Jyväskylän yliopiston julkaisuarkisto
Jyväskylä University Digital Archive

UNIVERSITY OF JYVÄSKYLÄ

This is an electronic reprint of the original article.
This reprint *may differ* from the original in pagination and typographic detail.

Author(s): Torppa, Minna; Georgiou, George; Salmi, Paula; Eklund, Kenneth; Lyytinen, Heikki

Title: Examining the double-deficit hypothesis in an orthographically consistent language

Year: 2012

Version:

Please cite the original version:

Torppa, M., Georgiou, G., Salmi, P., Eklund, K., & Lyytinen, H. (2012). Examining the double-deficit hypothesis in an orthographically consistent language. Scientific Studies of Reading, 16(4), 287-315. https://doi.org/10.1080/10888438.2011.554470

All material supplied via JYX is protected by copyright and other intellectual property rights, and duplication or sale of all or part of any of the repository collections is not permitted, except that material may be duplicated by you for your research use or educational purposes in electronic or print form. You must obtain permission for any other use. Electronic or print copies may not be offered, whether for sale or otherwise to anyone who is not an authorised user.

Running head: DOUBLE-DEFICIT HYPOTHESIS IN FINNISH

Examining the Double-Deficit Hypothesis in an Orthographically Consistent Language

Minna Torppa

University of Jyväskylä, Finland

George Georgiou

University of Alberta, Canada

Paula Salmi

University of Jyväskylä, Finland

Kenneth Eklund

University of Jyväskylä, Finland

Heikki Lyytinen

University of Jyväskylä, Finland

Torppa, M., Georgiou, G., Salmi, P., Eklund, K., & Lyytinen, H. (2012). Examining the double-deficit hypothesis in an orthographically consistent language. *Scientific Studies of Reading, 16*, 287-315.

This study was funded by a grant from the Academy of Finland to Finnish Centre of Excellence in Human Development and Its Risk Factors (# 44858) and a grant to Minna Torppa (#2106311).Correspondence concerning this article should be addressed to Minna Torppa, Department of Psychology, PL35, 40014 University of Jyväskylä, Jyväskylä, Finland. e-mail: minna.p.torppa@jyu.fi. Phone: + (358) 2602897.

Abstract

We examined the double-deficit hypothesis in Finnish. One-hundred-and five Finnish children with high familial risk for dyslexia and 90 children with low family risk were followed from the age of 3.5 years until grade 3. Children's phonological awareness, rapid naming speed, text reading and spelling were assessed. A deficit in RAN predicted slow reading speed across time and spelling difficulties after Grade 1. A deficit in phonological awareness predicted difficulties in spelling, but only in the familial risk sample. The effect of familial risk was significant in the development of phonological awareness, RAN, reading, and spelling. Our findings suggest that the basic premise of the double-deficit hypothesis – that RAN and phonological awareness are separable deficits with different effects on reading and spelling outcomes – holds also in a consistent orthography.

Examining the Double-Deficit Hypothesis in an Orthographically Consistent Language

Over the last three decades, phonological awareness, broadly defined as the ability to perceive and manipulate the sounds of spoken words, has dominated the research in probable causes of reading difficulties (e.g., Hulme, Snowling, Caravolas, & Carroll, 2005; Scarborough, 1998a; Vellutino, Fletcher, Snowling, & Scanlon, 2004). Several studies have found that children with impaired phonological awareness in kindergarten or in grade 1 experience reading difficulties later on (e.g., Puolakanaho et al., 2008; Savage, Carless, & Ferraro, 2008), that deficits in phonological awareness are likely a universal characteristic of children with dyslexia (e.g., Caravolas, Vólin, & Hulme, 2005), and that they persist into adulthood (e.g., Parrila, Georgiou, & Corkett, 2007; Wilson & Lesaux, 2001).

Although the phonological deficit hypothesis has been able to account for the large proportion of reading impairments, it has not been able to explain the heterogeneity of deficits observed in children and adults with dyslexia. Wolf and Bowers (1999) proposed an alternative conceptualization of dyslexia – *the double-deficit hypothesis* (DDH) – according to which phonological deficits and deficits in rapid automatized naming (RAN) speed are largely independent sources of reading difficulties. According to the DDH, the majority of poor readers can be classified into three groups: two with a single deficit in either phonological awareness or RAN, and one with a double deficit in both phonological awareness and RAN. Because phonological awareness and RAN deficits are assumed to have independent negative effects on reading, children in the double-deficit group are expected to experience the most severe reading problems.

To date, several studies have examined the DDH (see Kirby, Georgiou, Martinussen, & Parrila, 2010; Vukovic & Siegel, 2006; for reviews) and the evidence from these studies is

mixed. Several studies have found that children with the double-deficit are the poorest readers overall (e.g., Katzir, Kim, Wolf, Morris, & Lovett, 2008; Kirby, Parrila, & Pfeiffer, 2003; Lovett, Steinbach, & Frijters, 2000; Papadopoulos, Georgiou, & Kendeou, 2009; Powell, Stainthorp, Stuart, Garwood, & Quinlan, 2007; Wolf, O'Rourke, Gidney, Lovett, Cirino, & Morris, 2002). In contrast, other studies report that children in the double-deficit group did not perform significantly poorer on reading than children in the single-deficit groups or, in some cases, the no-deficit group (e.g., Ackerman, Holloway, Youngdahl, & Dykman, 2001; Pennington, Cardoso-Martins, Green, & Lefly, 2001; Vaessen, Gerretsen, & Blomert, 2009).

Mixed results may be due to variations in the sample characteristics, in the cut-off scores used to identify the groups, and in the orthographies in which the DDH was examined. For example, some studies have used unselected samples of children (e.g., Manis, Doi, & Bhadha, 2000; Papadopoulos et al., 2009; Powell et al., 2007; Sunseth & Bowers, 2002), whereas others have used reading-disabled children (e.g., Katzir et al., 2008; Lovett et al., 2000; Vaessen et al., 2009; Wolf et al., 2002). In addition, the cut-off score used to define disability in phonological awareness and RAN has varied. For example, Bowers, Sunseth, and Golden (1999) used the 35th percentile, Manis et al. (2000) the 25th percentile, and Papadopoulos et al. (2009) the 20th percentile. Finally, studies on the DDH have been conducted in languages varying in orthographic consistency (e.g., English: Wolf et al., 2002; Spanish: Escribano, 2007; German: Wimmer, Mayringer, & Landerl, 2000; Greek: Papadopoulos et al., 2009). In orthographically consistent languages such as Finnish, Greek, or German, the main difficulty for children with dyslexia is not decoding accuracy but reading speed (e.g., Holopainen, Ahonen, & Lyytinen, 2001; van Daal & van der Leij, 1999; Wimmer, 1993), whereas in less consistent writing systems both accuracy and speed are problematic (see reviews in Snowling, 2000; Vellutino et al., 2004).

Vukovic and Siegel (2006) suggested that longitudinal studies are needed to improve our understanding of the DDH. To our knowledge, only three studies have examined the DDH longitudinally and have provided partly inconsistent findings (Kirby et al., 2003; Papadopoulos et al., 2009; Wimmer et al., 2000). Wimmer et al. (2000) examined the DDH with two samples of German-speaking children; the first was followed from school entry to grade 3 (Study 1) and the second from kindergarten to grade 4 (Study 2). The three deficit groups along with the no-deficit group (identified at the first assessment) were compared on reading accuracy and rate, and on spelling in grades 3 or 4. The double-deficit and naming-deficit groups differed significantly from both the no-deficit and the phonological-deficit groups on most of the speed measures. With respect to spelling, all three deficit groups performed significantly poorer than the no-deficit group without any reliable differences observed between the deficit groups. In reading accuracy, all three deficit groups showed close to ceiling performance. Wimmer et al. (2000) interpreted the high scores of the deficit group on reading accuracy as a function of two factors: the easy orthographic system in which German-speaking children are learning to read and the synthetic phonics teaching approach that emphasizes phonological coding. Although Wimmer et al.'s (2000) findings illustrate the areas in which the deficit groups may differ in an orthographically consistent language, they provide little information as to how the three deficit groups developed across time.

This limitation was addressed by Kirby et al. (2003) who examined the DDH in a longitudinal study extending from kindergarten until grade 5 in English. The three deficit groups along with the no-deficit group were compared on word identification, word attack, and passage comprehension. The participants in the no-deficit group performed consistently well and the participants in the double-deficit group performed consistently poorly. Participants with single

phonological deficits performed poorly at the beginning, but then approached the no-deficit group on reading. The participants in the naming-deficit group did poorly throughout, almost as poorly as the double-deficit participants. Kirby et al. (2003) concluded that the double-deficit group lagged behind the no-deficit group by almost two years of achievement and were showing no sign of catching up.

Finally, Papadopoulos et al. (2009) examined the DDH in a longitudinal study with Greek-speaking children followed from kindergarten until grade 2. Papadopoulos et al. formed the deficit groups in grade 1 and then examined their performance retrospectively (in kindergarten) and prospectively (in grade 2). The three deficit groups along with the no-deficit group were compared in grades 1 and 2 on reading accuracy and speed, orthographic processing, and reading comprehension. In grade 1, the double-deficit group performed significantly poorer than the other three groups on all measures. The single-deficit groups also performed poorer than the no-deficit group on both reading accuracy and speed tasks and the phonological-deficit group also performed poorer than the no-deficit group in orthographic processing. In grade 2, there was a clear tendency to catch up, particularly with respect to reading accuracy. Gains were evident in reading accuracy, particularly for the phonological-deficit group, but the difficulties in reading speed persisted in the naming-deficit and double-deficit groups.

