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

Author(s): Heikkilä, Riikka; Torppa, Minna; Aro, Mikko; Närhi, Vesa; Ahonen, Timo

Title: Double-Deficit Hypothesis in a Clinical Sample : Extension Beyond Reading

Year: 2016

Version:

Please cite the original version:

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.
The double-deficit hypothesis in a clinical sample: extension beyond reading.

Riikka Heikkilä, Minna Torppa, Mikko Aro, Vesa Närhi, and Timo Ahonen

This study explored the double-deficit hypothesis (DDH) in a transparent orthography (Finnish), and extended the view from reading disabilities to comorbidity of learning-related problems in math and attention. Children referred for evaluation of learning disabilities in second through sixth grade (N = 205) were divided into four groups based on rapid automatized naming (RAN) and phonological awareness (PA) according to the DDH: the double-deficit group, the naming speed deficit only group, the phonological deficit only group, and the no deficit group. The results supported the DDH in that the prevalence and severity of reading disability were greatest in the double-deficit group. Despite the greater prevalence of reading disabilities in single-deficit groups compared to the no deficit group, the means of reading measures in the single-deficit groups were similar to those of the no deficit group. The PA single-deficit group was poorer in spelling than the no deficit group and single naming deficit group. Deficits in RAN or PA were primarily linked to reading disabilities, but not with math or attention problems. The results supported the double-deficit hypothesis partially and implied that deficits in RAN and PA are specific to reading disabilities.
1 INTRODUCTION

1.1 The double-deficit hypothesis

The double-deficit hypothesis (DDH; Wolf & Bowers, 1999, 2000) was presented as an alternative for the phonological deficit hypothesis that has long been the predominant theory to explain the background of reading disabilities (RD; e.g., Stanovich & Siegel, 1994; Vellutino, Fletcher, Snowling, & Scanlon, 2004). DDH states that there are two partially independent core deficits behind RDs: deficits in phonological awareness (PA) and rapid automatized naming (RAN). Combinations of these deficits lead to three subgroups: double deficit (DD, deficits in RAN and PA), phonological deficit (PD) only, and naming speed deficit (NSD) only. It is assumed that the negative effects of PD and NSD are additive, and thus, RDs are more severe in the DD group than in single-deficit groups (Wolf & Bowers, 1999). An additional linked hypothesis is that PA and RAN are associated with different aspects of literacy: PA has been primarily connected with reading and spelling accuracy, and RAN has been especially linked with reading fluency and rate (Compton, DeFries, & Olson, 2001; Cornwall, 1992; Furnes & Samuelsson, 2011; Kairaluoma, Torppa, Westerholm, Ahonen, & Aro, 2013; Moll, Ramus, et al., 2014; Papadopoulos, Georgiou, & Kendeou, 2009; Pennington, Cardoso-Martins, Green, & Lefly, 2001; Schatschneider, Carlson, Francis, Foorman, & Fletcher, 2002; Sunseth & Bowers, 2002; Torppa, Georgiou, Salmi, Eklund & Lyytinen, 2012; Vaessen, Gerretsen, & Blomert, 2009; Vaessen et al., 2010; Wimmer, Mayringer, & Landerl, 2000, however, see Moll, Ramus, et al., 2014, and Ziegler et al., 2010 for contradictory findings on RAN’s role concerning Finnish samples).

The double-deficit hypothesis serves as a plausible model for explaining reading problems, especially in more transparent orthographies. Due to the fast development of reading accuracy (Seymour, Aro, & Erskine, 2003), PA’s role in predicting reading skills seems to be limited in the first school grades in these orthographies (Aarnoutse, van Leeuwe, & Verhoeven, 2005; Papadopoulos et al.,
Torppa et al., 2012; Wimmer et al., 2000) whereas RAN as a predictor of reading rate and fluency seems to be a more robust predictor of reading after early grades than PA (see e.g. Kirby, Georgiou, Martinussen, & Parrila, 2010; Norton & Wolf, 2012; Wolf, Bowers, & Biddle, 2000, for reviews). That said, DDH has received support across orthographies in studies of school-aged children in unselected samples as well as among poor readers (see Georgiou & Parrila, 2013; Kirby et al., 2010; Norton & Wolf, 2012, and Wolf et al., 2000 for reviews, and Cronin, 2013; Norton et al., 2014; Steacy, Kirby, Parrila, & Compton, 2014; Torppa et al., 2012; Torppa et al., 2013, for latest results). Some studies, however, have not found support for the independent role of RAN in explaining RDs (for a review, see Vukovic & Siegel, 2006). One reason for the lack of a RAN-RD link may be the use of RD definitions that are based solely on accuracy leaving the variation between RAN and reading fluency uncontrolled. There are also contradictory findings on the role of single-deficit subgroups based on DDH: even though a group with double-deficit has often been the poorest in most literacy skills, not all studies have found significant differences between single-deficit groups in reading skills (e.g. Manis, Doi, Bhadha, 2000; Cronin, 2013; Norton et al., 2014, see also Wolf & Bowers, 1999 for review).

1.2 DDH and comorbid problems with RDs: math disabilities and attention deficit

In addition to RDs, deficits in phonological processing and RAN may influence comorbid problems such as math disabilities (MDs) and attention deficit (AD). RD is often comorbid with MD and AD (Czamara et al., 2013; DuPaul, Gormley, & Laracy, 2013; Landerl & Moll, 2010; see Boada, Willcutt, & Pennington, 2012; Germanò, Gagliano, & Curatolo, 2010; Sexton, Gelhorn, Bell, & Classi, 2012, and Willcutt, Pennington, et al., 2010 for reviews). Most studies, however, have studied RDs, MDs, and AD separately, or subjects with comorbid learning problems have been excluded from analyses. Even though the etiology of these disabilities seem to partly overlap (Hart, Petrill, Thompson, & Plomin, 2010; Kovas et al., 2007; Willcutt, Pennington, et al., 2010), consensus on the cognitive or behavioral factors associated with the comorbidity is lacking (see e.g. Miranda, Presentación, Siegenthaler, Colomer, & Pinto, 2011, for a discussion on the comorbidity between RD and AD). In addition, it has been suggested that the comorbidity of learning disabilities and attention deficit may be related with more severe neuropsychological deficits than single disabilities (Boada et al., 2012; Miranda et al., 2011). Therefore, attempts to discover the specific underlying deficits in each disability as well as the common factors associated with the comorbidity are important in detecting and diagnosing these disabilities and preventing the secondary causes of comorbid disabilities. Essential from the perspective of double-deficit hypothesis is the possibility that
deficits in phonological processing or RAN are linked not only with RDs but also with other challenges in learning, increasing the comorbidity of learning disabilities and AD. Although naming speed (Denckla & Rudel, 1976; Heikkilä, Närhi, Aro, & Ahonen, 2009; Willcutt, Betjemann, et al., 2010) and phonological deficits (Fletcher, 2005; Landerl, Fussenger, Moll, & Willburger, 2009) seem to be most strongly connected with RDs, links with MDs and AD have been found (see below).

