthographic structures were used in some words on the second card. For example, the rule of devoicing -d in the end position of a word (viz. *dak* is pronounced as [d-a-k], but *pad* is pronounced as [p-a-t]). For both reading cards, the child was asked to read out aloud the words as accurate and quickly as possible for 1 min. For each card, the number of accurately read words in 1 min was the word decoding efficiency score.

A combination score of the two cards was reported reliable, with a Cronbach's  $\alpha$  of .96 (Krom et al., 2010) and can be reliably used as an indicator of standardized word decoding efficiency in Grade 1.

### **Precursor Measures**

The precursor measures comprised seven tasks on cognitive and linguistic abilities that were administered by the end of kindergarten. The tasks were specifically designed for the purpose of the longitudinal study (Schaars et al., 2017). To obtain insight into the psychometrical quality of the kindergarten test, the Cronbach's alpha of each task was normed on the representative larger cohort of 1006 children. In addition, the variance from the mean and the deviation of the scores between students provided good distributions for sensible analyses of the individual variations among the children. No floor or ceiling levels for the tasks nor for the individual items were reached. Items within the tasks were considered coherent and item difficulty reached acceptable levels.

#### *Phoneme isolation*

To measure phonemic awareness, the child was asked to sound out the first phoneme of 10 orally presented monosyllabic CVC-structured words (e.g. *mu**i**s*, *s**o**e**p*; [*m*o*u*s*e*, *s*o*u*p]). The score on this task was the amount of correct responses, with a maximum score of 10. Reliability of the task was good (Cronbach's  $\alpha = .83$ ).

#### *Phoneme segmentation*

To measure phoneme segmentation skills, the child was asked to serially pronounce each single speech sound of an orally presented monosyllabic and highly consistent word. There were 10 words in this task with increasing difficulty, starting with CVC-structured words followed by CCVC- or CVCC-structured words and CCCVC-structured or CVCCC-structured words. The score on this task was the amount of correct responses, with a maximum score of 10. The reliability of the task was good (Cronbach's  $\alpha = .85$ ).

#### *Grapheme-phoneme knowledge*

To assess emergent literacy skills, we asked the child to sound out 34 graphemes used in Dutch. In this task, only the grapheme sound was considered correct; naming the grapheme's name was incorrect. It is common to measure grapheme-phoneme knowledge in the Netherlands, because in the Netherlands, children usually start with learning the speech sounds of graphemes before they learn the letter names (Melby-Lervåg et al., 2012). If a child named the grapheme's name, a second chance was offered once. The test assistant replied with 'Yes, that is the name of the letter. But do you know the sound of the letter?'. If the child could correctly reply to that second question, the item was counted as correct. Reliability of the task was excellent (Cronbach's  $\alpha = .93$ ).

#### *Lexical retrieval*

To measure the speed of lexical retrieval of visually presented objects, the child was asked to name pictures from a card with repeated rendering of five pictures. The five pictures corresponded with one-syllable high frequent Dutch words (viz. *saw*, *pot*, *thumb*, *trousers* and *tent*). The child was asked to name the pictures



as accurate and quickly as possible during 1 min. The amount of correct named pictures was the lexical retrieval score. Reliability of the task was excellent (Cronbach's  $\alpha = .95$ ).

#### *Verbal short-term memory*

To assess verbal short-term memory, we asked the children to repeat 20 orally presented pseudo-words with increasing complexity and number of syllables (increasing from 1 to 4 syllables). All phonemes and entire syllables in a word had to be repeated correctly. Repetitions containing stress differences or substitutions due to certain articulation errors in individuals were counted as correct. The reliability of the task was good (Cronbach's  $\alpha = .77$ ).

#### *Visual short-term memory*

To assess visual short-term memory, we asked the children to remember and rebuild the order of a series of visual figures (viz., *fish, cow, ship, chicken* and *sock*) that was presented shortly by the test assistant. The amount of visual figures in a series increased from two to five figures to remember. The entire series had to be remembered to be considered correct. In total 15 series were offered. This task followed the task design of the 'Visual Memory Span: Concrete Figure Sequences' sub-task of the Revisie Amsterdamse Kinder Intelligentie Test RAKIT-test kit (Pieters, Dek, & Kooij, 2013). The reliability of the visual short-term memory task was good (Cronbach's  $\alpha = .77$ ).

#### *Active vocabulary*

The Active vocabulary task was based on the Vocabulary task in the *Taaltoets Allochtone Kinderen* ('Language test Ethnic Minority Children' Verhoeven & Vermeer, 1986). A total of 29 words was assessed. The reliability was good (Cronbach's  $\alpha = .83$ ).

All precursor measures, except for the grapheme-phoneme knowledge and the lexical retrieval task, contained a cut-off score to avoid further frustration if the performance level of a child was reached.

### **Procedure**

We assessed baseline precursors by the end of kindergarten, so before formal reading instruction started in Grade 1. The precursor measures were administered by the first author and eight trained test assistants with Bachelor or Master Degrees in Educational Science, Psychology or Linguistics. The test assistants were extensively trained on judging the tasks in practice sessions prior to any data collection, and until agreement was reached among the first author and the eight test assistants. The test assistants were not aware of any child background information including risk status at the moment of the assessment. All tasks were administered individually in a quiet room at school during regular school hours.

In Grade 1, the word decoding instruction was provided in general classroom setting following the normal reading curriculum. All participating schools made use of the same systematic incremental phonics-based reading curriculum, called *Veilig Leren Lezen* ('Learning to Read Safely'; Mommers et al., 2003). This curriculum comprises 10 successive instruction blocks of 3–4 weeks. The first six instruction

blocks are characterized by a systematic incremental offering of grapheme–phoneme correspondences during which children incrementally learn to read simple CVC-structured one-syllable words. After the first six instruction blocks, 34 graphemes are covered by the reading instruction in classroom (i.e., all graphemes used in Dutch except from c, q, y, x). Thereafter, four more instruction blocks follow. During these following instruction blocks, explicit instruction focuses on complex internal word structures. Consequently, children incrementally learn to read words with consonant clusters (CC- and -CC), polysyllabic words and words with morphological units. They also learn specific rules, for example, that the pronunciation of a grapheme in a word is sometimes defined by the cluster to which it belongs or by its position within a word (e.g. in Dutch *sch-* and *-eeuw*). The instruction method consists of extensive manuals and schedules, and the lessons and materials are well defined to ensure consistency in education between schools.