In summary, the previous longitudinal studies have offered support for the DDH, although there are notable inconsistencies between the studies. For example, Kirby et al. (2003) showed that the double-deficit group performed consistently poorly on reading accuracy in English. In contrast, Wimmer et al. (2000) found no significant differences among the groups on reading accuracy in German. Inconsistent findings have been reported even between orthographically consistent languages (German and Greek): Papadopoulos et al. (2009) found no

differences between the naming-deficit and the phonological-deficit groups on reading speed, whereas Wimmer et al. (2000) found that the naming-deficit group performed significantly worse than the phonological-deficit group on reading speed. A potential reason for this inconsistency may be the different age and reading level of the children; in Wimmer et al.'s (2000) study, the DDH grouping was based on assessment prior to the beginning of reading instruction, whereas in Papadopoulos et al.'s (2009) study it was based on assessment at the end of grade 1 and after reading instruction had already exerted some influence on the classification measures.

Overview of the Present Study

The purpose of the present study was to examine longitudinally the DDH in an orthographically consistent language (Finnish). To our knowledge, this is the first study to examine the DDH in Finnish. Compared to previous longitudinal studies the current study makes three important contributions: First, we included a sample of children with high familial risk for dyslexia (and matched controls) as indexed through familial incidence in first degree relatives. The inclusion of the high-risk sample contrasts markedly with the relatively small unselected school samples used in previous longitudinal studies (e.g., Kirby et al., 2003; Papadopoulos et al., 2009; Wimmer et al., 2000). The probability to identify children with phonological and/or naming difficulties is higher in a familial risk sample than in unselected samples because the amount of children with genetic vulnerability for cognitive deficit is higher (e.g., Lyytinen et al., 2004; Scarborough, 1990; Snowling, Gallagher, & Frith, 2003). The incidence of dyslexia is also higher in a familial risk sample for dyslexia: In the present sample, 35.8% of the high-risk children were found to have dyslexia in grade 2 whereas the comparable percentage in the low-risk sample was 9.8% (Puolakanaho et al., 2008). In Puolakanaho et al.'s (2008) study, familial

risk was a significant predictor of dyslexia in addition to phonological awareness, RAN, and letter knowledge. In the present study, we compare and describe in more detail the prediction of reading and spelling development from a double-deficit viewpoint in the two samples.

Second, we trace the development of the double-deficit groups in RAN and phonological awareness from the age of 3.5 years until grade 3 (approaching age 10). This developmental span is significantly longer than in previous longitudinal studies, begins earlier than in previous studies and, importantly, begins prior to the implementation of formal reading instruction. Previous studies have examined across time correlations of phonological awareness and RAN (e.g., Puolakanaho et al., 2008; Scarborough, 1998b), but the consistency of phonological and RAN deficit across time has rarely been investigated. Spector (2005) reported a considerable amount of instability on the phonological awareness and RAN deficits. High instability of the groups may reflect either true changes in the skills or problems with the reliability of measures, or in the cut-off criteria, all of which weaken the prediction power. Spector's (2005) results were from a sample of school-age children and thus likely confounded by the effects of reading acquisition and teaching. In the present study we examined stability between ages 3.5 and grade 3 across five measurement occasions.

Third, we control for early reading skill (before school entry, age 6.5 years) as a possible confounding variable to group differences in literacy outcomes. It is critical that the groups are matched on their reading ability at the outset of the study; otherwise later differences in reading could simply reflect an early advantage in reading. In line with Kirby et al. (2003) and Wimmer et al. (2000), we formed our deficit groups prior to the commencement of formal reading instruction. However, neither Wimmer et al. (2000) nor Kirby et al. (2003) mentioned whether there were children in the no-deficit group who could already read some words in kindergarten.

The control for reading level even prior to school entry is important in this study because Finnish is a highly consistent orthography and about one third of Finnish children can read before they enter formal education (Lerkkanen, Rasku-Puttonen, Aunola, & Nurmi, 2004; Silvén, Poskiparta, & Niemi, 2004).

In this study we asked the following questions:

1. Can we identify the DDH groups in the two samples varying in familial risk for dyslexia?
2. How stable are deficits in RAN and phonological awareness between 3.5 years and grade 3?
3. Do phonological and/or naming deficits predict difficulties in reading and spelling development from grade 1 to grade 3?

Method

Participants

Participants were 194 children followed from birth as part of the Jyväskylä Longitudinal Study of Dyslexia (JLD, e.g., Lyytinen et al., 2004, 2008). Participating families were recruited with the help of maternity clinics throughout Central Finland. Children born to families with one or both parents diagnosed as dyslexic were assigned to the high-risk group (n = 105). Those born to families with no report of reading difficulties were assigned to the low risk group (n = 90). Parental risk status was confirmed with extensive individual assessment comprising reading/spelling, phonological, and orthographic processing (see Leinonen Müller, Leppänen, Aro, Ahonen, & Lyytinen, 2001). The high-risk and low-risk samples were matched on their parental educational level. All the children spoke Finnish as their native language and had no mental, physical, or sensory impairments. Of the total 222 children recruited, 22 withdrew from the study. Six children were omitted from this study due to missing data.

The geographic distribution of the sample across Central Finland ensured few of the participating children were classmates. However, Finnish schools follow the same national curriculum and the classroom level effect on student's reading skill variance is small (e.g., Torppa et al., 2007).

Demographic information (gender, age, IQ, and parents' dyslexia status and education) is presented in Table 1. No significant differences emerged between the high-risk and low-risk samples in terms of gender ($\chi^2(1) = .86$, $p = .35$), age (1st grade: $F(1, 194) = 1.34$, $p = .25$, 2nd grade: $F(1, 194) = .00$, $p = .99$, and 3rd grade: $F(1, 194) = .31$, $p = .58$), the education levels of the parents (mothers: $\chi^2(6) = 1.38$, $p = .29$, fathers: $\chi^2(6) = 7.04$, $p = .32$), or the WISC-III performance IQ ($F(1, 188) = .01$, $p = .92$). Samples were, however, different in the WISC-III verbal IQ ($F(1, 189) = 8.80$, $p < .001$). In addition, dyslexia in grade 2 was more common in the at-risk sample than in the low-risk sample ($\chi^2(1) = 17.83$, $p < .001$). Of the 41 children who were diagnosed with dyslexia in grade 2 (see Puolakanaho et al., 2007, for a detailed description of the criteria to identify the dyslexic children), 33 came from the high-risk sample and eight from the low-risk sample.

Measures and Procedure

RAN, phonological awareness, reading (accuracy and speed), spelling, and IQ estimates were all administered individually by a trained examiner. The assessments were scheduled according to children's age prior to school entry. From school entry onwards, the assessments were scheduled according to the month of a school year and thus children's ages varied. Consequently, the effect of age was examined at school age. Where standardized variables were used, all standardizations were based on the JLD low-risk sample's distribution because it was considered as being closer to the age cohort.

General Cognitive Ability

The Wechsler Intelligence Scale for Children – Third Edition (WISC-III, Wechsler, 1991) was administered in grade 2, and the scale scores of four performance quotient subtests (Picture Completion, Block Design, Object Assembly, and Coding) were used to form the performance IQ measure. Similarly, verbal IQ was calculated from the scale scores of five subtests: Similarities, Vocabulary, Comprehension, Series of Numbers, and Arithmetic. The total IQ measure was calculated from the Verbal and Performance scale scores according to the manual.

Phonological Awareness

At age 3.5 years, phonological awareness was measured by identification of word size segments, identification of subword segments, and blending. At ages 5.5 and 6.5 years, phonological awareness was measured by identification of subword segments, initial phoneme identification, initial phoneme production, and blending. Composite scores of these tasks were formed by calculating the means of the standardized scores of each task. Cronbach's alpha reliability coefficients for the composite scores at each age were .45 at age 3.5 years, .76 at age 5.5 years, and .82 at age 6.5 years. At school age, phonological awareness was measured by a common unit task (see below).

Identification of word size segments from compound words. The child was presented with three pictures of objects with simultaneous pronunciation of the name of the illustrated objects (e.g., 'lentokone' [aeroplane]; 'soutuvene' [rowing boat]; 'polkupyörä' [bicycle]). All targets were compound words and the child was required to identify the picture on a computer screen containing a specified part of the compound (e.g., In which picture can you hear the sound 'kone' (plane)?). Two practice items were administered for each child, and a third practice

followed automatically if the child failed the first two. There were eight items in this computer-aided task at age 3.5 years (for details, see Puolakanaho, Poikkeus, Ahonen, Tolvanen, Poikkeus, & Lyytinen, 2003) and the Cronbach's alpha reliability coefficient was .71.

Identification of subword segments. The task was the same as in the word level identification task but with the requirement to identify sub-word level units within the target (e.g., the /koi/ in the word 'koira' (dog)). No practice items were given as the procedure was identical to the previous task. The size of the segment to be identified varied from one to four phonemes, and from two syllables to one phoneme. Segments came from the beginning, end, or middle part of the word. There were eight items in this computer-aided task at age 3.5 years, and 20 items at 5.5 and 6.5 years. Cronbach's alpha reliability coefficients were .53 at age 3.5 years, .82 at 5.5 years, and .76 at age 6.5 years.

Blending. The child was presented with segments (words, syllables, or phonemes), each separated by 750 msec., with the task requirement to blend the segments and to pronounce the resulting word that was presented in small pieces (e.g., *per-ho-nen* (butterfly)). One test item consisted of a compound word, eight items required synthesis of syllables, and three items required synthesis of syllables and phonemes. Only a response containing the correctly assembled form was coded as correct and one point was awarded for each correct answer. The principle of assembly was reinforced in two practice items. There were nine items in this computer-aided task at age 3.5 years, and 16 items at 5.6 and 6.5 years. Cronbach's alpha reliability coefficients were .78 at age 3.5 years, .78 at 5.5 years, and .80 at age 6.5 years.

Initial sound identification. This task entailed the child being shown four pictures of objects with the simultaneous presentation of the object name. The child was then required to select the correct picture on the basis of the oral presentation of a subsequent initial phoneme

relating to one target (e.g., “In the beginning of which word do you hear ____?”). There were nine items measured at ages 5.5 and 6.5 years. Cronbach’s alpha reliability coefficients were .95 at age 5.5 years, and .65 at age 6.5 years.