Evidence for the effects of RAN and phonological skills on math is mixed. Some studies have shown a unique connection between RAN and math skills (Ackerman, Holloway, Youngdahl, & Dykman, 2001; Koponen, Aunola, Ahonen, & Nurmi, 2007; Koponen, Salmi, Eklund, & Aro, 2013; Murphy, Mazzeocco, Hanich, & Early, 2007; van Bergen, de Jong, Maassen, & van der Leij, 2014; van der Sluis, de Jong, & van der Leij, 2004). A unique effect of phonological skills on math skills has also been reported (De Smedt & Boets, 2010; Krajewski & Schneider, 2009; Simmons & Singleton, 2008), even after controlling for reading skills (Hecht, Torgesen, Wagner, & Rashotte, 2001). It has also been suggested that a major part of the shared variance between reading and math performance is explained by RAN among children with RDs and MDs (Geary, 1993; Geary, Hamson, & Hoard, 2000), or phonological processing and RAN (Hecht et al., 2001). However, not all studies have supported the unique link between RAN and math. In some studies, the effect of RAN on math performance disappeared after other relevant cognitive variables (such as reading, IQ, attention, or processing speed) were controlled (Georgiou, Tziraki, Manolitsis, & Fella, 2013; Landerl et al., 2009; Moll, Göbel, Gooch, Landerl, & Snowling, 2014; Willcutt et al., 2013; Willburger, Fussenger, Moll, Wood, & Landerl, 2008). Similarly, not all studies have supported a unique connection between PA and math skills when reading performance has been controlled (Durand, Hulme, Larkin, & Snowling, 2005; van Bergen et al., 2014; Willcutt et al., 2013).

In regard to attention deficit, phonological skills and AD seem not to be related (Gooch, Snowling, & Hulme, 2011; McGee, Brodeur, Symons, Andrade, & Fahie, 2004; Purvis & Tannock, 2000; Willcutt et al., 2001). Instead, a link between naming speed deficit and attention deficit has been suggested (Tannock, Martinussen, & Frijters, 2000) especially with AD without impulsivity and hyperactivity (Arnett et al., 2012; Hynd et al., 1991; Thomson et al., 2005). Some studies suggest that naming speed deficits are most severe in the comorbid group with RDs and AD (Bental & Tirosh, 2007; Rucklidge & Tannock, 2002). However, in many studies that controlled for reading, a connection between RAN and attention was not evident (Ackerman & Dykman, 1993; Felton & Wood, 1989; Raberger & Wimmer, 2003; Semrud-Clark, Griffin, & Hynd, 2000) or the attention deficit hyperactivity disorder (ADHD) group outperformed the RD group in RAN (Felton, Wood, Brown, Campbell, & Harter, 1987; Närhi & Ahonen, 1995; Willcutt, Betjemann, et al., 2010).
To our knowledge, only three studies have investigated the double-deficit hypothesis and included measures for attention and/or math. First, Ackerman and colleagues (2001) explored double deficits in children with RDs. The researchers found RAN had only a minor effect on RDs after PA was controlled. The researchers also showed that comorbid problems in mathematics were common in a sample with RDs and that RAN was significantly correlated with math performance, attention ratings, and processing speed. However, reading was measured based on accuracy only. Second, Waber, Forbes, Wolff, and Weiler (2004) investigated whether groups based on the double-deficit hypothesis showed differences in other skills than reading. A group for single phonological deficit was not found, but the three groups (DD, NSD, and NoD) differed in several skills, such as motor, visuospatial, or oral language skills. However, there were no group differences in attention. It was suggested that because RAN and PA seemed to be linked to skills other than reading, it was probable that learning problems other than reading were accumulated in the double-deficit group. However, in Waber and colleagues’ (2004) study, RD was defined based solely on word reading accuracy, and measures for math skills were not included in the study. In a third study (Torppa et al., 2013), an unselected sample of Finnish children was explored in the framework of DDH including measures for attention and using reading measures based on fluency. The results showed more attention deficits in the double-deficit group indicating that they were more prone to comorbid problems.

1.3 The present study

In this study, we examine the DDH in Finnish, which is at the extreme transparent end of the transparency continuum of orthographies (see Seymour et al., 2003). Our sample consisted of children who all have learning difficulties, and the majority of the children have RDs. In a clinical sample, measures that typically reach ceiling in transparent orthographies (i.e., reading accuracy and PA) may still show variance, especially with tasks complex enough (Caravolas, Volin, & Hulme, 2005; De Jong & van der Leij, 2003; Landerl, Wimmer, & Frith, 1997; Treinta, Benassi, Di Filippo, Pontillo, & Zoccolotti, 2013). Although the DDH has been shown to predict RD in Finnish (Torppa et al., 2012, 2013), we extended the focus from RDs to comorbidity with mathematical disabilities and attention deficit. The use of a clinical sample ensures that the sample includes children with severe learning difficulties. To further broaden the scope of previous studies, we include accuracy and rate aspects in reading assessment, which is a more relevant approach in transparent orthographies. Majority of the previous studies investigating the connection of RAN and PA with math or attention performance have either not controlled the effect of reading at all or used accuracy only. Thus, the connections found may be mediated by reading fluency.
The research questions of this study are as follows: (a) Will the assumptions of the double-deficit hypothesis be replicated in a clinical sample with comorbid deficits in reading, math, and attention? If RAN and PA are independent and additive contributors behind RD, it is expected that the correlation between RAN and PA is modest, RAN and PA differ from each other in their relation to reading and spelling performance, the groups based on the double-deficit hypothesis (DDH groups) differ in reading and spelling abilities, and the reading disabilities are most severe and the prevalence of RDs is greatest in the DD group. (b) Do the DDH groups differ in math performance or attention ratings? Group differences in these measures would indicate that deficits in PA and/or RAN were not restricted only to reading and spelling domains. (c) Is the comorbidity of RD, MD, and AD linked to the prevalence of PD, NSD, and DD? If RAN and PA have a unique link with MDs and AD excluding reading, the frequency of PD, NSD, and DD should be high in comorbid groups with RDs, MDs, and AD. However, if the connection of NSD and PD with MDs and AD is mediated by reading ability, the prevalence of NSD, PD, and DD will be higher only in the groups with RDs.
2 METHODS