To capture each incremental step that the children made during the first 5 months of reading instruction, the children's ability to decode words was assessed after each training block of 3–4 weeks. Assessments were conducted using curriculum-embedded word decoding tasks of 1 min, assessed by certified and well instructed classroom teachers of the participating schools (mostly the daily teachers of the children). This is part of their normal education pursuits, and the curriculum-embedded tasks were designed for assessment by classroom teachers.

Standardized word decoding tasks were administered after explicit instruction of phonological recoding skills (halfway Grade 1) and after explicit instruction of complex word decoding (at the end of Grade 1). Because we intended to measure the decoding level of both simple structured words and more complex structured words, both reading cards (respectively, simple one-syllable words and complex one-syllable words) were assessed individually at both time points.

### Analytic Approach

Data were evaluated on missing data patterns and distributional characteristics before analyses. To give answer to the first research question, *t*-tests for independent measures and two Repeated Measures Analyses of Variance (ANOVAs) were conducted with Time as the within subject factor and Group as the between subject factor. Interactions were further explored using one-way ANOVAs and independent samples *t*-tests (two-sided). Non-parametric approaches were used to check the outcomes of the analyses with curriculum-embedded incremental word decoding (Jöreskog & Sörbom, 1996). To answer the second research question, Linear Structural Relations (LISREL) group comparison models were used. The fit of the group model was evaluated using chi-squared statistics ( $\chi^2$ ). The root-mean-square error of approximation (RMSEA) and the relative chi-squared ( $\chi^2_{rel}$ ), calculated as the ratio of the chi-squared with the degrees of freedom, were also evaluated. As a guideline for accepting the model, the RMSEA cut-off criterion was set at  $<.06$  (Hu & Bentler, 1999; Tabachnick & Fidell, 2014), and the relative chi-squared should be lower than 3 (Carmines & McIver, 1981).



## RESULTS

The data were missing for 2.26% of the values. Less than 1.5% of the values were missing in the curriculum-embedded word decoding measurements, while 9.9% of the values were missing in the standardized word decoding measurements. No data were missing in the kindergarten cognitive and linguistic measurements. The missing values were mostly due to incidental absence of individual children, due to illness during a measurement moment in Grade 1. For six individuals (three matched pairs), no data were collected on the standardized test halfway Grade 1, because the participating school was not able to collect the measurements in time. In the current sample, no children were excluded from analyses owing to longer lasting illness or movement to other schools.

To prevent from unnecessary relevant data loss, missing values were estimated by the expectation maximalization method of SPSS (EM; IBM SPSS 23). The dataset was suitable for this method, because the missing pattern was considered at random (Little's missing completely at random test for the curriculum-embedded measurements:  $\chi^2(20) = 19.87, p = .47$  and Little's Missing Completely At Random (MCAR) test for the standardized measurements:  $\chi^2(11) = 15.97, p = .14$ ).

Prior to analysis, the variables were examined for relevant assumptions separately for the 73 children at risk and the 73 controls. The assumption of sphericity was violated (Mauchly's test  $\chi^2(14) = 380.53, p < .001$ ) in both groups for the curriculum-embedded word decoding tasks. Therefore, Greenhouse Geiser was reported ( $\epsilon = .43$ ) in the Repeated Measures analysis. The residuals of the data for the curriculum-embedded word decoding variables were not normally distributed for both groups. Overall, *F*-tests are robust to deviations from normality (see Lindman, 1974), so we first conducted classic ANOVA. We checked the main effects of Time and of Group of the curriculum-embedded variables with non-parametric Mann–Whitney tests and a Wilcoxon signed-rank test. Non-parametric methods are not particularly suitable for testing interaction effects in Repeated Measures designs, however. To approach the interaction analysis that was conducted in the classic (parametric) Repeated Measures ANOVA, we analysed the Group \* Time interaction effect by analysing the total difference score from WDC1 to WDC6 in one step in the non-parametric Mann–Whitney analysis.

### Differences in Kindergarten Precursors

The first research question first concerned differences between children at risk and controls on the kindergarten precursor measures. Table 1 presents the means and standard deviations of phoneme isolation, phoneme segmentation, grapheme-phoneme knowledge, lexical retrieval, vocabulary and verbal and visual short-term memory per group. Independent samples *t*-tests (two-sided) showed that the groups differed significantly on Phoneme isolation (equal variances not assumed),  $t(115.23) = 2.10, p = .04, d = .3$  and on Phoneme segmentation (equal variances assumed)  $t(144) = 2.11, p = .04, d = .35$ , with medium effect sizes. There were no significant differences on the other precursors measures between the groups, as can be seen in Table 1.

Table 1. Means, standard deviations, and independent samples t-tests of the kindergarten precursors

Measurement	M (SD)	Min–max	$M_{\text{difference}}$	$df_{\text{(equal variances assumed/not assumed)}}$	t	p	Cohen's d
Phoneme isolation							
Control	8.78 (1.37)	3–10	.67	115.23	2.10	.04	.39
At risk	8.11 (2.37)	0–10					
Phoneme segmentation							
Control	4.81 (2.47)	0–10	.84	144	2.11	.04	.35
At risk	3.97 (2.32)	0–9					
Grapheme-phoneme knowledge							
Control	18.92 (6.64)	5–32	2.05	144	1.72	.09	.29
At risk	16.86 (7.75)	2–31					
Lexical retrieval							
Control	41.26 (8.51)	18–66	1.92	144	1.42	.16	.24
At risk	39.34 (7.81)	22–61					
STM verbal							
Control	15.79 (2.33)	10–20	.55	138.43	1.27	.21	.22
At risk	15.25 (2.85)	8–20					
STM visual							
Control	8.12 (3.02)	0–13	.03	144	−.05	.96	.01
At risk	8.15 (2.99)	1–14					
Active vocabulary							
Control	14.70 (4.19)	2–23	.04	144	−.07	.95	.01
At risk	14.74 (3.48)	6–23					

STM, short-term memory.