Initial sound production. The experimenter showed a picture to the child and asked what she/he saw in the picture. Next, the child was asked to listen the word and to articulate the initial sound (phoneme or letter name) of the object. The sum of the correct sounds or initial letter answers formed the score of the task. There were eight items at age 5.5 years and 10 items at age 6.5 years. Cronbach’s alpha reliability coefficients were .87 at age 5.5 years, and .94 at age 6.5 years.

Common unit. In grade 1, the child was asked to say aloud the sound which she/he heard to be similar in both words of a word pair heard via headphones. The length of the similar sound was always one phoneme, and it was situated either at the beginning (e.g., **l**ahja – **l**intu) or at the middle (e.g., muu**t**os – kei**t**in) of the word (four word pairs of each type were used). The length of the words varied between four to six phonemes. Before formal assessment, the children performed two practice trials. Correct responses were praised with emphasis on the common sound and the desired response. Throughout testing, the question put to the child was “which sound is the same in ...?” No reference was made to positional concepts, such as “beginning” or “in the middle”. Responses were scored as correct if they corresponded to the shared segment of sound in the word pairs. Both the letter names and phoneme sounds were scored as correct. The sum of correct responses on the eight items was used as the individual’s score. In grade 3, the task was similar to the grade 1 task, with the exception that nonwords (5-12 phonemes in length) were used instead of words. There were three practice nonword pairs followed by 15 experimental items. The position of the common phoneme varied between the words in a

nonword pair as well as between the pairs (e.g., nohdi**t** – la**t**sukurje, tuvin**o**iski – r**o**lla). The sum of correct responses out of 15 trials was used as an individual's score. Cronbach's alpha reliability coefficient was .74 in grade 1 and .81 in grade 3.

Rapid Naming Speed

RAN objects and colors were assessed at age 3.5, 5.5, 6.5, 1 st grade January and 3rd grade November using the standard procedure (see Denckla & Rudel, 1976) in which the child is asked to name as rapidly as possible a series of five different visual stimuli. All the Finnish names for the objects were two syllable high frequency words and all color names were either two or three syllables long. The child's performance was timed, and the errors and self-corrections were recorded. The score at ages 3.5 and 5.5 years was the time to name 30 items dispersed in five rows of six (5 stimuli by 6 times). From age 6.5 years onwards the traditional matrix of 50 items (5 stimuli by 10 times) was used. Total matrix completion naming time (in seconds) was used as the child's score. Because only few naming errors occurred, they were not considered further. The RAN composite score was the mean on the standardized object and color naming tasks at ages 5.5 and 6.5 years. At age 3.5 and at school assessments color naming was not available and the RAN score was only the standardized object naming task. All RAN tasks were multiplied by -1 to ease interpretation (a higher RAN score means better performance). The test-retest correlations were high: between the composite scores of RAN objects and colors between ages 5.5 and 6.5 correlation coefficient was .67, between RAN objects at grade 1 and grade 3 it was .66 and the correlation between the 6.5 year RAN composite and grade 1 RAN objects was .63. However, 3.5 year RAN objects correlation coefficient was .30 with 5.5 year RAN composite.

Nonword Reading

In kindergarten (age 6.5 years), the children were asked to read a list of nine bisyllabic (vcv, cvcv, vcvc) nonwords (e.g., ame, hopa, olus). Before each task the participants were presented with a practice list with three items. The nonword items were derived from a test battery compiled as part of a pan-European collaboration (COST A8 Action 'Learning Disorders as a Barrier to Human Development', for more details, see Aro & Wimmer, 2003 and Seymour et al., 2003). The score in each task was the number of items the child read accurately. Cronbach's alpha reliability coefficient was .87.

Text-Reading Accuracy and Speed

A text reading task was used to measure reading accuracy and speed. This task was administered individually in a laboratory setting as a part of JLD assessment phase at first grade (May), second grade (June), and third grade (April). The text was a fictional story consisting of 124 words / 901 letters in first and second grade, and a more difficult story in third grade consisting of 189 words / 1189 letters. Reading accuracy score is the percentage of correct words read and the reading speed measure is the amount of words per minute a child was able to read.

Spelling

Spelling was assessed by spelling of nonwords. Nonwords were presented separately via headphones by using the sound files accessed through the computer, and the children were asked to write them on a piece of paper. An individual's score was the number of correctly spelled nonwords. There were 18 items (nine bi-syllabic and nine more complex nonwords) in first grade and 12 items (all four-syllabic) in second and third grade. The spelling items where the same in grades 2 and 3 but different in grade 1. Cronbach's alpha reliability coefficients were .85 in grade 1, .86 in grade 2, and .71 in grade 3.

Results

Preliminary Data Analyses

The distributions of RAN, reading accuracy, and grade 1 spelling deviated from normality. The few extremely low response times for RAN were moved to the tails of the distributions. A ceiling effect emerged for reading accuracy with 70%, 75%, and 80% of grades 1, 2, and 3 children, respectively, exceeding 90% accuracy. Grade 1 spelling accuracy also produced a ceiling effect. Transformations did not correct the skewness in these measures. Because of the ceiling effect in text-reading accuracy across ages, it was not considered in further analyses. The distribution of nonword reading prior to school entry was U-shaped and for this reason we classified the children into readers and nonreaders. Ability to read at least two short nonwords correctly was considered as a proof of reading ability. This cut-off score was also supported by the distribution. Table 2 presents the descriptive statistics and the results of group comparisons on all the measures used in the study. In all of the measures included in this study the high-risk group showed, on average, lower mean level performance. In many of the measures, there was also more variability in the high-risk sample than in the low-risk sample.

Identifying the Double-Deficit Hypothesis (DDH) Groups

First, we examined whether we could identify the DDH groups. The classification was based on the RAN composite (mean of standardized colors and objects scores) and the phonological awareness composite (mean of standardized scores of segment identification, initial phoneme identification, initial phoneme production, and blending) at age 6.5 years. We chose the 20th percentile as a cut-off score from the low-risk sample distribution for the identification of the deficit groups in RAN and phonological awareness (see Figure 1 for within sample scatter plots with cut-off lines). This cut-off score is similar to the one used in previous longitudinal

studies in orthographically consistent languages (Papadopoulos et al., 2009; Wimmer et al., 2000). By using the 20th percentile cut-off score we identified a group of children (n = 25) scoring at or below the 20th percentile in both phonological awareness and RAN; they constituted the Double Deficit (DD) group. Those (n = 29) with a score at or below the 20th percentile on the RAN composite but higher than the 20th percentile in phonological awareness constituted the Naming Deficit (ND) group and those (n = 27) with a score at or below the 20th percentile in the phonological awareness composite but higher than the 20th percentile in RAN constituted the Phonological Deficit (PD) group. There were 113 children with a score above the 20th percentile on both RAN and phonological awareness that were allocated to the Double Asset (DA) group.

Because 52 children were reading at least two nonwords prior to school entry, we further divided the children based on their early reading skill. Of the 52 children who were able to read two or more nonwords at age 6.5, 47 belonged to the DA group and five belonged to the ND group. The DA group was further divided into the Double Asset Readers group (DAR, n = 47) and the Double Asset non-Readers group (DAnR, n = 66). The five ND group children who could read were omitted from further analyses. In order to avoid a potential confounding effect of low IQ, we omitted two additional children from the group comparisons because their performance IQ was below 80 at age 8. One of these children came from the DD group and the other one from the DAnR group.

Thus, the final group sizes were: DD (n = 24), PD (n = 27), ND (n = 24), DAnR (n = 65), and DAR (n = 47). There were no significant differences between the final groups in gender or in the average age at any school assessment phase, in performance IQ at age 8 years, or in parental education level. The DDH groups were found to be differently represented in the familial risk samples ($\chi^2(4) = 10.63$, $p < .05$, see Table 3 for the cross-tabulation of samples and the DDH

classification). Two thirds of children in the deficit groups were coming from the high-risk sample, while there was a slight overrepresentation of children from the low-risk sample (55% vs. 45 %) in the double asset groups. The examination of the adjusted standardized residuals revealed that the high-risk sample children were significantly over-represented in the ND group (adjusted standardized residual = 2.2).

Note that the effect of the familial risk status was examined in all of the subsequent analyses for three reasons: the sample differences were significant in each measure used, the correlations between RAN and phonological awareness at age 6.5 were lower (z = 1.87, two-sided p = .06) in the high-risk sample (r = .34) than in the low-risk sample (r = .56), and because of generalizability concerns if the two samples were to be combined.

Comparison of RAN and Phonological Awareness in the DDH Groups

Before examining the differences between the groups on reading speed and spelling, we compared them on RAN and phonological awareness. This was done to address Schatschneider, Carlson, Francis, Foorman, and Fletcher's (2002) concern that the double-deficit subtype may show the most severe deficits in reading and spelling merely because their RAN and phonological awareness difficulties are also more severe than in the single-deficit groups. The means, standard deviations, and pairwise comparisons for phonological awareness and RAN in the five DDH groups are reported in Table 4. Note that at the age of classification, the RAN performance of the DD and ND groups were almost identical. Similarly, at the age of classification, the phonological awareness performance of the DD and PD groups was almost identical. Thus, the separation of the groups on the two measures was optimal. Table 4 presents the group comparisons for RAN and phonological awareness also at other available measurement occasions. The group differences in RAN were very stable: the DD and ND groups did not differ

from each other at any of the assessment points and were poorer than DAnR and DAR groups at every point. Groups with no naming deficits (PD, DAnR, and DAR) did not differ from each other in RAN at any age. Group differences in phonological awareness, on the other hand, were less stable, but the DD and PD did not differ from each other from age 5.5 years onwards. DD group was poorer than DAnR and DAR groups at every assessment of phonological awareness, and the PD group was poorer than DAnR and DAR groups at the ages 5.5 years, 6.5 years and grade 3.