2.1 Participants

The clinical database used in this study consisted of children referred to a child neuropsychological clinic for closer evaluation of learning disabilities or of the need for neuropsychological intervention. Children were referred by a psychologist or special health care services after basic examination of cognitive skills and learning disabilities conducted by a psychologist. All children had an acknowledged specific learning disability and all of them had received special education services in school: 7% of children were situated in a small group and 25% of children had Individualized Education Program in some of the school subjects. 20% of the children had been retained in school. Some children used additional services like speech therapy or occupational therapy. After the evaluation at child neuropsychological clinic the proper support for learning and adjustments in school were suggested, including support provided in classroom, special education services, and neuropsychological intervention. After applying the exclusion criteria of native language other than Finnish, score below 80 in both verbal and performance IQ on the Wechsler Intelligence Scale for Children (WISC-R or WISC-III; Wechsler, 1974, 1991), or a child’s neurological disorder reported by parents, the resulting sample consisted of 205 children. All children were affected by single or comorbid disabilities in reading (RDs, \( n = 158 \)), math (MDs, \( n = 96 \)), and/or attention (AD, \( n = 78 \)) (see the criteria for definitions, and see Figure 1 for the comorbidity of RDs, MDs, and AD where areas represent the proportion of children in each sector). All children came from central Finland, and the distribution of the educational level of their parents was comparable to the Finnish population. The mean age of the sample was 10 years 4 months (standard deviation 13 months, range = 8 years 0 months to 13 years 8 months), and the proportion of boys was 62%. Information on the general cognitive ability, phonological skills, RAN, academic performance, and evaluation of the attention for the sample is presented in Table 1.
FIGURE 1 Combinations and comorbidity of reading disabilities (RD), math disabilities (MD), and attention deficit (AD).

2.2 Measures

All measures used in this study were obtained from the neuropsychological assessment of the children, with the exception of attention, which was evaluated by the teachers and parents of the child. The typical assessment procedure took place over the course of two 3-hour sessions. Measures with available normative data (IQ, RAN, PA, reading fluency, and spelling) were standardized according to the age level, and z-scores were used in the analyses.

IQ. The total score on the Wechsler Intelligence Scale for Children (WISC-R or WISC-III; Wechsler, 1974, 1991) was used as a measure for intelligence.

Rapid Automatized Naming. RAN was assessed using letters (O, A, S, T, P) and digits (2, 4, 6, 7, 9) from a Finnish version of the test of rapid automatized naming (Ahonen, Tuovinen, & Leppäsaari, 1999). The stimulus cards consisted of a total of 50 items arranged in five rows. Each of the five items was repeated in pseudorandom order; none of the items were presented successively. The mean of the letter and number naming times ($r = .789$) was standardized according to age and used as the RAN score. Cronbach’s alpha for the letter and
digit naming rate based on standardized items in the sample used in this study was $\alpha = .882$ (cp. $\alpha = .892$ in normative sample).

**Phonological awareness.** For phonological awareness, we used the Phonological Processing subtask of the Developmental Neuropsychological Assessment (NEPSY; Korkman, Kirk, & Kemp, 1997). It consists of tasks for phoneme or word segment deletion and replacement (maximum score of 36). For children under 9 years of age, the task begins with a subtask in which child must identify one of the three pictures that contains the word segment the experimenter says. The standardized score of correct answers was used as the PA score. Cronbach alpha for this task is reported to be .97 in a normative sample.

**Reading fluency.** Reading fluency (i.e., a combined measure for reading accuracy and rate) was defined with a standardized time-limited word list reading (Häyrinen, Serenius-Sirve, & Korkman, 1999). Children were instructed to read the words aloud as accurately and quickly as possible. The standardized score from number of words read correctly within 2 minutes was used as an outcome score.

**Reading accuracy and rate.** In addition to the measure for reading fluency, we included separate accuracy and rate measures. In the absence of normative data for the separate accuracy and rate measures, age was controlled in the statistical analysis of these measures. The separate composites for reading accuracy and rate were formed of three measures: word list reading, pseudoword list reading, and age appropriate informational text reading (Niilo Mäki Institute, 2004). The word and pseudoword lists consisted of 20 items with a varying length (from 4 to 18 letters), including multimorphemic words and structures requiring decoding of phonemic length (e.g., words like veneeseen and lannistumatonta, and pseudowords like tapekkaat and kaalluspastikki), which are problematic for Finnish dyslexic readers (Pennala et al., 2010) and adults with poor reading skills (Lyytinen, Leinonen, Nikula, Aro, & Leiwo, 1995). In all tasks, the children were instructed to read aloud as quickly and accurately as possible from the beginning to the end. Reading accuracy was defined as the mean percentage of accurately read items of these three measures. Cronbach’s alpha for the accuracy score based on the three standardized items was $\alpha = .787$ and the correlations between accuracy measures controlled for age varied between .51 and .55. The accuracy for text reading was high ($M = 93\%$, $SD = 6$; 66% of the sample exceeding 90% accuracy), whereas for words and pseudowords, there was more variation ($M = 82\%$, $SD = 15$, $M = 70\%$, $SD = 18$, respectively). For the reading rate, the completion time for word list reading, pseudoword list reading, and text reading was measured. For each task, the number of items read per minute was calculated, and the mean rate of the three tasks served as a measure for the reading rate (Cronbach alpha based on standardized items, $\alpha = .899$, correlations controlled for age varied between .69 and .84).