### Differences in Incremental Word Decoding Development

Table 2 shows the mean word decoding efficiency (correct per minute) and the standard deviations for each repeated measurement of the curriculum-embedded word decoding tasks and for the standardized word decoding tasks. The mean scores on the curriculum-embedded word decoding efficiency per group are presented graphically in Figure 1.

To answer the research question concerning the differences in incremental word decoding development, the curriculum-embedded word decoding measurements were analysed in a Repeated Measures design. Repeated Measures ANOVA with Time (WDc1–WDc6) as within-subjects factor and Group (at risk, controls) as between subject factor revealed a Group \* Time interaction,  $F(2.12, 305.66) = 3.14, p = .042, \eta_p^2 = .02$ . This indicated that there was a (small) effect of genetic risk factor on incremental word decoding development with simple CVC-structured words during the first months of formal education, which can also be seen in the slightly divergent developing lines in Figure 1. A main effect of Time,  $F(2.12, 305.66) = 82.63, p < .001, \eta_p^2 = .37$  and a main effect of Group was found,  $F(1, 144) = 5.24, p = .024, \eta_p^2 = .04$ . Post hoc independent comparisons on all measurement moments showed group differences, except for WDc1 and WDc3, with the at-risk group scoring below the control group (as can be seen in Table 2).

Non-parametric tests were conducted to check the results in the classic analyses procedures as described earlier. First, a one-step Mann–Whitney analysis

Table 2. Mean word decoding efficiency (correct per minute), standard deviations and independent samples t-tests for each Repeated Measurement of the curriculum-embedded tasks and the standardized tasks for word decoding

Measurement	M (SD)	min-max	M <sub>difference</sub>	df <sub>(equal variances assumed)</sub>	t	p	Cohen's d
WDc1							
Control	18.50 (9.89)	4-64.44	2.91	144	1.61	.11	.27
At risk	15.580 (11.83)	1-85.71					
WDc2							
Control	21.59 (10.04)	6-70.59	3.81	144	2.11	.04	.35
At risk	17.78 (11.69)	1-78.62					
WDc3							
Control	22.49 (13.08)	5-82.76	3.89	144	1.68	.09	.28
At risk	18.59 (14.79)	1-114					
WDc4							
Control	26.12 (13.55)	10-80	4.82	144	2.08	.04	.36
At risk	21.30 (14.43)	6-96					
WDc5							
Control	27.87 (14.77)	8-73.13	6.12	144	2.39	.02	.40
At risk	21.76 (16.10)	4-92.31					
WDc6							
Control	34.89 (18.48)	7-81.43	8.10	144	2.57	.01	.43
At risk	26.79 (19.56)	8-108.57					
WDs Halfway CVC							
Control	34.46 (15.01)	9-71	7.65	144	3.13	.002	.52
At risk	26.81 (14.48)	11-87					
WDs Halfway CCVC							
Control	15.70 (8.130)	3-42	3.04	144	2.05	.04	.34
At risk	12.66 (9.71)	1-57					
WDs End CVC							
Control	49.11 (17.81)	14-87	9.54	144	3.00	.003	.50
At risk	39.56 (20.55)	11-109					
WDs End CCVC							
Control	35.16 (17.03)	6-78	9.32	144	3.27	.001	.54
At risk	25.84 (17.43)	3-91					

WDc, curriculum-embedded word decoding task; WDs, standardized word decoding task; CVC and CCVC resp. indicate CVC/CVVC-structured task or task with words containing at least one consonant cluster.

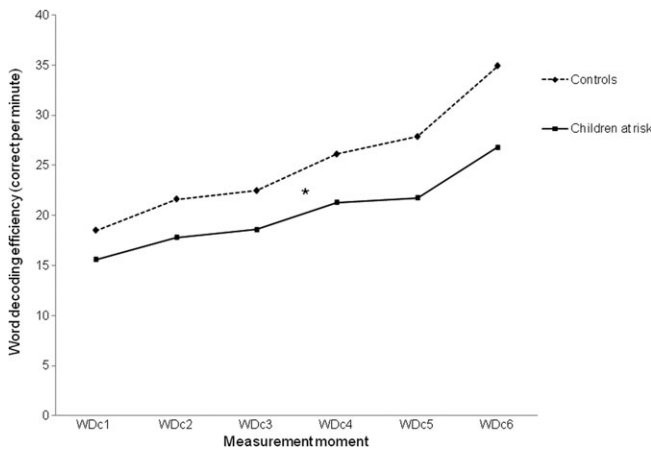


Figure 1. Mean scores on curriculum-embedded word decoding (WDc) efficiency per group.

with difference scores from WDc1 to WDc6 was used to check the found interaction of Group and Time. In line with the previous finding, we found a significant difference between the groups in word decoding development over time ( $U = 1974.00$ ,  $Z = -2.70$ ,  $p = .007$ ). Next, a Wilcoxon signed-rank test showed that there was a significant difference between WDc1 and WDc6, indicating an overall effect of Time ( $Z = -9.72$ ,  $p < .001$ ). A Mann–Whitney analysis on the overall average of the consecutive measurement moments per group, revealed again a significant main effect of Group, with lower scores for the at-risk group ( $U = 1792.00$ ,  $Z = -3.42$ ,  $p = .001$ ). To check the results of the post hoc independent comparisons, we conducted a Mann–Whitney analysis on the six consecutive measurement moments (Table 3). This analysis revealed significant differences in word decoding efficiency between the groups for all measurement moments. This was somewhat different from the results in the classic post hoc independent analyses, because no significant difference was found in WDc1 and WDc3 in that approach. Overall, non-parametric analyses showed an interaction effect and a main effect of Group and Time, confirming our findings in the classic analyses.