To address our second research question about the stability of the deficits in RAN and phonological awareness, we assessed how many children in the DD, ND and PD groups could be identified with RAN or phonological awareness deficit (with using the same 20th percentile cut-off criteria from the low-risk sample as we used at age 6.5 years) also at ages 3.5 years, 5.5 years, in grade 1, and in grade 3. The analysis showed reasonable stability: of the children who were classified as having a RAN deficit at age 6.5, 90.1% were classified as having a RAN deficit also in at least one other measurement point and 77.5% in three or more measurement points out of the five assessment occasions. Of the children who were classified as having a phonological awareness deficit at age 6.5, 92.2% were classified as having phonological awareness deficit also in at least one other measurement point and 64.7% in three or more measurement points out of the five assessment occasions.

Prediction of Reading and Spelling Development

Our third research question concerned the prediction of reading and spelling development. The analysis included: (a) pairwise mean comparisons of reading and spelling performance between the five DDH groups; (b) mixed methods ANOVA (see below); and (c) the cross-tabulation of DDH groups and dyslexia diagnosis at the end of grade 2.

The means and standard deviations for the five DDH groups, and the results of ANOVA pairwise comparisons of reading and spelling are presented in Table 5. In Table 5, results for the combined sample of the high-risk and low-risk samples are reported because of the small sample sizes within each DDH group. However, all group comparisons were conducted also within the samples and described in text. The comparisons between the five DDH groups on reading speed in each assessment point (see Table 5) indicated that all deficit groups had slower reading speed than the DAR group, but the two DDH groups with a RAN deficit (DD and ND) showed significant differences also from the DAnR group in grade 1, 2, and 3 reading speed. The ND group showed the slowest development in reading speed and their mean reading speed at the end of grade 3 was about half of that of the DAR group. The groups with single or double deficit did not differ significantly from each other in reading speed. The DDH group comparison results were very similar in the high-risk and low-risk samples.

In spelling, the DAR and DAnR groups were more accurate at each age than the groups with single or double deficits (with the exception of the ND group in grade 1; see Table 5). The DDH group comparisons within the familial risk samples showed that in the low-risk sample children rarely had spelling difficulties and the DDH groups were different only in grade 1 spelling where the ND group children made on average more spelling errors than the other groups. In the high-risk sample, the DDH group comparison results were similar to those of the combined sample reported in Table 5.

In order to verify the results derived from the pairwise group comparisons, we examined next the effects of familial risk status, phonological deficit, and RAN deficit on reading and spelling development with mixed method ANOVAs. The analyses included time as a within-subjects factor and three between-subjects factors: RAN deficit (no deficit, deficit), phonological

deficit (no deficit, deficit), and familial risk (high-risk, low-risk). The analysis for reading speed included three time points (grade 1, grade 2, and grade 3) but the analysis for spelling included only two time points (grade 2 and grade 3) because the nonword spelling measure was different in grade 1. Only children who were nonreaders prior to school entry were included to avoid the confounding effect of reading on the results. The results are reported in Table 6. The results for reading speed indicated that the main effects of time, familial risk status, and RAN deficit, and the interaction of RAN and time were significant. The main effects suggest that reading speed developed across time, the high-risk sample had slower reading speed than the low-risk sample, and the children with a RAN deficit were slower readers than children without a RAN deficit. A further analysis of the RAN x time interaction revealed that the reading speed development of the children with a RAN deficit was slower than that of the children without RAN deficit (the details of this analysis are available from first author).

The mixed method ANOVA for nonword spelling indicated significant main effects of time, familial risk status, and RAN deficit, and interactions for familial risk status and phonological deficit, and for RAN deficit and phonological deficit. The main effects suggest that there was a significant spelling accuracy development from grade 2 to grade 3, the high-risk sample was poorer in nonword spelling than the low-risk sample, and the children with RAN deficit were poorer in nonword spelling than the children without RAN deficit. A further examination of the interactions revealed that the interaction between familial risk status and phonological deficit reflected the difference in the effect of phonological deficit in the two samples: phonological deficit predicted spelling difficulties only in the high-risk sample (details of this analysis available from the first author). The further examination of the interaction between RAN deficit and phonological deficit revealed that it reflected the equally poor

performance of the children with any deficit (either RAN, phonological, or double) as compared to the children without deficits (see Figure 2. Details of this analysis are available from the first author).

Finally, we examined how many of the DDH group children were diagnosed as dyslexics at the end of grade 2. Table 7 reports the results separately for the high-risk, low-risk, and the combined sample. There was a statistically significant link between dyslexia diagnosis and DDH grouping ($\chi^2(4) = 18.60$, $p < .001$, for the high-risk sample, $\chi^2(4) = 24.30$, $p < .001$, for the low-risk sample, and $\chi^2(4) = 40.05$, $p < .001$, for the combined sample). None of the DAR group children (who had no phonological or naming deficits and who were readers before school age) were diagnosed as having dyslexia. The ND group had more children diagnosed with dyslexia than what would have been expected by chance alone (adjusted standardized residuals were 4.1, 2.8, and 2.1 for the total, high-risk, and low-risk samples, respectively). In addition, in the analyses with combined and low-risk samples, the DD group had more children diagnosed with dyslexia than what would have been expected by chance alone (adjusted standardized residuals were 3.2 and 4.2 for total and low-risk samples, respectively). There were two interesting differences between the high-risk and the low-risk samples; first, phonological deficit alone did not seem to predict dyslexia in the low-risk sample as only one out of ten PD children had dyslexia whereas in the high-risk sample seven out of 17 PD children (41.2%) had dyslexia later on. Second, the DAnR group children do not seem be at high risk for dyslexia if they come from a low familial risk group as only one out of 35 children (2.8%) were diagnosed with dyslexia whereas in the high-risk sample almost one third of the DAnR children (eight out of 29 children (27.6%)) were diagnosed with dyslexia at the end of grade 2.

Discussion

This study examined the DDH in Finnish, a highly consistent orthography. We asked three research questions: Can we identify the DDH groups in a sample of 6.5-year-old Finnish children? How stable are RAN and phonological deficits across time? And, does phonological and/or naming deficit predict difficulties in reading and spelling development from grade 1 to grade 3? In all analyses, we controlled for the effect of familial risk and reading ability prior to school entry.

We identified five DDH groups based on the 20th percentile cut-off in phonological awareness and RAN at age 6.5 that in Finland is just prior to the school entry at the fall of the year children turn seven years. The comparison of the development of phonological and RAN skill in the DDH groups from age 3.5 to the end of third grade showed that the group differences were most pronounced at the age of classification, but the differences were evident already at 3.5 years of age. For RAN, the group differences remained stable across school age. For phonological awareness, the group differences were less stable.

As expected, the percentage of children with RAN and/or phonological deficit was higher in the high-familial risk sample than in the low-familial risk sample, and the average skill level of RAN, phonological awareness, reading, and spelling was lower in the high-risk sample (e.g., Elbro et al., 1998; Gallagher et al., 2000, Lyytinen et al., 2004; Lyytinen et al., 2006; Puolakanaho et al., 2008; Scarborough, 1990; Snowling et al., 2003). In the high-risk group, 50% of the children had RAN, phonological or double deficit whereas the comparable percentage was 28% in the low-risk sample. Our findings complement the recent behavioral genetic studies showing that rapid naming and, to a lesser extent, phonological awareness difficulties are heritable (e.g., Petrill et al., 2010; Samuelsson et al., 2007) and that the genetic effects of RAN

and phonological awareness are partially independent (e.g., Compton, Davis, DeFries, Gayan, & Olson, 2001; Grigorenko et al., 1997). It should be noted, however, that there was more variation in the high-risk sample than in the low-risk sample. This difference in variation stems from the heterogeneity of the high-risk sample; the cases with the most severe deficits in phonological awareness and RAN difficulties were high-risk children while many others in the high-familial risk sample had above average skills in RAN, phonological awareness, reading, and spelling.

The examination of the phonological awareness and RAN deficit stability showed that over 90% of the children who were classified as having RAN or phonological awareness deficits at age 6.5 were also performing below 20th percentile in some other measurement occasion irrespective of the risk sample they came from. A more enduring deficit identified in three or more assessment points was found for 78% of the children with a RAN deficit, and 65% of the children with a phonological awareness deficit. These numbers are similar to those reported by Spector (2005) for school-aged children. The lower stability of the phonological awareness deficit may be attributed to two factors: a change in the phonological awareness measures at school entry, and the effects that formal reading instruction has on phonological awareness but not on RAN (when measured with objects and colors). Korkman, Barron-Linnankoski, and Lahti-Nuuttila (1999) demonstrated that reading instruction influences subsequent phonological awareness in Finnish, but does not have an effect on subsequent RAN. A similar nonsignificant effect of reading instruction on subsequent RAN (when measured with colors and objects) was observed by Compton (2003).

Our final research question addressed the effects of phonological and naming deficits on subsequent reading and spelling development. The results indicated that RAN deficit was predictive of slow development of text-reading speed. The effect of familial risk was also

significant, a finding that reflects the overall slower reading speed observed in the high-risk sample. These findings are in line with previous studies that have shown a unique association between RAN and reading fluency (e.g., de Jong & van der Leij, 1999; Georgiou, Parrila, & Liao, 2008; Georgiou, Parrila, & Papadopoulos, 2008; Katzir et al., 2008; Savage & Frederickson, 2005; see also Kirby et al., 2010, for a review).

In spelling, there were two interesting interactions: an interaction between phonological deficit and RAN deficit and another one between risk status and phonological deficit. A further examination of the phonological deficit and RAN deficit interaction revealed that the interaction likely resulted from the fact that there was no cumulative effect of having a double deficit – children with RAN or PA deficit alone performed equally poor in spelling than children with a double deficit. This finding of both RAN and phonological deficit being linked to spelling difficulties is similar to Wimmer et al.'s (2000) findings. The interaction between familial risk and phonological deficit reflected the fact that, in the high-risk sample, phonological deficit predicted poor performance in nonword spelling whereas in the low-risk sample it did not. The finding of phonological awareness being a good predictor of spelling among the high-risk children is easy to understand because spelling in Finnish requires sensitivity to the phoneme-level information. Finnish morphology is agglutinative in nature with very rich and complex sequential inflections and frequent stem variations (see for more detailed description of Finnish morphology in e.g. Lyytinen et al., 2006; Lyytinen & Lyytinen, 2004). Many of the morphological variations of the same words often differ only by one phoneme (e.g., the inflections of the word talo (a house) include forms such as talo*ssa* (in a house) and talo*sta* (from a house). In order to be able to use such inflections, children must have well specified phonological representations. The explanation why the spelling was not similarly predicted in the

low-risk sample needs more examination. One explanation could be the small number of children with spelling and/or phonological difficulties; however, although severe spelling and phonological difficulties were not as common in the low-risk as in the high-risk sample, there was a reasonable amount of variation also in the low-risk sample.