**Spelling accuracy.** Spelling accuracy was defined by a standardized dictation task with 20 items with increasing difficulty (Häyrinen et al., 1999). For second graders, the items were words, and for older children, the items consist-
ed of words and phrases. The standardized score of words spelled correctly was used as a score for spelling.

**Math skills.** Math skills were assessed with a time-limited test of arithmetic fluency containing multidigit calculations from all four basic operations, RMAT (Räsänen, 2004). The Cronbach alpha for the RMAT is reported to be .92–.95 in a normative sample, depending on the school grade. If the RMAT score was not available, a subtest for arithmetic from another standardized test battery (Häyrinen et al., 1999) was used for the MD definition \((n = 10)\). For analyses of variance, a raw RMAT score was used, with age controlled.

**Attention.** Attention problems were evaluated with the scale for Attention-deficit/hyperactivity problems in the Child Behavior Checklist (CBCL) filled in by the child’s parents or the Child Behavior Checklist Teacher’s Report Form (TRF) filled in by the child’s primary school teacher (Achenbach & Rescorla, 2001). For the attention scale, the CBCL has seven and the TRF 13 statements, which are rated as three categories \((0 = \text{not true}, 1 = \text{somewhat or sometimes true}, \text{and} 2 = \text{very true or often true})\). In the absence of published local norms, the raw scores were changed to \(t\) scores based on U.S. norms (Achenbach & Rescorla, 2001), and the mean score of the available assessments was used as in the analyses (Cronbach alpha based on standardized items, \(\alpha = .754\)).

### 2.2.1 Criteria for DDH, RD, MD, and AD categorization.

The criterion for naming speed deficit was a RAN performance 1.0 SD below the normative mean for age. For phonological awareness, a criterion for phonological deficit was a mean accuracy score 1.0 SD below the normative mean for age. Based on these criteria, four groups according to the DDH were formed: DD, NSD, PD, and NoD.

The criterion used for classifying RD and MD was the performance level 1.0 SD below the normative mean for the grade. The criteria were based on an efficiency score, that is, the number of correct responses within a time limit. A cutoff of \(t\) score greater than 60 (i.e., 16th percentile in U.S. norms; Achenbach & Rescorla, 2001) on the Attention scale on either the parental or teacher evaluation in the Child Behavior Checklist served as the criterion for AD.

### 2.2.2 Analysis Overview

Outliers that deviated from the distribution (more than 2.5 SD below the mean of the sample) in text reading accuracy \((n = 4)\), phonological awareness \((n = 2)\), and RAN \((n = 6)\) were moved to the tail of the distribution in their original order to avoid bias in the analyses. The distributions of the RAN, PA, reading, and spelling measures were skewed (with an exception for reading fluency measure, which did not need any adjustments) and were normalized with transformations: RAN with natural logarithm and PA, reading, and spelling with square root transformation. Measure for attention was skewed (floor effect) in spite of the attempts with square root or natural logarithmic transformation.
The descriptive statistics for the measures before transformations are reported in Table 1.

The between-group analyses were performed with analyses of variance (ANOVA). First, the performance in RAN and PA (dependent variables) in the DDH groups (fixed factors) were compared to verify the grouping based on double-deficit hypothesis and to address the concern that possibly lower performance in the double-deficit group may be due to the more severe PD and NSD in this group compared to single-deficit groups (Schtschneider et al., 2002).

In the second set of analyses, the effects of phonological deficit and naming speed deficit on reading, spelling, and math skills were examined with ANCOVA, with age set as the covariate. First, phonological deficit (deficit, no deficit) and naming speed deficit (deficit, no deficit) were included as separate fixed factors in order to examine their main effects and interaction effect. Second, pairwise comparisons of the four DDH subgroups (DD, NSD, PD, NoD) were conducted. Since the attention measure was skewed and not possible to normalize with the transformations, the between-group analyses including attention were conducted with the Kruskal-Wallis test.

Finally, the prevalence and comorbidity of RDs, MDs, and AD were computed in the DDH groups. Groups for different combinations of RDs, MDs, and AD (LD groups) were cross-tabulated with the DDH groups.
3 RESULTS

3.1 Comparison of age, gender, IQ, phonological awareness, and RAN in the DDH groups

Age and IQ in the DDH groups (DD, \( n = 90 \); NSD, \( n = 48 \); PD, \( n = 37 \); NoD, \( n = 30 \)) were compared with ANOVA analyses, and gender distribution was explored with cross-tabulation (Gender \( \times \) DDH Groups). The DDH groups did not differ from each other in any of these measures (see Table 1). Results from the group comparisons in PA and RAN showed that the DD group did not have more severe deficits in PA and RAN than the single-deficit groups. Accordingly, the single-deficit groups were well-matched with the NoD group since there was no difference between the NSD group and the NoD group on PA, and there was no difference between the PD group and the NoD group on RAN (Table 1).

The correlations (see Table 2) confirmed the group comparison results. The correlation between PA and RAN was not significant. PA was most strongly related to reading and spelling accuracy and RAN to reading fluency and rate.