### Differences in Standardized Word Decoding Development

We conducted an overall Repeated Measures ANOVA with Complexity (CVC and CCVC) and Time (halfway Grade 1, end Grade 1) as within-subjects factors and Group (at risk, controls) as between subject factor to analyse the standardized word decoding efficiency of simple and advanced word decoding halfway Grade 1 and end Grade 1. An overall Group \* Time \* Complexity interaction,  $F(1, 144) = 7.91$ ,  $p = .006$ ,  $\eta_p^2 = .05$ , indicated that the groups differed in their overall development over time, that the groups differed on their overall scores on reading the two word complexities and that the groups overall increased in word decoding efficiency over time. The analyses further revealed a Group \* Time interaction,  $F(1, 144) = 6.05$ ,  $p = .015$ ,  $\eta_p^2 = .04$ , as well as a Group \* Complexity interaction,  $F(1, 144) = 6.13$ ,  $p = .014$ ,  $\eta_p^2 = .04$ , and a Time \* Complexity interaction  $F(1, 144) = 11.32$ ,  $p = .001$ ,  $\eta_p^2 = .073$ . Main effects of complexity,  $F(1, 144) = 964.118$ ,  $p < .001$ ,  $\eta_p^2 = .87$ , Time,  $F(1, 144) = 325.99$ ,  $p < .001$ ,  $\eta_p^2 = .69$  and of Group  $F(1, 144) = 9.88$ ,  $p = .002$ ,  $\eta_p^2 = .06$  were found.

We further discuss the results of both Word decoding complexities separately to find out what the exact differences between the groups were on both complexities. First, a Repeated Measures ANOVA for the simple structured CVC-words showed a main effect of Time,  $F(1, 144) = 263.10$ ,  $p < .001$ ,  $\eta_p^2 = .65$  and a main effect of Group  $F(1, 144) = 10.07$ ,  $p = .002$ ,  $\eta_p^2 = .07$ . No Group \* Time interaction was found,  $F(1, 144) = 1.26$ ,  $p = .26$ ,  $\eta_p^2 = .01$ . This means that the

Table 3. Non-parametric (Mann–Whitney U) test for post hoc independent comparisons of the incremental word decoding development

	WDc1	WDc2	WDc3	WDc4	WDc5	WDc6
Mann–Whitney $U$	1927.500	1860.500	1960.000	1895.000	1753.500	1802.000
$Z$	-2.889	-3.151	-2.760	-3.014	-3.568	-3.377
$p$	.004	.002	.006	.003	<.001	.001

standardized efficiency scores of both groups increased similarly over time but that the children at risk showed lower overall performance levels than controls.

Next, for the word decoding task containing consonant clusters (CCVC), there was a Group \* Time interaction,  $F(1, 144) = 10.14, p = .002, \eta_p^2 = .07$ , as well as a main effect of Time,  $F(1, 144) = 273.9, p < .001, \eta_p^2 = .66$  and a main effect of Group  $F(1, 144) = 9.11, p = .003, \eta_p^2 = .06$ . Post hoc independent comparisons showed lower performance levels for children at risk on both measurement moments, respectively (equal variances assumed)  $t(144) = 2.05, p = .042, d = .34$ ;  $t(144) = 3.27, p = .001, d = .54$ . The interaction can be explained by children at risk showing less progress than controls (independent groups comparison of difference scores (equal variances assumed),  $t(144) = 3.19, p = .002, d = .53$ ).

Mean scores per group on word decoding efficiency with both word complexities are presented graphically in Figure 2.

### Precursors of Word Decoding Development

To answer the second research question, we used multiple-group comparison models in LISREL (Jöreskog & Sörbom, 1996). These models are powerful ways to testing hypotheses about potential differences between groups (Little, 2013). We tested the path model from relevant cognitive skills (Phonemic awareness and Lexical retrieval) to emergent literacy (grapheme-phoneme knowledge) to word decoding at the start, halfway and in the end of Grade 1 (Figure 3). The word decoding measure at the Start of Grade 1 is constructed of only the first curriculum-embedded word decoding measurement (WDc1). Word decoding Halfway Grade 1 is constructed with Principal Axis Factoring (PAF) promax factor analysis, containing both the simple standardized word decoding task and the task containing consonant clusters. The word decoding measure in the End of Grade 1 is constructed with PAF promax factor analysis containing both the simple word decoding task and the word decoding task containing consonant clusters that

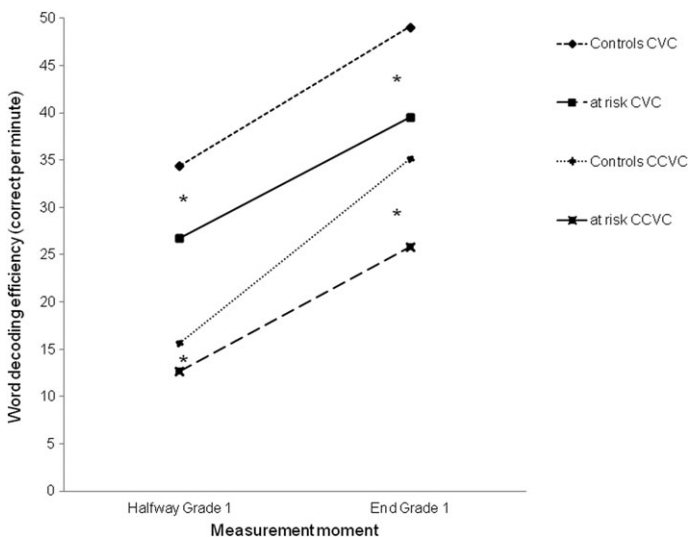


Figure 2. Mean scores on word decoding efficiency with different word complexities per group.

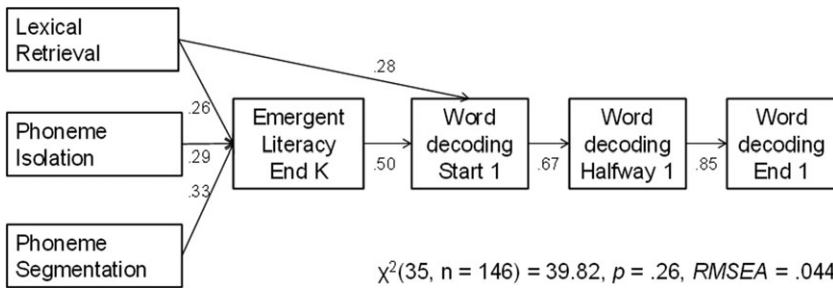


Figure 3. Structural equation multiple-group comparison model of early word decoding development. Numbers represent the standardized coefficients. End K, end of Kindergarten. Followed by start, halfway and end of Grade I.