The effects of RAN deficit on reading speed and spelling development were persistent. Evidence showing the persistence of RAN effects has also been reported by Kirby et al. (2003). The naming speed deficit may impact the development of reading fluency and of spelling through its effects on the formation of orthographic representations, which are essential in the development of both skills. Wolf and Bowers (1999) proposed that an inadequate development of the ability to form orthographic representations for commonly seen letter strings may be caused by slow retrieval of letter identities, which is reflected in the performance on the RAN tasks. According to Wolf and Bowers' hypothesis, processes underlying slow naming speed may contribute to reading failure in three ways: (a) by impeding the appropriate amalgamation of connections between phonemes and orthographic patterns at sub–word and word levels of representation, (b) by limiting the quality of orthographic codes in memory, and (c) by increasing the amount of repeated practice needed to unitize codes before representations of adequate quality are achieved. Moll et al. (2009), however, argued that RAN is not a measure of orthographic processing; they found that RAN predicted unique variance in reading fluency even after accounting for spelling, whereas RAN accounted hardly any variance in word-reading fluency when it was entered after nonword-reading fluency.

Overall, the DDH group comparisons suggested that naming deficit predicted slow reading speed and spelling development after grade 1 and phonological deficit predicted early difficulties in spelling but only in the high-risk sample. It was also found that the double-deficit

group performed significantly more poorly in reading and in spelling than the no-deficit groups, but the performance level did not differ from the single-deficit groups.

It is also worth noting that 27.6% of the high-risk children (but only 2.8% of the low-risk sample children) who were classified to the double asset non-readers group in kindergarten were diagnosed with dyslexia in grade 2. This finding suggests that familial risk has a negative impact on reading acquisition despite adequate performance on RAN and phonological awareness (see also Puolakanaho et al., 2008), suggesting that additional risk factors contribute to the development. What these additional risk factors are is an important issue for future studies to examine. Our results suggest, however, that when assessments are carried out in kindergarten to identify children at-risk for future reading difficulties, information on familial risk should also be obtained.

Some limitations of the present study are worth reporting. First, our sample size was relatively small. In a larger dataset, more children with deficits may be identified with a more stringent cut-off score than the one used in the current study or with a more sophisticated analysis method, such as mixture models (e.g., Torppa et al., 2007). Second, we used different measures of phonological awareness across time and the interpretations regarding its development should be viewed with some caution. However, in an assessment of phonological awareness between the ages of 3.5 and 10, the measures must necessarily change. Third, the distribution of the reading and spelling accuracy measures deviated from normality. Measurement of reading and spelling accuracy in Finnish is challenging given the almost one-to-one consistent mapping between graphemes and phonemes. Fourth, because of the age of the participants at the first two measurement points, only non-alphanumeric (Colors and Objects) RAN tasks were administered. Previous studies have shown that RAN digits and letters are more

strongly related to reading than RAN colors and objects (e.g., Georgiou, Parrila, Kirby, & Stephenson, 2008; Georgiou, Parrila, & Papadopoulos, 2008; Wolf, Bally, & Morris, 1986).

To conclude, the present study offered support for the view that RAN and phonological awareness are partly independent deficits (Wolf & Bowers, 1999). Our findings indicated that RAN and phonological awareness were separable deficits and they predict different kinds of literacy problems. These findings may be useful in the development of screening and intervention strategies. In screening, the assessment of both skills, as well as information on the familial risk for dyslexia, is valuable for predicting reading and spelling development. In addition, because different problems in literacy development were predicted by partly different cognitive deficits (RAN and/or phonological awareness) and differently in the two samples, individually designed reading instruction and intervention may be needed.

ACKNOWLEDGEMENTS

The Jyvaskyla Longitudinal study of Dyslexia (JLD) has belonged to the Finnish Center of Excellence Program for two consecutive periods (2000-2005 and 2006-2011) and is supported by the Academy of Finland (#213486), the Niilo Mäki Foundation and the University of Jyväskylä. We would like to thank the families who participated in the study.

References

Aarnoutse, C., Van Leeuwe, J., & Verhoeven, L. (2005). Early literacy from a longitudinal perspective. *Educational Research and Evaluation, 11,* 253-275.

Ackerman, P., Holloway, C., Youngdahl, P., & Dykman, R. (2001). The double deficit theory of reading does not fit all. *Learning Disabilities Research and Practice, 16,* 152-160.

Aro, M., & Wimmer, H. (2003). Learning to read: English in comparison to six more regular orthographies. *Applied Psycholinguistics, 24,* 621-635.

Bowers, P. G., Sunseth, K., & Golden, J. (1999). The route between rapid naming and reading progress. *Scientific Studies of Reading, 3,* 31-53.

Caravolas, M., Vólin, J., & Hulme, C. (2005). Phoneme awareness is a key component of alphabetic literacy skills in consistent and inconsistent orthographies: Evidence from Czech and English children. *Journal of Experimental Child Psychology, 92,* 107-139.

Compton, D. L. (2003). Modeling the relationship between growth in rapid naming speed and growth in decoding skill in first-grade children. *Journal of Educational Psychology, 95,* 225-239.

Compton D. L., Davis C. J., DeFries J. C., Gayan J., & Olson R. K. (2001). Genetic and environmental influences on reading and RAN: An overview of results from the Colorado Twin Study. In M. Wolf (Ed.), *Time, fluency and the brain* (pp. 277-303). Timonium, MD: York Press.

de Jong, P. F., & van der Leij, A. (1999). Specific contributions of phonological abilities to early reading acquisition: Results from a Dutch latent variable longitudinal study. *Journal of Educational Psychology, 91,* 450-476.

Denckla, M. B., & Rudel, R. G. (1976). Rapid "automatized" naming (R.A.N.): Dyslexia differentiated from other learning disabilities. *Neuropsychologia, 14*, 471-479.

Dufva, M., Niemi, P., & Voeten, M. J. (2001). The role of phonological memory, word recognition, and comprehension skills in reading development: From preschool to grade 2. *Reading and Writing: An Interdisciplinary Journal, 14*, 91-117.

Elbro, C., Borstrøm, I., & Petersen, D. K. (1998). Predicting dyslexia from kindergarten: The importance of distinctness of phonological representations of lexical items. *Reading Research Quarterly, 33*, 36-60.

Escribano, C. (2007). Evaluation of the double-deficit hypothesis subtype classification of readers in Spanish. *Journal of Learning Disabilities, 40*, 319-330.

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

Georgiou, G., Parrila, R., & Liao, C.-H. (2008). Rapid naming speed and reading across languages that vary in orthographic consistency. *Reading and Writing: An Interdisciplinary Journal, 21*, 885-903.

Georgiou, G., Parrila, R., Kirby, J., & Stephenson, K. (2008). Rapid naming components and their relationship with phonological awareness, orthographic knowledge, speed of processing, and reading. *Scientific Studies of Reading, 12*, 325-350.

Georgiou, G., Parrila, R., & Papadopoulos, T. C. (2008). Predictors of word decoding and reading fluency in English and Greek: A cross-linguistic comparison. *Journal of Educational Psychology, 100*, 566-580.

Goswami, U. (2002). Phonology, reading development, and dyslexia: A cross-linguistic perspective. *Annals of Dyslexia, 52*, 141-163.

Grigorenko, E. L., Wood, F. B., Meyer, M. S., Hart, L. A., Speed, W. C., Shuster, A., et al. (1997). Susceptibility loci for distinct components of developmental dyslexia on chromosomes 6 and 15. *American Journal of Human Genetics, 60*, 27-39.

Holopainen, L., Ahonen, T., & Lyytinen, H. (2001). Predicting delay in reading achievement in a highly transparent language. *Journal of Learning Disabilities, 34*, 401-413.

Hulme, C., Snowling, M., Caravolas, M., & Carroll, J. (2005) Phonological skills are (probably) one cause of success in learning to read: A comment on Castles and Coltheart. *Scientific Studies of Reading, 9*, 351-365.

Katzir, T., Kim, Y.-S., Wolf, M., Morris, R., & Lovett, M. W. (2008). The varieties of pathways to dysfluent reading: Comparing subtypes of children with dyslexia at letter, word, and connected text levels of reading. *Journal of Learning Disabilities, 41*, 47-66.

Kirby, J., Georgiou, G., Martinussen, R., & Parrila, R. (2010). Naming speed and reading: From prediction to instruction. *Reading Research Quarterly, 45*, 341-362.

Kirby, J., Parrila, R., & Pfeiffer, S. (2003). Naming speed and phonological awareness as predictors of reading development. *Journal of Educational Psychology, 95*, 453-464.

Korkman, M., Barron-Linnankoski, S., & Lahti-Nuuttila, P. (1999). Effects of age and duration of reading instruction on the development of phonological awareness, rapid naming, and verbal memory span. *Developmental Neuropsychology, 16*, 415-431.

Landerl, K., & Thaler, V. (2006). Reading and spelling acquisition and dyslexia in German. In R. M. Joshi & P. G. Aaron (Eds.), *Handbook of orthography and literacy* (pp. 121-134). Mahwah, NJ: Erlbaum.

Landerl, K., & Wimmer, H. (2008). Development of word reading fluency and spelling in a consistent orthography: An 8-year follow-up. *Journal of Educational Psychology, 100*, 150-161.

Leinonen, S., Müller, K., Leppänen, P. H. T., Aro, M., Ahonen, T., & Lyytinen, H. (2001). Heterogeneity in adult dyslexic readers: Relating processing skills to the speed and accuracy of oral text reading. *Reading and Writing: An Interdisciplinary Journal, 14*, 265-296.