3.2 Comparison of reading and spelling in the DDH groups

Group comparisons of reading fluency (ANCOVA, age as a covariate) revealed that NSD status (deficit, no deficit) had the main effect on reading fluency, \( F(1, 164) = 5.41, p = .021, \eta^2 = .032 \). Neither the main effect of phonological deficit status (deficit, no deficit) nor the interaction between NSD and PD status was significant. The pairwise comparisons of the four DDH groups in reading fluency revealed that the DD group performed significantly worse than all the other groups.
<table>
<thead>
<tr>
<th></th>
<th>Double-deficit</th>
<th>Phonological deficit (n = 37)</th>
<th>Naming speed deficit (n = 48)</th>
<th>No deficit (n = 30)</th>
<th>F Value</th>
<th>df</th>
<th>η²</th>
</tr>
</thead>
<tbody>
<tr>
<td>% of Girls</td>
<td>40.0</td>
<td>35.1</td>
<td>41.7</td>
<td>33.3</td>
<td>n. s.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age (months)</td>
<td>119.4</td>
<td>121.1 (88.4)</td>
<td>122.5 (91.3)</td>
<td>119.2</td>
<td>n. s.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IQ</td>
<td>90.6</td>
<td>88.4</td>
<td>91.3</td>
<td>89.2</td>
<td>n. s.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PA (z score)</td>
<td>-2.7&lt;sup&gt;a&lt;/sup&gt;</td>
<td>-2.3 (0.8)</td>
<td>-0.5 (0.5)</td>
<td>-0.4 (0.6)</td>
<td>122.37***</td>
<td>3, 201</td>
<td>.646</td>
</tr>
<tr>
<td>RAN (z score)</td>
<td>-2.9&lt;sup&gt;a&lt;/sup&gt;</td>
<td>-0.2 (0.5)</td>
<td>-2.5 (1.9)</td>
<td>-0.4 (0.5)</td>
<td>86.00***</td>
<td>3, 201</td>
<td>.562</td>
</tr>
<tr>
<td>Reading</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fluency (z score)</td>
<td>-2.1&lt;sup&gt;a&lt;/sup&gt;</td>
<td>-1.5 (1.0)</td>
<td>-1.6 (1.2)</td>
<td>-1.3 (1.3)</td>
<td>4.51**</td>
<td>3, 164</td>
<td>.076</td>
</tr>
<tr>
<td>Accuracy (%)</td>
<td>77.7&lt;sup&gt;a&lt;/sup&gt;</td>
<td>83.6 (9.7)</td>
<td>83.7 (10.2)</td>
<td>85.5 (6.9)</td>
<td>4.86**</td>
<td>3, 184</td>
<td>.073</td>
</tr>
<tr>
<td>Rate (items/min)</td>
<td>24.5&lt;sup&gt;a&lt;/sup&gt;</td>
<td>32.8 (11.7)</td>
<td>29.4 (11.7)</td>
<td>31.6 (12.7)</td>
<td>6.22***</td>
<td>3, 184</td>
<td>.092</td>
</tr>
<tr>
<td>Spelling (z score)</td>
<td>-1.7&lt;sup&gt;a&lt;/sup&gt;</td>
<td>-1.7 (1.2)</td>
<td>-1.0 (1.4)</td>
<td>-0.6 (0.8)</td>
<td>4.93**</td>
<td>3, 151</td>
<td>.089</td>
</tr>
<tr>
<td>Math (raw score)</td>
<td>19.7</td>
<td>21.5</td>
<td>21.7</td>
<td>21.8</td>
<td>n. s.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Attention (t score)</td>
<td>55.3</td>
<td>56.3</td>
<td>55.6</td>
<td>58.4</td>
<td>n. s.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### TABLE 2 Correlations Among Tasks Controlling for Age

<table>
<thead>
<tr>
<th>Variable</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. PA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. RAN</td>
<td>.14</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Reading Fluency</td>
<td>.18*</td>
<td>.30***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Reading Accuracy</td>
<td>.19**</td>
<td>.16*</td>
<td>.42***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Reading Rate</td>
<td>.12</td>
<td>.38***</td>
<td>.58***</td>
<td>.32***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Spelling</td>
<td>.36***</td>
<td>.01</td>
<td>.32***</td>
<td>.44***</td>
<td>.10</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. Math</td>
<td>.18*</td>
<td>.09</td>
<td>.04</td>
<td>.12</td>
<td>.04</td>
<td>.04</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. Attention&lt;sup&gt;a&lt;/sup&gt;</td>
<td>.04</td>
<td>.12</td>
<td>.20**</td>
<td>.01</td>
<td>.08</td>
<td>.03</td>
<td>.05</td>
<td></td>
</tr>
</tbody>
</table>

*Note.* Variables are coded by numbers in table head. PA = phonological awareness, RAN = rapid automatized naming. *Non-parametric correlations (Spearman)*

*<sup>*p </sup> < .05, **<sup>p</sup> < .01, ***<sup>p</sup> < .001.

Group comparisons of reading accuracy (ANCOVA, age as a covariate) revealed that NSD status, \(F(1, 184) = 4.82, p = .029, \eta^2 = .026\), and PD status, \(F(1, 184) = 3.95, p = .048, \eta^2 = .021\) had the main effect, but the NSD × PD interaction was not significant. The pairwise comparisons for the four DDH groups revealed that the DD group performed significantly worse in reading accuracy than all the other groups. No other group comparisons were significant.

Group comparisons of the reading rate (ANCOVA, age as a covariate) revealed that NSD status, \(F(1, 184) = 10.62, p = .001, \eta^2 = .055\), had the main effect. Neither the main effect for PD status nor the NSD × PD interaction was significant. Pairwise comparisons for the four DDH groups revealed that there were significant group differences between the DD group and all other groups in reading rate. No other group comparisons were significant.

Group comparisons of spelling (ANCOVA, age as a covariate) revealed that PD status had a significant main effect, \(F(1, 151) = 13.43, p < .001, \eta^2 = .082\). NSD status had no main effect, and there was no NSD × PD interaction. Pairwise comparisons showed that the DD and PD groups performed worse than the NSD and NoD groups in spelling. To insure that the results found in analyses of variance were not reached due to the alterations made for the data (transformations, covarying the age), the data was reanalyzed with untransformed data and with nonparametric methods where normal distribution was not required nor age controlled. The results of these analyses paralleled the ones reported here.