had been conducted in the end of Grade I. Standardized coefficients of the model are presented in Figure 3. The model shows a strong relation of lexical retrieval, phoneme isolation and phoneme segmentation to the measure of emergent literacy by the end of kindergarten (represented by the grapheme-phoneme knowledge). A high predictive power of these kindergarten precursor measures is given through by the emergent literacy measure (.50) to the word decoding performances at the start of Grade I (WDcI), with an additional and independent direct contribution of lexical retrieval to word decoding at the start of Grade I. The coefficients from one word decoding measurement to the other (from the start to halfway and from halfway to the end of Grade I) are within-construct prediction coefficients, and therefore, they can be interpreted as stability coefficients of the longitudinally measured word decoding development (auto regression). The model shows high stability of word decoding in Grade I. The model turned out to fit the data on both groups very well ( $\chi^2(35, n = 146) = 39.82, p < .26, RMSEA = .044$ ). This multiple-group analysis showed that the contributions to the  $\chi^2$  were equal for both groups. The group of children without known risk contributed 48.23% to the  $\chi^2$  of the conceptual model. The group of children at risk contributed 51.77% to the  $\chi^2$ .

## DISCUSSION

The purpose of the current longitudinal study was to examine the early word decoding development of children with and without known genetic risk at micro level, from kindergarten to the end of Grade I. Children at risk were less skilled than controls in phonemic awareness in kindergarten. No differences were found for grapheme-phoneme knowledge, lexical retrieval, short-term memory and vocabulary. Furthermore, the results during the first months of learning to read showed that word decoding efficiency of the children at risk lagged behind that of the control group, with the at-risk group showing slower progression. After the explicit instruction of phonological recoding was finished, the discrepancy between the groups remained stable for simple word decoding. During the following months, decoding efficiency of more advanced words (containing consonant clusters and orthographic complexities) consistently differed between groups. Although both groups increased their advanced word decoding skills, the control group developed faster than the group at risk. Finally, it was found that



for both groups to the same extent, phonemic awareness and lexical retrieval predicted emergent literacy in kindergarten, which was highly predictive for the early word decoding development during phonics-based instruction in Grade 1.

With respect to our research question about differences in precursor measures, the results showed that children at risk had less skilled phonemic awareness than controls in kindergarten. This early delay in phonemic awareness has been evidenced previously (e.g. Snowling, Gallagher, & Frith, 2003; Snowling & Melby-Lervåg, 2016) and can be argued to show early delays in sublexical processing. This finding is in line with the phonological deficit hypothesis for children with dyslexia (Dehaene, 2009; but see Castles & Friedmann, 2014). No differences were found in the performances on lexical retrieval, suggesting that reading might not be hampered by early problems in lexical processing. Furthermore, children at risk did not differ on grapheme-phoneme knowledge, vocabulary and short-term memory skills. This was somewhat surprising compared with findings in general genetic risk literature (see the meta-analysis and review of Snowling & Melby-Lervåg, 2016). However, there are other risk studies describing no differences between at risk and control groups in, for example, grapheme-phoneme knowledge (Blomert & Willems, 2010; Carroll & Snowling, 2004). This finding may be partly explained by a shifting trend in education. Education in kindergarten tends to give more attention to the development of linguistic characteristics in recent years. It might be possible that children who did not playfully catch up skills like grapheme-phoneme knowledge and vocabulary in kindergarten received extra explicit help and attention from kindergarten teachers or parents. In line with that, questionnaires in the study of Gallagher, Frith, and Snowling (Gallagher, Frith, & Snowling, 2000) even showed that parents of children at risk are more aware of the risk status and tend to spend more time on practicing letters as compared with parents of controls. Such additional practicing in kindergarten may have caused decreasing individual variation in the precursor measures.

Results on the early word decoding development showed that children at risk already lagged behind in word decoding efficiency during incremental phonological recoding instruction, as compared with the controls. This finding confirms the discrepancy that has been found in genetic risk studies that focused on later phases of reading development. The current study shows that the previously evidenced discrepancy already emerges from the very beginning of learning to read. Although emerging, group differences seem small during the first 3 months, indicating that children at risk might keep pace with typical readers as long as reading is not too demanding yet. However, overall differences in progression rate were found during the first months. After the explicit instruction of phonological recoding was finished, we found that the discrepancy between groups stabilized for simple word decoding measures. Both groups showed ongoing efficiency gains over time, indicating that both groups kept developing their simple reading skills with similar progression.

Results on the advanced word decoding development showed that both groups increased their advanced word decoding over time. The decoding efficiency scores differed consistently between groups, with controls developing faster than children at risk. The finding that children at risk systematically lag in word decoding efficiency is in line with our expectations and in accordance with results reported by, for example, Dandache et al. (2014) and Pennington and Lefly (2001). Both

groups were more efficient in simple as compared with advanced word decoding tasks. This finding elaborates on the findings in children with dyslexia (e.g. Tilanus et al., 2013). The results show a notable difference between the simple word decoding development and the more advanced word decoding development. These different developmental paths clearly indicate that at-risk children's development differs slightly from the control children from early on (during simple word decoding instruction), while the difference is clearly demonstrated and growing larger once more complex words need to be processed. This sensitivity to increasing task demands has, although in different conditions, previously been described by Van der Leij and Van Daal (1999).

Regarding our second research question, we showed that the separated phases as described earlier were all highly related and could be embedded in one path model. Kindergarten phonemic awareness and lexical retrieval predicted the emergent literacy in kindergarten, which was highly predictive for the early word decoding development during phonics-based instruction in Grade 1. These findings are in line with the literature on predictors of reading development (e.g. Moll et al., 2014; Verhoeven, Van Leeuwe, Irausquin, & Segers, 2016). Early word decoding during the first weeks of formal phonics-based reading instruction was found to be highly predictive for later development of word decoding efficiency during Grade 1. The path model from precursors to word decoding is similar for both groups, showing that although word decoding development in both the initial and more advanced phases of word decoding differed between the groups, the groups do not differ in the predictability of one phase by another. Furthermore, it indicates that the predictive power of precursors of word decoding development is similar to both children at risk and typical developing readers. These findings are in line with our hypotheses.