Leppänen, U., Niemi, P., Aunola, K., & Nurmi, J.-E. (2006). Development of reading and spelling Finnish from preschool to grade 1 and grade 2. *Scientific Studies of Reading, 10*, 3-30.

Lerkkanen, M.-K., Rasku-Puttonen, H., Aunola, K. & Nurmi, J.-E. (2004). Predicting reading performance during the first and the second year of primary school. *British Educational Research Journal, 30*, 67-92.

Lovett, M. W., Steinbach, K. A., & Frijters, J. C. (2000). Remediating the core deficits of developmental reading disability: A double-deficit perspective. *Journal of Learning Disabilities, 33*, 334-358.

Lyytinen, H., Aro, M., Eklund, K., Erskine, J., Guttorm, T., Laakso, M., et al. (2004). The development of children at familial risk for dyslexia: Birth to early school age. *Annals of Dyslexia, 54*, 184-220.

Lyytinen, H., Aro, M., Holopainen, L., Leiwo, M., Lyytinen, P., & Tolvanen, A. (2006). Children's language development and reading acquisition in a highly transparent orthography. In R. M. Joshi & P. G. Aaron (Eds.), *Handbook of orthography and literacy* (pp. 47-62). Mahwah, NJ: Erlbaum.

Lyytinen, H., Erskine, J., Ahonen, T., Aro, M., Eklund, K., Guttorm, T., et al. (2008). Early identification and prevention of dyslexia: Results from a prospective follow-up study of children at familial risk for dyslexia. In G. Reid, A. Fawcett, F. Manis, & L. Seigel (Eds.), *The SAGE handbook of dyslexia* (pp. 121-146). London: Sage Publishers.

Lyytinen, P., & Lyytinen, H. (2004). Growth and predictive relations of vocabulary and inflectional morphology in children with and without familial risk for dyslexia. *Applied Psycholinguistics, 25*, 397-411.

Manis, F. R., Doi, L. M., & Bhadha, B. (2000). Naming speed, phonological awareness, and orthographic knowledge in second graders. *Journal of Learning Disabilities, 33*, 325-333.

Moll, K., Fussenegger, B., Willburger, E. & Landerl, K. (2009). RAN is not a measure of orthographic processing. Evidence from the asymmetric German orthography. *Scientific Studies of Reading, 13*, 1-25.

Muthén, L. K., & Muthén, B. O. (2004). *Mplus User's Guide*. Los Angeles: Authors.

Papadopoulos, T. C., Georgiou, G., & Kendeou, P. (2009). Investigating the double-deficit hypothesis in Greek: Findings from a longitudinal study. *Journal of Learning Disabilities, 42*, 528-547.

Parrila, R., Georgiou, G., & Corkett, J. (2007). University students with a significant history of reading difficulties: What is and is not compensated? *Exceptionality Education Canada, 17*, 195-220.

Parrila, R., Kirby, J. R., & McQuarrie, L. (2004). Articulation rate, naming speed, verbal short-term memory, and phonological awareness: Longitudinal predictors of early reading development. *Scientific Studies of Reading, 8*, 3-26.

Pennington, B. F., Cardoso-Martins, C., Green, P. A., & Lefly, D. L. (2001). Comparing the phonological and the double deficit hypotheses of developmental dyslexia. *Reading and Writing: An Interdisciplinary Journal, 14*, 707-755.

Pennington, B. F., & Lefly, D. L. (2001). Early reading development in children at family risk for dyslexia. *Child Development, 72,* 816-833.

Petrill, S. A., Hart, S. A., Harlaar, N., Logan, J., Justice, L. M., Schatschneider, C. et al. (2010). Genetic and environmental influences on the growth of early reading skills. *Journal of Child Psychology and Psychiatry, 51,* 660-667.

Powell, D., Stainthorp, R., Stuart, M, Garwood, H., & Quinlan, P. (2007). An experimental comparison between rival theories of rapid automatized naming performance and its relationship to reading. *Journal of Experimental Child Psychology, 98,* 46-68.

Puolakanaho, A., Ahonen, T., Aro, M., Eklund, K., Leppänen, P. H. T., Poikkeus, A.-M., et al. (2007). Very early phonological and language skills: Estimating individual risk of reading disability. *Journal of Child Psychology and Psychiatry, 48,* 923-931.

Puolakanaho, A., Ahonen, T., Aro, M., Eklund, K., Leppänen, P. H. T., Poikkeus, A.-M., et al. (2008). Developmental links of very early phonological and language skills to second-grade reading outcomes: Strong to accuracy but only minor to fluency. *Journal of Learning Disabilities, 41,* 353-370.

Puolakanaho, A., Poikkeus, A.-M., Ahonen, T., Tolvanen, A., & Lyytinen, H. (2003). Assessment of three-and-a-half-year-old children's emerging phonological awareness in a computer animation context. *Journal of Learning Disabilities, 36,* 416-423.

Samuelsson, S., Olson, R., Wadsworth, S., Corley, R., Defries, J. C., Wilcutt, E. et al. (2007). Genetic and environmental influences on prereading skills and early reading and spelling

development in the United States, Australia, and Scandinavia. *Reading and Writing: An Interdisciplinary Journal*, *20*, 51-75.

Savage, R., Carless, S., & Ferraro, V. (2008). Predicting curriculum and test performance at age 11 years from pupil background, baseline skills, and phonological awareness at age 5years. *Journal of Child Psychology and Psychiatry, 48*, 732-739.

Savage, R., & Frederickson, N. (2005). Evidence of a highly specific relationship between rapid automatic naming and text-reading speed. *Brain and Language, 93*, 152-159.

Savage, R., Pillay, V., & Melidona, S. (2008). Rapid serial naming is a unique predictor of spelling in children. *Journal of Learning Disabilities, 41*, 235-250.

Scarborough, H. S. (1990). Very early language deficits in dyslexic children. *Child Development, 61*, 1728-1743.

Scarborough, H. S. (1998a). Early identification of children at risk for reading disabilities: Phonological awareness and some other promising predictors. In B. Shapiro, P. Accardo, & A. Capute (Eds.), *Specific reading disability: A view of the spectrum* (pp. 75-119). Timonium, MD: York Press.

Scarborough, H. S. (1998b). Predicting the future achievement of second graders with reading disabilities: Contributions of phonemic awareness, verbal memory, rapid naming, and IQ. *Annals of Dyslexia*, *48*, 115-137.

Schatschneider, C., Carlson, C. D., Francis, D. J., Foorman, B., & Fletcher, J. (2002). Relationships of rapid automatized naming and phonological awareness in early reading development: Implications for the double deficit hypothesis. *Journal of Learning Disabilities, 35*, 245-256.

Sénéchal, M., & LeFevre, J. (2002). Parental involvement in the development of children's reading skill: A five-year longitudinal study. *Child Development, 73*, 445-460.

Seymour, P. H. K., Aro, M., & Erskine, J. M. (2003). Foundation literacy acquisition in European orthographies. *British Journal of Psychology, 94*, 143-174.

Silvén, M., Poskiparta, E. & Niemi, P. (2004). The odds of becoming a precocious reader of Finnish. *Journal of Educational Psychology, 96*, 152-164.

Snowling, M. J. (2000). *Dyslexia* (2nd Edition). Oxford, UK: Blackwell Publishers.

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*, 358-373.

Spector, J. E. (2005). Instability of double-deficit subtypes among at-risk first grade readers. *Reading Psychology, 26*, 285-312.

Sunseth, K., & Bowers, P. G. (2002). Rapid naming and phonemic awareness: Contributions to reading, spelling, and orthographic knowledge. *Scientific Studies of Reading, 6,* 401-429.

Torppa, M., Tolvanen, A., Poikkeus, A.-M., Eklund, K., Lerkkanen, M.-K., Leskinen, E., & Lyytinen, H. (2007). Reading Development Subtypes and Their Early Characteristics. *Annals of Dyslexia, 57,* 3-52.

Vaessen A., Gerretsen P., & Blomert L. (2009). Naming problems do not reflect a second independent core deficit in dyslexia: double deficits explored. *Journal of Experimental Child Psychology, 103*, 202-221.

van Daal, V. H. P., & van der Leij, A. (1999). Developmental dyslexia: Related to specific or general deficits? *Annals of Dyslexia, 49*, 71-104.

Vellutino, F. V., Fletcher, J. M., Snowling, M. J., & Scanlon, D. M. (2004). Specific reading disability (dyslexia): What we have learned in the past four decades. *Journal of Child Psychology and Psychiatry, 45*, 2-40.

Vukovic, R. K. & Siegel, L. S. (2006). The double deficit hypothesis: A comprehensive analysis of the evidence. *Journal of Learning Disabilities, 39*, 25-47.

Wagner, R. K., Torgesen, J. K., & Rashotte, C. A. (1994). Development of reading-related phonological processing abilities: New evidence of bidirectional causality from a latent variable longitudinal study. *Developmental Psychology, 30*, 73-87.

Wechsler, D. (1991). Wechsler Intelligence Scale for Children-Third Edition. San Antonio, TX: The Psychological Corporation.

Wilson, A. M., & Lesaux, N. K. (2001). Persistence of phonological processing deficits in college dyslexics with age-appropriate reading skills. *Journal of Learning Disabilities, 34*, 394-400.

Wimmer, H. (1993). Characteristics of developmental dyslexia in a regular writing system. *Applied Psycholinguistics, 14*, 1-33.

Wimmer, H., Mayringer, H., & Landerl, K. (2000). The double-deficit hypothesis and difficulties in learning to read a regular orthography. *Journal of Educational Psychology, 92*, 668-680.

Wolf, M., Bally, H., & Morris, R. (1986). Automaticity, retrieval processes, and reading: A longitudinal study of average and impaired readers. *Child Development, 57*, 988-1000.

Wolf, M., & Bowers, P. G. (1999).The double-deficit hypothesis for the developmental dyslexias. *Journal of Educational Psychology, 91*, 415-438.