### 3.3 Prevalence of RDs in the DDH groups

To explore the prevalence of RDs in the DDH groups, cross-tabulation with the DDH groups (DD, NSD, PD, NoD) and RD status (RD, no RD) was conducted. The cross-tabulation revealed a significantly greater prevalence of RDs in the DD group (90%, adjusted standardized residual = 3.9) and significantly smaller in the NoD group.
(53%, adjusted standardized residual = −3.3) than expected, \( \chi^2(3) = 19.49, p < .001 \). There was comparable prevalence of RDs in the single-deficit groups (73% in the PD group and 71% in the NSD group). The prevalence of RDs in the NoD group (53%), however, was remarkably greater than expected in the normative group (17%, equivalent of the 1.0 SD cutoff from the normative mean, which was a criterion for RDs in this study), a result that was predictable in the clinical sample.

### 3.4 Comparison math skills and attention in the DDH groups

To examine whether the DDH group differences were not restricted only to reading and spelling skills, the DDH groups were compared regarding their math performance and attention ratings. Group comparisons of math performance (ANCOVA, age as a covariate) revealed no significant main effect for NSD or PD status, and no significant PD status × NSD status interaction. Similarly, the group comparisons for attention (Kruskal-Wallis test) revealed no differences between the DDH groups in attention.

### 3.5 Prevalence of MDs and AD in the DDH groups

To explore the prevalence of MDs and AD in the DDH groups, cross-tabulations with the DDH groups (DD, NSD, PD, NoD) across MD status (MD, no MD), and with the DDH groups across AD status (AD, no AD) were conducted. The results revealed an equal distribution across the sample in both analyses; thus, AD and MDs were not related to the DDH grouping.

### 3.6 DDH groups and comorbidity of RDs, MD, and AD

To explore the link of comorbidity of RDs, MDs, and AD to the DDH, groups presenting different kind of combinations of reading, math, and attention problems were formed (comorbidity groups). As can be seen in Figure 1, the comorbidity of RDs, MDs, and AD was high overall. Roughly 50% of the children in this sample fulfilled the criteria for more than one disability.

Cross-tabulation for the DDH groups and the comorbidity groups revealed an uneven cell distribution (see Table 3) indicating that double deficits were less pronounced in groups without RD (13–25%) than in groups with RD (43–55%). Thus, the results from the comorbidity groups as well as pairwise correlations between measures suggest that DD seems to be related especially to RDs but not to MDs or AD.
<table>
<thead>
<tr>
<th>LD group</th>
<th>Double-deficit</th>
<th>Phonological deficit</th>
<th>Naming speed deficit</th>
<th>No deficit</th>
<th>$\chi^2$</th>
<th>$\Phi$</th>
</tr>
</thead>
<tbody>
<tr>
<td>RD+MD+AD</td>
<td>10 (0.6)</td>
<td>4 (0.2)</td>
<td>5 (0.2)</td>
<td>1 (-1.3)</td>
<td>29.7*</td>
<td>0.38*</td>
</tr>
<tr>
<td>RD+MD</td>
<td>25 (1.8)</td>
<td>7 (-0.5)</td>
<td>9 (-0.6)</td>
<td>4 (-1.2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RD+AD</td>
<td>13 (-0.1)</td>
<td>5 (-0.2)</td>
<td>7 (0.0)</td>
<td>5 (0.3)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RD</td>
<td>33 (1.6)</td>
<td>11 (-0.1)</td>
<td>13 (-0.6)</td>
<td>6 (-1.4)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MD+AD</td>
<td>3 (-1.4)</td>
<td>1 (-0.9)</td>
<td>4 (0.8)</td>
<td>4 (1.9)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MD</td>
<td>4 (-2.1)</td>
<td>5 (1.0)</td>
<td>7 (1.5)</td>
<td>3 (0.1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AD</td>
<td>2 (-2.6)</td>
<td>4 (0.8)</td>
<td>3 (-0.5)</td>
<td>7 (3.4)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* = $p < .05$. RD = reading disability, MD = math disability, AD = attention deficit. Adjusted standardized residuals appear in parentheses below group frequencies.

**Note.**
4 DISCUSSION

In this study, the double-deficit hypothesis (DDH) was examined in a Finnish clinical sample of children with various combinations of developmental difficulties in reading, math, and/or attention. The specificity of PA and RAN to reading was examined by analyzing whether difficulties in PA and RAN were linked to RDs, MDs, or AD, or to comorbid difficulties. Our findings partly supported the DDH and gave further support for specificity of the connection between RAN, PA, and reading.

We found partial support for the double-deficit hypothesis (Wolf & Bowers, 1999) as the DD group with deficits in PA and RAN performed significantly poorer in reading accuracy, rate, and fluency relative to all the other groups. However, in contrast to the expectation that the single-deficit groups would show different manifestations of reading deficits (NSD poor in reading fluency and rate and PD group in reading accuracy), single-deficit groups did not differ from each other nor from no-deficit group in reading skills. The only difference between single-deficit groups was the poorer performance of PD group than NSD or NoD group in spelling. The result that reading skills were similar in the single-deficit and NoD groups indicates either that the combination of NSD and PD had the most dramatic negative effects on reading performance whereas the negative effect of single deficits may to some extent be compensated or that in clinical samples there are other additive predictors over PA and RAN that hinder the reading performance. The finding for the lowest performance in DD group was not due to the poorer performance in PA and RAN compared to single-deficit groups in this sample, a concern addressed by Schatschneider et al. (2002). In accordance with our hypothesis, the prevalence of RDs was higher in the DD group (90%) compared to the single-deficit groups (both around 70%) and the NoD group (53%). In the present sample of children who were referred to clinic due to their difficulties in learning, the prevalence of RDs was significantly higher even in the NoD group (53%) than expected in the normative sample with 1.0 SD cutoff (17%). It follows that even though double deficits unquestionably had the closest link to RDs, there were also children who had RDs without PD or NSD (10% of the children with RDs in this study, n = 16). Because a proportionally large number of children in this sample had poor reading skills regardless of the children’s position in the DDH grouping, multiple problems (e.g., cognitive, socioemotional, mo-
tivational, mental) that are manifest in clinical samples might cause problems in literacy skills over and above PA and RAN in all groups.