The study shows that genetic risk indication should be used to identify children who should be carefully monitored during the initial reading process, so that right from the beginning, each child receives the optimal (additional) instruction. Of course, it should be acknowledged that although genetic risk factors might be useful in early identification of children at risk for reading problems, risk status cannot be used exclusively to diagnose children with dyslexia. Although children at risk have been found to be poorer readers as compared with controls (Van Bergen et al., 2011), not all children at risk do become dyslexic readers (Bishop, 2015). Furthermore, in the current study, we selected families at risk based on a previously issued official dyslexia diagnoses. The diagnosis system in the Netherlands is highly protocolled. Nevertheless, we recommend selection based on and controlled for actual word decoding measurements in parents and siblings in future research, to reduce noise in the sampling. Given these limitations, the present study makes it clear that children with genetic risk show to be less efficient in phonemic awareness and word decoding efficiency from the very beginning of phonics instruction and that the differences in word decoding tend to increase in the course of the grade. Accordingly, the current study provides insight into the genetically based variation in early word decoding development, and in how genetic disadvantages unfold during the initial phases of learning to read in first grade. The current study provides prospective insights in the early identification of children at risk for later reading problems. It would be interesting to follow

up on this prospective study by providing retrospective data on the early reading development of children who actually will be diagnosed with dyslexia.

The finding that from the very beginning of learning to read children at risk perform poorer than controls despite systematic and high-quality reading education, suggests that it could be difficult to prevent them from experiencing disadvantages in reading (Van der Leij et al., 2013). In addition, it might not be easy to remediate reading difficulties by didactical and remedial efforts (e.g. Dandache et al., 2014). The findings underscore the persistence of dyslexia and provide further support for the suggestion that the nature of dyslexia lies in cognitive and biological factors. To put everything in the right perspective though, it should be mentioned that both groups continue to develop their reading skills over time. This progression shows their learning potential and emphasizes the urge for good education and guidance of children at risk.

The current study has several educational implications. First, results show that children at risk for later reading problems should be screened early. Second, curriculum-embedded measurement is a sensitive and efficient method to identify individual differences in beginning first-grade reading development. Furthermore, explicit instruction and extra attention for children at risk needs to be provided from the very beginning. Finally, it is important to continue explicit instruction and practicing beyond the first months of explicit instruction, because children at risk are developing significantly slower as soon as word decoding becomes more demanding. In other words, practitioners should be alert on the fact that children at risk are more sensitive to increasing reading complexity as compared with controls.

The current study disentangled the early phonics-based word decoding development in children at risk for dyslexia. The development from precursors to emergent literacy to early word decoding was studied during the year in which children learn to read. Children at risk showed less efficient phonemic awareness skills in kindergarten, and from the very beginning of reading instruction, the control group performed increasingly better than the children at risk on both simple and more advanced word decoding. The development of word decoding abilities showed great stability over time with phonological awareness in kindergarten as the strongest predictor measure.

## REFERENCES

- Al Otaiba, S., & Fuchs, D. (2002). Characteristics of children who are unresponsive to early literacy intervention: A review of the literature. *Remedial and Special Education, 23*(5), 300–316 <https://doi.org/10.1177/07419325020230050501>.
- Al Otaiba, S., Folsom, J. S., Schatschneider, C., Wanzek, J., Greulich, L., Meadows, J., ... Connor, C. M. (2011). Predicting first-grade reading performance from kindergarten response to tier I instruction. *Exceptional Children, 77*(4), 453–470 Retrieved from: <http://ecx.sagepub.com/content/77/4/453.full.pdf>.
- Bishop, D. V. M. (2015). The interface between genetics and psychology: Lessons from developmental dyslexia. *Proceedings Royal Society Publishing, 282*, 20143139. <https://doi.org/10.1098/rspb.2014.3139>.
- Blomert, L., & Willems, G. (2010). Is there a causal link from a phonological awareness deficit to reading failure in children at familial risk for dyslexia? *Dyslexia, 16*(4), 300–317 <https://doi.org/10.1002/dys.405>.

- Boets, B., Smedt, B., Cleuren, L., Vandewalle, E., Wouters, J., & Ghesquiere, P. (2010). Towards a further characterization of phonological and literacy problems in Dutch-speaking children with dyslexia. *British Journal of Developmental Psychology*, 28(1), 5–31 <https://doi.org/10.1348/026151010X485223>.
- Carmines, E. G., & McIver, J. P. (1981). Analyzing models with unobserved variables: Analysis of covariance structures. In G. W. Bohrnstedt, & E. F. Borgatta (Eds.), *Social Measurement* (pp. 65–116). Beverly Hills, CA: Sage.
- Carroll, J. M., & Snowling, M. J. (2004). Language and phonological skills in children at high risk of reading difficulties. *Journal of Child Psychology and Psychiatry*, 45(3), 631–640 <https://doi.org/10.1111/j.1469-7610.2004.00252.x>.
- Carroll, J. M., Mundy, I. R., & Cunningham, A. J. (2014). The roles of family history of dyslexia, language, speech production and phonological processing in predicting literacy progress. *Developmental Science*, 17(5), 727–742 <https://doi.org/10.1111/desc.12153>.
- Castles, A., & Friedmann, N. (2014). Developmental dyslexia and the phonological deficit hypothesis. *Mind & Language*, 29(3), 270–285 <https://doi.org/10.1111/mila.12050>.
- Centraal Bureau voor de Statistiek [Statistics Netherlands] (2013), December 12. Retrieved from: <http://statline.cbs.nl/Statweb/publication/?DM=SLNL&PA=71822NED&D1=0-1&D2=0&D3=0&D4=04&D5=a&D6=0&D7=1&HDR=T,G2,G1,G6,G5,G4&STB=G3&VW=T>
- 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), 204–256 <https://doi.org/10.1037//0033-295x.108.1.204>.
- Dandache, S., Wouters, J., & Ghesquiere, P. (2014). Development of reading and phonological skills of children at family risk for dyslexia: A longitudinal analysis from kindergarten to sixth grade. *Dyslexia*, 20(4), 305–329 <https://doi.org/10.1002/dys.1482>.
- Dehaene, S. (2009). *Reading in the brain: The new science of how we read*. New York, USA: Penguin.
- Ehri, L. C. (2005). Learning to read words: Theory, findings, and issues. *Scientific Studies of Reading*, 9(2), 167–188 <https://doi.org/10.1207/s1532799xssr0902>.
- Ehri, L. C. (2014). Orthographic mapping in the acquisition of sight word reading, spelling memory, and vocabulary learning. *Scientific Studies of Reading*, 18(1), 5–21 <https://doi.org/10.1080/10888438.2013.819356>.
- Eklund, K., Torppa, M., Aro, M., Leppänen, P. H., & Lyytinen, H. (2015). Literacy skill development of children with familial risk for dyslexia through Grades 2, 3, and 8. *Journal of Educational Psychology*, 107(1), 126 <https://doi.org/10.1037/a0037121>.
- Fuchs, D., Fuchs, L. S., & Compton, D. L. (2004). Identifying reading disabilities by responsiveness-to-instruction: Specifying measures and criteria. *Learning Disability Quarterly*, 27(4), 216–227 <https://doi.org/10.2307/1593674>.
- Gallagher, A., Frith, U., & Snowling, M. J. (2000). Precursors of literacy delay among children at genetic risk of dyslexia. *Journal of Child Psychology and Psychiatry*, 41, 203–213 <https://doi.org/10.1111/1469-7610.00601>.
- Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1), 1–55 <https://doi.org/10.1080/10705519909540118>.
- Jöreskog, K. G., & Sörbom, D. (1996). *LISREL 8 user's reference guide* (2nd ed.). Chicago, IL: Scientific Software International.
- Kirby, J. R., Desrochers, A., Roth, L., & Lai, S. S. V. (2008). Longitudinal predictors of word reading development. *Canadian Psychology/Psychologie Canadienne*, 49(2), 103–110 <https://doi.org/10.1037/0708-5591.49.2.103>.
- Krom, R., Jongen, I., Verhelst, N., Kamphuis, F., & Kleintjes, F. (2010). *Wetenschappelijke verantwoording DMT en AVI* [Scientific justification Three Minutes Test and Analysis of Individualization Formats]. Arnhem: Stichting CITO Instituut voor Toetsontwikkeling.
- Landerl, K., Ramus, F., Moll, K., Lyytinen, H., Leppänen, P., Lohvansuu, K., ... Schulte-Körne, G. (2013). Predictors of developmental dyslexia in European orthographies with varying complexity. *Journal of Child Psychology and Psychiatry*, 54(6), 686–694 <https://doi.org/10.1111/jcpp.12029>.