Wolf, M., O'Rourke, A. G., Gidney, C., Lovett, M. W., Cirino, P., & Morris, R. (2002). The second deficit: An investigation of the independence of phonological and naming-speed deficits in developmental dyslexia. *Reading and Writing: An Interdisciplinary Journal, 15*, 43-72.

Table 1

Demographic Information for JLD Parents and Their Children

	High-risk families	Low-risk families
Number of children	105	90
Gender (boys : girls)	53 : 52	39 : 51
Number of cases with dyslexic		
mother	54	0
father	48	0
both parents	3	0
Education[a]		
mother	4.08 (1.51)	4.46 (1.36)
father	3.70 (1.31)	3.79 (1.41)
WISC-III [a]		
VIQ (5 years)	104.38 (16.88)	110.84 (12.52)
PIQ (5 years)	101.97 (14.15)	102.17 (13.31)
Child age (years) [a)]		
1st grade May	7.90 (.30)	7.92 (.27)
2nd grade June	8.98 (.32)	8.98 (.27)
3rd grade April	9.89 (.36)	9.83 (.28)

Note. VIQ = Verbal IQ; PIQ = Performance IQ.

[a] Means and standard deviations (in parenthesis). Parental questionnaire based education was classified using a 7-point scale: 1 = only comprehensive school (CS); 2 = CS and short-term vocational courses; 3 = CS and a vocational school degree; 4 = CS and a vocational college degree; 5 = CS and a lower university degree / a polytechnic degree; 6 = upper secondary general school and a lower university degree / a polytechnic degree; 7 = CS or upper secondary general school diploma and a higher university degree (Master's or a Doctorate-level degree).

Table 2
Sample Comparisons in RAN, Phonological Awareness, Reading, and Spelling

	High-Risk Sample		Low-Risk Sample		Sample comparisons	
	M	*SD*	*M*	*SD*	*F*	Levene
RAN 6.5 y [a]	-0.67	1.71	0.03	0.92	$F(1,185) = 11.70$ ***	***
Phonological awareness 6.5 y [a]	-0.41	0.98	-0.04	0.83	$F(1,185) = 7.54$ **	ns.
Text reading accuracy (%)						
1st grade May	87.65	14.54	92.78	10.99	$F(1,177) = 6.97$ **	**
2nd grade June	91.07	8.50	94.30	4.91	$F(1,178) = 9.20$ **	***
3rd grade April	92.45	7.47	95.94	4.67	$F(1,178) = 13.75$ ***	***
Text reading speed (words/minute)						
1st grade May	30.63	20.47	41.65	23.87	$F(1,175) = 10.93$ ***	*
2nd grade June	56.14	24.67	68.38	24.75	$F(1,181) = 11.12$ ***	ns.
3rd grade April	66.98	28.61	79.97	26.23	$F(1,181) = 10.32$ **	ns.
Spelling accuracy (%)						
1st grade May	85.20	20.01	90.92	15.87	$F(1,181) = 4.48$ *	ns.
2nd grade June	62.83	27.05	73.09	20.96	$F(1,181) = 7.97$ **	*
3rd grade April	76.82	19.95	85.76	15.12	$F(1,185) = 11.60$ ***	**

[a] Note that standardized variables are reported here for RAN and phonological awareness composites. Standardization was based on the total low-risk sample's distribution at each measure at each time point. The mean values for the low-risk sample in this study are not exactly zero because all children were not included in the analyses of this paper because of missing data in the key variables. Note also that the spelling measure on grade 1 was different from the spelling measure used on grades 2 and 3..
*$p < .05$. **$p < .01$. ***$p < .001$.

Table 3

Cross-tabulation of DDH Group and Sample

Sample		DDH Group					
		Double Deficit	Phonological Deficit	Naming Deficit	Double Asset Non-Readers	Double Asset Readers	Total
High-risk	Count	16	17	18	29	21	101
	% within group	15.8 %	16.8 %	17.8 %	28.7 %	20.8 %	100.0 %
	% within cluster	66.7 %	63.0 %	75.0 %	44.6 %	44.7 %	54.0 %
	Adj. stand residual	1.3	1.0	2.2	-1.9	-1.5	
Low-risk	Count	8	10	6	36	26	86
	% within group	9.3 %	11.6 %	7.0 %	41.9 %	30.2 %	100.0 %
	% within cluster	33.3 %	37.0 %	25.0 %	55.4 %	55.3 %	46.0 %
	Adj. stand residual	-1.3	-1.0	-2.2	1.9	1.5	
	Total	24	27	24	65	47	187
		12.8 %	14.4 %	12.8 %	34.8 %	25.1 %	100.0 %

Note. χ^2 (4) = 10.63, p = .03.

Table 4

Phonological Awareness and RAN Development in the DDH Groups

	1. Double Deficit (n = 24)		2. Phonological Deficit (n = 27)		3. Naming Deficit (n = 24)		4. Double Asset Non-Readers (n = 65)		5. Double Asset Readers (n = 47)		*F*	Lev. sig.
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*		
RAN												
3.5 y [a]	$-1.00^{2,4,5}$	1.01	-0.10^{1}	0.82	$-0.70^{4,5}$	0.99	$0.24^{1,3}$	0.83	$0.27^{1,3}$	0.80	$F(4, 157)=11.21$***	ns.
5.5 y [a]	$-1.48^{2,4,5}$	1.08	$-0.09^{1,3}$	0.69	$-1.79^{2,4,5}$	1.79	$-0.07^{1,3}$	1.27	$0.43^{1,3}$	0.85	$F(4, 180)=20.91$***	***
6.5 y	$-2.25^{2,4,5}$	1.57	$-0.02^{1,3}$	0.41	$-2.20^{2,4,5}$	1.56	$0.28^{1,3}$	0.60	$0.52^{1,3}$	0.56	$F(4, 182)=68.53$***	***
Grade 1	$-1.10^{2,4,5}$	1.32	$-0.08^{1,3}$	0.72	$-0.90^{2,4,5}$	1.10	$0.17^{1,3}$	0.55	$0.41^{1,3}$	0.53	$F(4, 180)=22.32$***	***
Grade 3	$-1.32^{2,4,5}$	1.33	-0.12^{1}	0.85	$-0.71^{4,5}$	0.97	$0.18^{1,3}$	0.89	$0.32^{1,3}$	0.81	$F(4, 170)=15.13$***	ns.
PA												
3.5 y	$-0.95^{2,3,4,5}$	0.66	-0.32^{1}	0.71	0.05^{1}	0.53	-0.05^{1}	0.62	0.14^{1}	0.71	$F(4, 151)=10.48$***	ns.
5.5 y	$-1.28^{3,4,5}$	0.55	$-0.82^{4,5}$	0.64	-0.37^{1}	0.63	$-0.30^{1,2}$	0.57	$0.58^{1,2}$	0.67	$F(4, 162)=35.55$***	ns.
6.5 y	$-1.47^{3,4,5}$	0.49	$-1.34^{3,4,5}$	0.49	$-0.22^{1,2,5}$	0.40	$-0.11^{1,2,5}$	0.48	$0.84^{1,2,3,4}$	0.41	$F(4, 182)=150.50$***	ns.
Grade 1	$-1.22^{4,5}$	1.39	-0.55^{5}	1.22	-0.65^{5}	1.38	-0.13^{1}	0.96	$0.31^{1,2,3}$	0.93	$F(4, 177)=8.77$***	*
Grade 3	$-1.04^{4,5}$	0.86	$-0.75^{4,5}$	0.94	-0.66^{5}	0.91	$-0.06^{1,2}$	0.93	$0.43^{1,2,3}$	0.91	$F(4, 170)=13.95$***	ns.

Note. The superscript number indicate that group means differ significantly ($p \leq .05$), two-tailed ANOVA F tests. Reference of superscript numbers: 1 = Double Deficit, 2 = Phonological Deficit, 3 = Naming deficit, 4 = Double Asset Non-Readers, and 5 = Double Asset readers. The pairwise comparisons are based on Bonferroni correction. The F test results were confirmed with a robust test of Equality of means (Welch) when equality of variances was not obtained. There were no differences between F-test and Welch. Therefore Fs are reported.

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

Table 5

Descriptive Statistics and Group Comparisons of the DDH Groups in Reading and Spelling Development

	1. Double Deficit (n = 24)		2. Phonological Deficit (n = 27)		3. Naming Deficit (n = 24)		4. Double Asset Non-Readers (n = 65)		5. Double Asset Readers (n = 47)		*F*	Leve -ne sig.
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*		
Text-reading speed (words/minute)												
Grade 1	$20.08^{4,5}$	15.72	23.18^{5}	13.50	$21.74^{4,5}$	14.87	$34.86^{1,3,5}$	17.82	$57.65^{1,2,3,4}$	21.73	$F(4, 172)=28.17$***	**
Grade 2	$43.48^{4,5}$	20.76	52.57^{5}	21.03	$40.56^{4,5}$	18.22	$65.73^{1,3,5}$	22.02	$82.13^{1,2,3,4}$	19.70	$F(4, 178)=24.06$***	ns.
Grade 3	$54.33^{4,5}$	22.75	64.60^{5}	20.92	$48.04^{4,5}$	22.06	$78.70^{1,3,5}$	25.24	$92.04^{1,2,3,4}$	24.93	$F(4, 182)=19.44$***	ns.
Spelling accuracy (%)												
Grade 1	$72.47^{4,5}$	32.23	$78.63^{4,5}$	22.97	82.87^{5}	18.89	$92.36^{1,2}$	7.70	$96.57^{1,2,3}$	5.97	$F(4, 178)=11.80$***	***
Grade 2	$52.08^{4,5}$	30.22	$57.10^{4,5}$	24.21	$47.57^{4,5}$	27.53	$75.40^{1,2,3}$	17.10	$81.34^{1,2,3}$	15.53	$F(4, 178)=16.66$***	***
Grade 3	$67.71^{4,5}$	23.09	$73.15^{4,5}$	20.59	$71.18^{4,5}$	19.03	$86.28^{1,2,3}$	13.62	$89.72^{1,2,3}$	11.29	$F(4, 182)=12.54$***	***

Note. The superscript number indicate that group means differ significantly ($p \leq .05$), two-tailed ANOVA F tests. Reference of superscript numbers: 1 = Double Deficit, 2 = Phonological Deficit, 3 = Naming deficit, 4 = Double Asset Non-Readers, and 5 = Double Asset readers. The pairwise comparisons are based on Bonferroni correction. The F test results were confirmed with a robust test of Equality of means (Welch) when equality of variances was not obtained. There were no differences between F-test and Welch. Therefore Fs are reported.