PA and RAN had unique connections with reading and spelling skills so that PD status (deficit, no deficit) had a significant main effect on reading accuracy and spelling but not on reading rate or fluency. NSD status had a main effect on reading accuracy, rate, and fluency, but not on spelling. The correlations showed the parallel trend: PA was most strongly related to reading and spelling accuracy (and also to fluency which was defined as a rate of accurate responses), whereas RAN was especially linked with reading fluency and rate. These results were in accordance with several previous studies from various orthographies (Compton et al., 2001; Cornwall, 1992; Furnes & Samuelson, 2010; Kairaluoma et al., 2013; Landerl & Wimmer, 2008; Moll, Ramus et al., 2014; Papadopoulos et al., 2009; Pennington et al., 2001; Schatschneider et al., 2002; Sunseth & Bowers, 2002; Torppa et al., 2012; 2013; Vaessen et al., 2009; Wimmer et al., 2000). However, the results presented above do not give a full closure to the discussion as single PD and NSD groups did not differ in any other aspect of literacy than spelling, which is in contrast to some previous findings from transparent orthographies. Wimmer et al., (2000) for example showed that the two single-deficit groups differed from each other in reading rate. This may be partly due to the power lost in analyses with dichotomous variables compared to the continuous ones (Branum-Martin, Fletcher, & Stuebing, 2013; Compton et al., 2001; Schatschneider et al., 2002) but also due to the sample (normative vs. RD, see Kirby et al., 2010; McBride-Chang & Manis, 1996; Meyer, Wood, Hart, & Felton, 1998; cf. Katzir et al., 2006). Even though some questions on this concern are left unanswered, the results of the present study, along with the increasing research evidence from various orthographies, support the importance of including both reading accuracy and rate measures in studies on reading development, reading disability, and double deficits (discussed e.g. in Moll, Ramus, et al., 2014; Papadopoulos et al., 2009).

The second main part of this study explored the comorbidity of RDs, MDs, and AD in the context of the double-deficit hypothesis. The results of this study support the view of a specific connection of RAN and PA with reading and spelling difficulties. First, the DDH groups differed from each other especially by reading ability, not by math performance or attention ratings. Second, the prevalence of double deficits was higher in the comorbidity groups (i.e., groups with different combinations of RDs, MDs, and AD) with RDs but not in the comorbidity groups without RDs. Third, RAN and PA were correlated with reading and spelling measures but not with math performance or attention ratings (with the exception of a small but significant correlation between PA and math). Based on these results, the deficits in naming speed and phonological awareness are related to RDs, but not to MDs, AD, or the comorbidity of RDs, MDs, and AD. These results are in line with several studies in which RAN and/or PA were specifically linked with reading, not with math or attention in normative populations (Durand et al., 2005; Georgiou et al., 2013) as well as in children with RDs, MDs, and/or AD (Denckla & Rudel, 1976; Fletcher, 2005; Heikkilä et al., 2009; Landerl et al., 2009; Moll, Göbel, & Snowling, 2014; Willburger et al., 2008; Willcutt, Betjemann et al., 2010; Willcutt et al., 2013; Wise et al., 2008). These results support the view that in explaining the comorbidity of RDs, MDs, and AD, predictors other than RAN and PA seem to be more essential. There are candidates such as
processing speed (Boada et al., 2012; Compton, Fuchs, Fuchs, Lambert, & Hamlett, 2012; McGrath et al., 2011; Willcutt, Betjemann et al., 2010; Willcutt, Pennington, et al., 2010; Willcutt et al., 2013) and short-term memory (Landerl et al., 2009; Willcutt et al., 2013), which were not included in this study but should be in the scope of further studies. Because many previous studies have found significant connections between RAN or PA and math or attention, the results are discussed next in more detail.

In contrast to previous studies (Ackerman et al., 2001; Koponen et al., 2007; 2013; van Bergen et al., 2014; van der Sluis et al., 2004), we failed to find a link between RAN and math skills even though both measures were timed and even though digits were included in a composite measure of RAN, which usually increases the correlation between these tasks (Hart, Petrill, Thompson, & Plomin, 2009; Landerl, Bevan, & Butterworth, 2004; Willburger et al., 2008). There are possible explanations for the nonexistent connection between RAN and math. First, the type of reading measure used may have had an effect on the results. In previous studies where RDs were defined based on reading fluency rather than accuracy, the group with MDs alone did not differ from the control group in RAN (Landerl et al., 2009; van der Sluis et al., 2004, however, see Koponen et al., 2013; van Bergen et al., 2014), and RAN has been uniquely associated with RDs but not MDs (Georgiou et al., 2013; Landerl et al., 2009; Willburger et al., 2008; Willcutt et al., 2013). Second, the connection between RAN and calculation may be stronger in normative samples than in the RD group (Koponen et al., 2013). This may partly explain why the RAN-calculation association was weaker in this study (the majority of the children had RDs) than the association observed in studies conducted with unselected groups (Koponen et al., 2007; van der Sluis, de Jong, & van der Leij, 2007) or those in which normative and RD groups were analyzed together (van Bergen et al., 2014). Third, the connection between RAN and calculation may diminish with development (Hecht et al., 2001; Mazzocco & Grimm, 2013) and thus may not be significant in children after the primary grades, the group that formed the majority of this sample. Finally, the result may be at least partly explained by the characteristics of the math task. Although RAN seems to be a strong predictor of simple arithmetic fluency that relies on retrieval skills (e.g., counting, calculation in single-digit tasks), the connection has been weaker with more complex calculation tasks (Hecht et al., 2001; Koponen et al., 2007, 2013) like the one used in this study. In sum, our results support the notion that RAN has no unique connection with more complex math skills in a sample with a great representation of RDs when RDs was defined based on reading fluency. The small but significant correlation between math performance and PA is in line with previous studies (De Smedt & Boets, 2010; Krajewski & Schneider, 2009; Simmons & Singleton, 2008) but as this connection was not very strong and PA did not seem to have any main effects for math performance, our results support the conclusion that PA was more closely connected to literacy skills than to math performance. In further studies, the factors presented above concerning the sample selection, the measures used and the ones that are controlled, and the developmental factors should be acknowledged in studies exploring the background skills of math performance.
Although many scholars have found a connection between attention and RAN (Arnett et al., 2012; Hynd et al., 1991; Tannock et al., 2000; Thomson et al., 2005), we failed to find a link between the two. Our results support the view that RAN is especially linked with reading and RDs, not with attention deficit. The difference between these previous results and the results of this study may derive from the fact that in many studies investigating the link between attention and RAN, reading was not controlled, or was measured only with an accuracy measure, which leaves the common variance of RAN and reading fluency uncontrolled. Thus, the results of this study are in line with studies in which no connection between RAN and attention was found when reading was controlled (Ackerman & Dykman, 1993; Felton & Wood, 1989; Raberger & Wimmer, 2003; Semrud-Clikeman et al., 2000; see also Boada et al., 2012 for review). The results of this study also confirmed the findings of previous studies where no connection between PA and attention was found (Gooch et al., 2011; McGee et al., 2004; Purvis & Tannock, 2000, Willcutt et al., 2001).