- Lindman, H. R. (1974). *Analysis of Variance in complex experimental design*. New York: W.H. Freeman and Co.
- Little, T. D. (2013). *Methodology in the social sciences: Longitudinal structural equation modeling*. New York, USA: The Guilford Press.
- Lyytinen, H., Erskine, J., Tolvanen, A., Torppa, M., Poikkeus, A.-M., & Lyytinen, P. (2006). Trajectories of reading development: A follow-up from birth to school age of children with and without risk for dyslexia. *Merrill-Palmer Quarterly*, *52*(3), 514–546 <https://doi.org/10.1353/mpq.2006.0031>.
- McCardle, P., Scarborough, H. S., & Catts, H. W. (2001). Predicting, explaining, and preventing children's reading difficulties. *Learning Disabilities Research and Practice*, *16*(4), 230–239 <https://doi.org/10.1111/0938-8982.00023>.
- Melby-Lervåg, M., Lyster, S. A. H., & Hulme, C. (2012). Phonological skills and their role in learning to read: A meta-analytic review. *Psychological Bulletin*, *138*(2), 322–352 <https://doi.org/10.1037/a0026744>.
- Moll, K., Loff, A., & Snowling, M. J. (2013). Cognitive endophenotypes of dyslexia. *Scientific Studies of Reading*, *17*(6), 385–397 <https://doi.org/10.1080/10888438.2012.736439>.
- Moll, K., Ramus, F., Bartling, J., Bruder, J., Kunze, S., Neuhoff, N., ... Landerl, K. (2014). Cognitive mechanisms underlying reading and spelling development in five European orthographies. *Learning and Instruction*, *29*, 65–77 <https://doi.org/10.1016/j.learninstruc.2013.09.003>.
- Mommers, C. L., Verhoeven, L., Koekebacker, E., Van der Linden, S., Stegeman, W., & Warnaar, J. (2003). *Veilig leren lezen [Learning to read safely]*. Tilburg, Netherlands: Zwijzen.
- Nelson, R. J., Benner, G. J., & Gonzalez, J. (2003). Learner characteristics that influence the treatment effectiveness of early literacy interventions: A meta-analytic review. *Learning Disabilities Research and Practice*, *18*(4), 255–267 <https://doi.org/10.1111/1540-5826.00080>.
- Olson, R.K., Keenan, J. M., Byrne, B., & Samuelsson, S. (2014). Why do children differ in their development of reading and related skills? *Scientific Studies of Reading*, *18*(1), 38–54 <https://doi.org/10.1080/10888438.2013.800521>.
- Oslund, E. L., Simmons, D. C., Hagan-Burke, S., Kwok, O. M., Simmons, L. E., Taylor, A. B., & Coyne, M. D. (2015). Can curriculum-embedded measures predict the later reading achievement of kindergarteners at risk of reading disability? *Learning Disability Quarterly*, *38*(1), 3–14 <https://doi.org/10.7319/ldq.2014.38.1.14524752>.
- Pennington, B. F., & Lefly, D. L. (2001). Early reading development in children at family risk for dyslexia. *Child Development*, *72*(3), 816–833 <https://doi.org/10.1111/1474.227.30>.
- Pennington, B. F., & Olson, R. K. (2005). Genetics of Dyslexia. In M. J. Snowling, & C. Hulme (Eds.), *The science of reading: A handbook* (pp. 453–472). Malden, USA: Blackwell Publishing <https://doi.org/10.1002/9780470757642.ch24>.
- Perfetti, C. A. (1992). The representation problem in reading acquisition. In P. B. Gough, L. C. Ehri, & R. Treiman (Eds.), *Reading acquisition* (pp. 145–174). Hillsdale, England: Lawrence Erlbaum Associates.
- Perfetti, C. A. (2007). Reading ability: Lexical quality to comprehension. *Scientific Studies of Reading*, *11*(4), 357–383 <https://doi.org/10.1080/10888430701530730>.
- Pfost, M., Hattie, J., Dörfler, T., & Artelt, C. (2013). Individual differences in reading development: A review of 25 years of empirical research on matthew effects in reading. *Review of Educational Research*, *84*(2), 203 <https://doi.org/10.1080/00346543.2013.800949>.
- Pieters, S., Dek, J.E., & Kooij, A.P. (2013). RAKIT-2 psychometrische eigenschappen deel 2. Retrieved from: <http://www.pearsonclinical.nl/rakit-2-revisie-amsterdamse-kinder-intelligentietest> [26 April 2014]
- Puolakanaho, A., Ahonen, T., Aro, M., Eklund, K., Leppänen, P. H., Poikkeus, A. M., ... Lyytinen, H. (2007). Very early phonological and language skills: Estimating individual risk of reading disability. *Journal of Child Psychology and Psychiatry*, *48*(9), 923–931 <https://doi.org/10.1111/j.1469-7610.2007.01763.x>.
- Raven, J. C. (1958). *Guide to using the coloured progressive matrices*. Oxford, England: Lewis & Co.
- Scarborough, H. S. (1989). Prediction of reading disability from familial and individual differences. *Journal of Educational Psychology*, *81*(1), 101 <https://doi.org/10.1037/0022-0663.81.1.101>.