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

Table 6

The F-values for the Repeated Measures of ANOVAs Predicting Reading Speed in Grade 1, Grade 2, and Grade 3, and Spelling Development in Grade 2 and Grade 3 (Only Data of Children Who Were Non-Readers at Age 6.5 Years (n=125) Are Included)

	Text-reading speed		Spelling Accuracy	
	Grade 1 à Grade 2 à Grade 3		Grade 2 à Grade 3	
Time	291.93	***	55.79	***
Familial Risk	3.86	*	11.69	***
RAN deficit	12.36	***	12.31	***
PA deficit	2.37		.55	
Familial Risk x RAN deficit	.10		2.84	
Familial Risk x PA deficit	.67		3.96	*
RAN deficit x PA deficit	2.36		6.15	*
Time x Familial Risk	.76		.01	
Time x RAN deficit	5.59	**	2.17	
Time x PA deficit	.15		.42	

Note. PA = phonological awareness, RAN = rapid serial naming.Note that all three-way and four-way interactions were non-significant and not reported here.
*$p < .05$. **$p < .01$. ***$p < .001$.

Table 7

Cross-tabulation of DDH Group and Dyslexia in Grade 2

	High-Risk Sample		Low-Risk Sample		Total Sample	
	No Dyslexia	Dyslexia	No Dyslexia	Dyslexia	No Dyslexia	Dyslexia
Double deficit	9 (56.3%)	7 (43.8%)	4 (50%)	4 (50%)	12 (52.2%)	11 (47.8%)
Phonological deficit	10 (58.8%)	7 (41.2%)	9 (90%)	1 (10%)	19 (70.4%)	8 (29.6%)
Naming deficit	7 (38.9 %)	11 (61.1%)	4 (66.7%)	2 (33.3%)	11 (45.8%)	13 (54.2%)
Double Asset Non-Readers	21 (72.4%)	8 (27.6%)	35 (97.2%)	1 (2.8%)	56 (86.2%)	9 (13.8 %)
Double Asset Readers	21 (100%)	0 (0%)	26 (100%)	0 (0%)	47 (100%)	0 (0%)
Total	68 (67.3%)	33 (32.7%)	78 (90.7%)	8 (9.3%)	145 (78%)	41 (22%)

Note. $\chi^2(4) = 40.05$, $p < .001$ for the total sample, $\chi^2(4) = 18.60$, $p < .001$ for the high-risk sample, and $\chi^2(4) = 24.30$, $p < .001$ for the low-risk sample.

Figure Captions

Figure 1. Scatter plots of RAN and phonological awareness at 6.5 years in the high-risk and low-risk sample. Values are standardized based on the low-risk sample's distribution (0 = low-risk sample average, 1= one standard deviation above low-risk sample average). DD = double deficit group, PD = phonological deficit group, ND = naming deficit group, DAnR = double asset non-reader group, and DAR = double asset reader group

Figure 2. Bar chart describing the interaction of phonological awareness (PA) deficit and rapid automatized naming (RAN) deficit in mixed method ANOVA for spelling

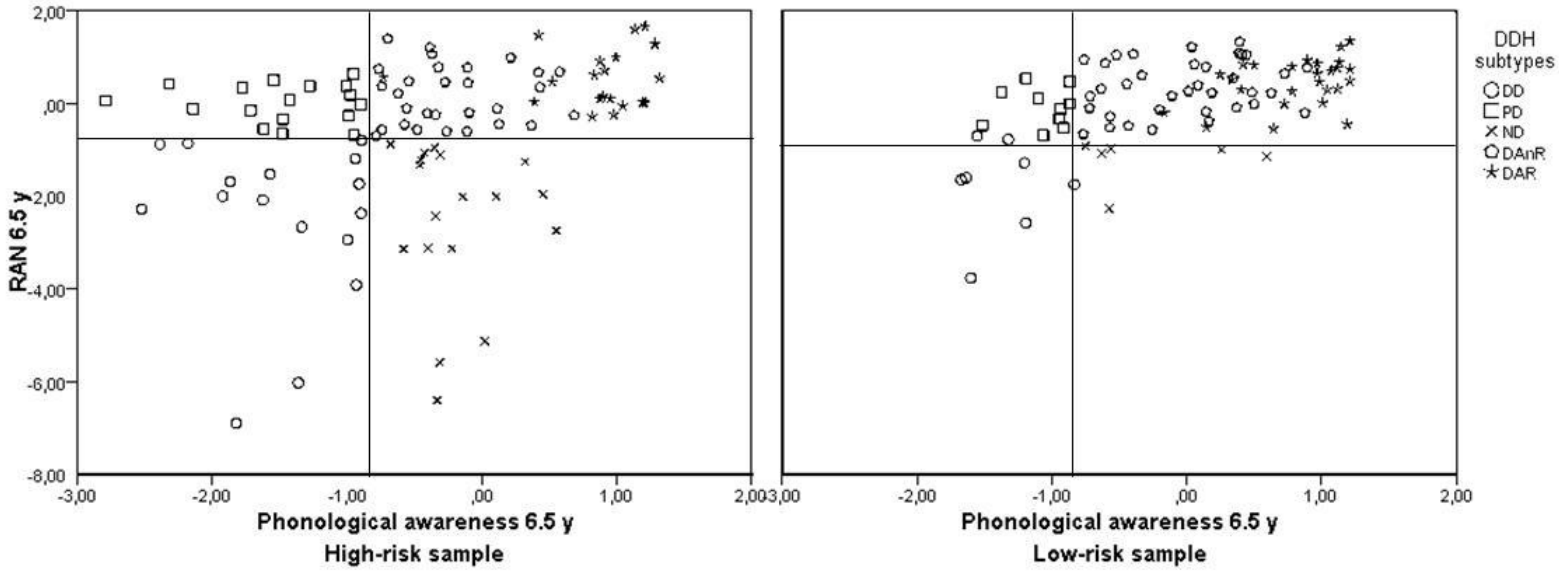
2,00
,00
-2,00
-4,00
-6,00
-8,00
RAN 6.5 y
-3,00
-2,00
-1,00
,00
1,00
2,00
Phonological awareness 6.5 y
High-risk sample
3,00
-2,00
-1,00
,00
1,00
2,00
Phonological awareness 6.5 y
Low-risk sample
DDH subtypes
DD
PD
ND
DAnR
DAR

Figure 1.

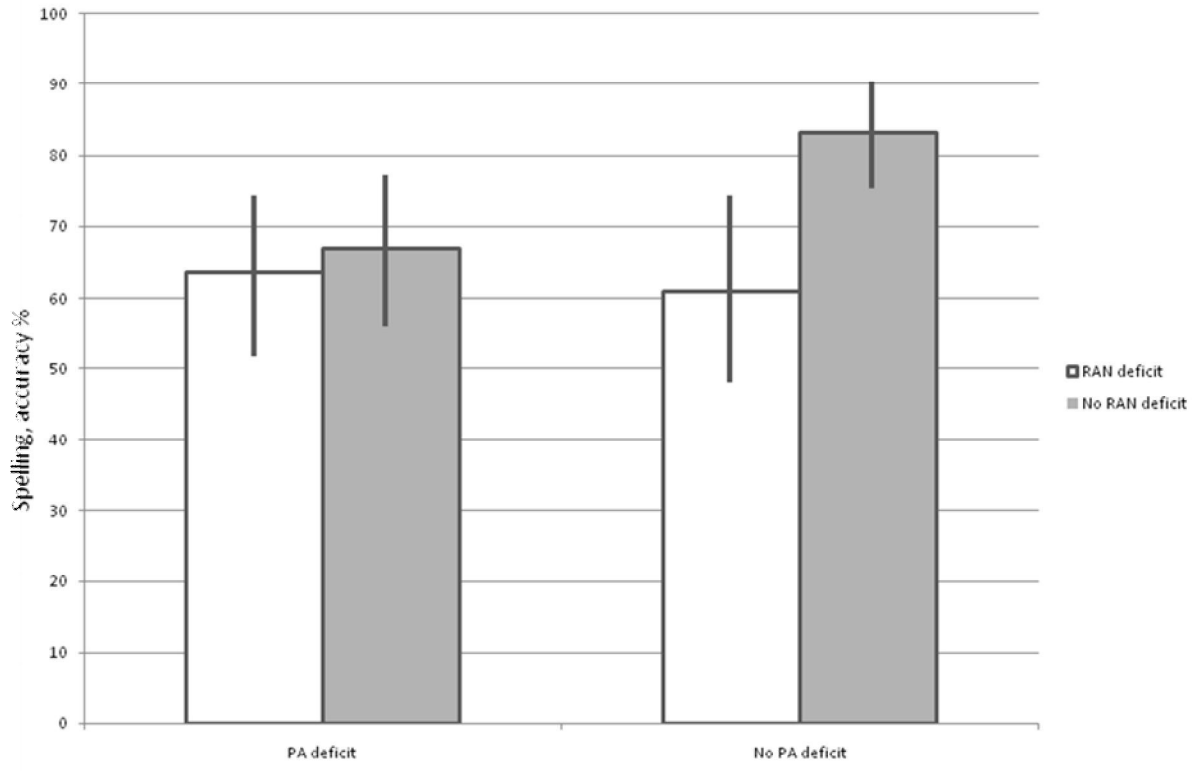
100
90
80
70
60
50
40
30
20
10
0
Spelling, accuracy %
PA deficit
No PA deficit
RAN deficit
No RAN deficit

Figure 2.