Several practical implications for evaluating and treating reading disabilities can be derived from the results of this study. First, in assessing reading disabilities, the knowledge that RAN and PA are linked with unique aspects of reading gives an additional tool for identifying RDs early and provides relevant information when planning interventions for children with RDs. Deficits in RAN refer to problems in reading fluency, and the intervention could be targeted at automatizing reading skills. Phonological deficits instead are connected with problems in reading and spelling accuracy especially in the initial stage of reading acquisition, whereupon the intervention could primarily emphasize basic decoding and spelling skills. Children with a double deficit have a greater risk for more severe reading disabilities than children with a single deficit in PA or RAN. Since measures for PA and RAN have been shown to predict reading skills before schooling (Papadopoulos et al., 2009; Puolakanaho, Poikkeus, Ahonen, Tolvanen, & Lyytinen, 2004; Puolakanaho et al., 2008; Torppa et al., 2012, 2013), preventive interventions could be targeted especially at this group of children with the greatest risk of RD.

Second, consistent with previous findings from transparent orthographies reading accuracy of age-appropriate text was high (93%) even in a clinical sample that showed a high prevalence of RDs. Our results for the children with learning difficulties thus solidify the results of previous studies in transparent orthographies with unselected samples (Aro & Wimmer, 2003; Landerl & Wimmer, 2008; Seymour et al., 2003), since the RDs were manifested primarily as fluency problems with relatively high accuracy in age-appropriate reading material.

Finally, even though RAN and PA seem to be especially connected to RDs and not to explain the comorbidity of disabilities in reading, math, or attention, the high comorbidity of these disorders should be acknowledged in learning assessment and intervention planning. As DuPaul and colleagues (2013) have suggested, screening of academic problems as well as the problems in attention should always be included in assessment among children with LDs or with attention problems. Even though these deficits probably have a shared genetic basis (e.g., Willcutt, Pennington, et al., 2010), they have unique characteristics that should be acknowledged in the intervention since different combinations of learning problems may respond differently to intervention (Boada et al., 2012; Fuchs et al., 2013). For example, for children with
comorbid RDs and AD, including components of attention training in intervention may also benefit reading skills (Aro, Ahonen, Tolvanen, Lyytinen, & de Barra, 1999; see Sexton et al., 2012 for a review).

As in all clinical samples, limitations affect the generalizability of the results. Especially notable is that the comorbidity rate may be inflated because children with multiple deficits may be more likely to be referred for evaluation and intervention (Semrud-Clikeman et al., 1992; Waber et al., 2003), especially those with hyperactive and impulsive types of attention problems (Willcutt & Pennington, 2000). Thus, generalizing these results outside clinical samples may lead to overemphasis on comorbidity. Other types of referral bias may also be involved, and it is possible that some disabilities are overemphasized in the sample (e.g., prevalence of RDs in this study seems to be relatively large compared to the prevalence of MDs and AD). This also leads to methodological problems since many of the comorbidity groups without RDs were small and only trend-level analyses for comorbidity could be made. Second, we are aware that categorizing continuous variables such as RAN or reading in dichotomous groups reduces the power to detect relationships between variables. However, unreported regression analyses conducted with the data revealed the same pattern reported here.

In this study, the issue of comorbidity of RDs, MDs, and AD was explored in the framework of the double-deficit hypothesis for the first time. In contrast to many studies that have explored comorbidity, reading fluency was used as a definition for RD, which is a key factor in explaining the relationship between RAN, RDs, MDs, and AD especially in transparent languages. In future studies searching for the predictors for comorbidity between these deficits, the scope should shift more towards the multiple deficit explanations of learning disabilities (Pennington, 2006; 2012) and include more protective and risk factors in addition to RAN and PA. These factors should include working memory and processing speed that have been shown to explain a significant part of the common variation between developmental disorders (Boada et al., 2012; Compton et al., 2012; Landerl et al., 2009; McGrath et al., 2011; Willcutt, Betjemann, et al., 2010; Willcutt, Pennington, et al., 2010; Willcutt et al., 2013). In addition, longitudinal studies following the precursors of comorbidity are important in investigating the primary and secondary effects of these disorders. In future studies on double deficits, longitudinal studies that include the phase before emerging reading skills should be conducted to explore the causality or reciprocity of the precursors of reading and emerging reading skills (see e.g. Castles & Coltheart, 2004, for a discussion on PA and reading). Finally, intervention studies for RDs derived from different background deficiencies (NSD, PD, DD) should also be developed and studied.

In sum, PA or RAN alone was not sufficient in explaining RD, but when they were measured together, they detected the majority of the RDs in this sample: 90% of children with RD had a deficit in phonological awareness, rapid automatized naming, or both. PA and RAN were especially associated with RD, not MD or AD. These results strengthen the position of RAN and PA as background skills of reading disabilities also in clinical samples with comorbid math and attention deficit.
ACKNOWLEDGEMENTS

This paper is part of the primary author’s doctoral dissertation. Financial support for this research was provided for Riikka Heikkilä by National Doctoral Programme of Psychology. The authors want to acknowledge Kenneth Eklund, Tuire Koponen, Kaisa Lohvansuu and two anonymous reviewers for their valuable comments on the manuscript and the Child Research Clinic of the Niilo Mäki Institute for the provision of the data.
REFERENCES