- Schaars, M. M. H., Segers, E., Verhoeven, L. (2017). Word decoding development in incremental phonics instruction in a transparent orthography. *Reading & Writing*, 1–22 <https://doi.org/10.1007/s1145-017-9735-3>.
- Seymour, P. H. K., Aro, M., Erskine, J. M., & COST Action A8 network (2003). Foundation literacy acquisition in European orthographies. *British Journal of Psychology*, 94, 143–174 <https://doi.org/10.1348/000712603321661859>.
- Share, D. L. (1995). Phonological recoding and self-teaching: Sine qua non of reading acquisition. *Cognition*, 55, 151–218 <https://doi.org/0010-0277/94/00645-8>.
- Share, D. L. (2004). Orthographic learning at a glance: On the time course and developmental onset of self-teaching. *Journal of Experimental Child Psychology*, 87(4), 267–298 <https://doi.org/10.1016/j.jecp.2004.01.001>.
- Snowling, M. J., & Melby-Lervåg, M. (2016). Oral language deficits in familial dyslexia: A meta-analysis and review. *Psychological Bulletin*, 1, 1–48 <https://doi.org/10.1037/bul0000037>.
- Snowling, M. J., Gallagher, A., & Frith, U. (2003). Family risk of dyslexia is continuous: Individual differences in the precursors of reading skill. *Child Development*, 74(2), 358–373 <https://doi.org/10.1111/1467-8624.7402003>.
- Snowling, M. J., Muter, V., & Carroll, J. (2007). Children at family risk of dyslexia: A follow-up in early adolescence. *Journal of Child Psychology and Psychiatry*, 48(6), 609–618 <https://doi.org/10.1111/j.1469-7610.2006.01725.x>.
- SPSS, I. (2015). IBM SPSS statistics for Windows, version 23.0. New York: IBM Corp.
- Swagerman, S. C., Van Bergen, E., Dolan, C., de Geus, E. J., Koenis, M. M., Pol, H. E. H., & Boomsma, D. I. (2015). Genetic transmission of reading ability. *Brain and Language*. <https://doi.org/10.1016/j.bandl.2015.07.008>.
- Tabachnick, B. G., & Fidell, L. S. (2014). *Using Multivariate statistics*. Pearson Education Limited.
- Tilanus, E. A., Segers, E., & Verhoeven, L. (2013). Diagnostic profiles of children with developmental dyslexia in a transparent orthography. *Research in Developmental Disabilities*, 34(11), 4194–4202 <https://doi.org/10.1016/j.ridd.2013.08.039>.
- Van der Leij, A., & Van Daal, V. H. (1999). Automatization aspects of dyslexia: Speed limitations in word identification, sensitivity to increasing task demands, and orthographic compensation. *Journal of Learning Disabilities*, 32(5), 417–428 <https://doi.org/10.1177/002221949903200507>.
- Van der Leij, A., Bergen, E., Zuijlen, T., De Jong, P., Maurits, N., & Maassen, B. (2013). Precursors of developmental dyslexia: An overview of the longitudinal Dutch dyslexia programme study. *Dyslexia*, 19(4), 191–213 <https://doi.org/10.1002/dys.1463>.
- Van Bergen, E., De Jong, P. F., Regtvoort, A., Oort, F., Van Otterloo, S., & Van der Leij, A. (2011). Dutch children at family risk of dyslexia: Precursors, reading development, and parental effects. *Dyslexia*, 17(1), 2–18 <https://doi.org/10.1002/dys.423>.
- Van Bergen, E., De Jong, P. F., Maassen, B., & Van der Leij, A. (2014). The effect of parents' literacy skills and children's preliteracy skills on the risk of dyslexia. *Journal of Abnormal Child Psychology*, 42(7), 1187–1200 <https://doi.org/10.1007/s10802-014-9858-9>.
- Verhoeven, L., & Perfetti, C. A. (2011). Morphological processing in reading acquisition: A cross-linguistic perspective. *Applied Psycholinguistics*, 32, 457–466 <https://doi.org/10.1017/S0142716411000154>.
- Verhoeven, L., & Van Leeuwe, J. (2008). Prediction of the development of reading comprehension: A longitudinal study. *Applied Cognitive Psychology*, 22, 407–423 <https://doi.org/10.1002/acp.1414>.
- Verhoeven, L., & Vermeer A. (1986). *Taaltoets allochtone kinderen* [Language test ethnic minority children]. Tilburg, Netherlands: Uitgeverij Zwijzen.
- Verhoeven, L., Van Leeuwe, J., Irausquin, R., & Segers, E. (2016). The unique role of lexical accessibility in predicting kindergarten emergent literacy. *Reading and Writing*, 29(4), 591–608 <https://doi.org/10.1007/s1145-015-9614-8>.
- Ziegler, J. C., Perry, C., & Zorzi, M. (2014). Modelling reading development through phonological decoding and self-teaching: Implications for dyslexia. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 369 <https://doi.org/20120397>.