Published online 11 October 2005 in Wiley InterScience
(www.interscience.wiley.com). DOI: 10.1002/dys.296

How Persistent are Phonological Difficulties? A Longitudinal Study of Reading Retarded Children

Idor Svensson* and Christer Jacobson

School of Education, Växjö University, 351 95 Växjö, Sweden

The present study examined the persistency of phonological deficiencies over time. The participants were 40 pupils in grade 2 with documented reading and writing difficulties and a comparison group of 30 pupils. The participants were followed over a 10-year period by word- and non-word-reading tests and tests of cognitive ability. The persistence of phonological deficits was indicated by a high correlation between non-word-reading tests in grades 3 and 12 in the reading-disabled group. A dyslexia cut-off definition based on phonological ability was the most consistent definition over time compared to a word-decoding definition or multiple cut-off definition based on IQ. Phonological decoding abilities were remarkably stable over time, and non-word-reading was found to be a valid instrument in diagnosing and discerning dyslexia both in children and adults.

Keywords: phonological development; non-word-reading; word recognition; dyslexia definitions

INTRODUCTION

An estimated 3–7% of the population shows specific reading and writing disorders (dyslexia). The variation in number reported as dyslexic depends on the definition of the concept. There is still no consensus on a definition of dyslexia that can dissociate between dyslexia and generally poor literacy skills. However, most researchers agree that a phonological deficit is an underlying factor that accounts for dyslexics' initial poor word-decoding, which is a main manifest symptom of dyslexia, and for their poor ability to learn the relationship between letters and sounds (Høien & Lundberg, 2000; Snowling, 2001; Stanovich, 1988; Stanovich & Siegel, 1994; Vellutino & Scanlon, 1987). Phonological difficulties are related to an information-processing mechanism at a cognitive level. Høien and Lundberg (2000) proposed the following definition

*Correspondence to: Idor Svensson, School of Education, Växjö University, 351 95 Växjö, Sweden. E-mail: Idor.Svensson@iped.vxu.se

'Dyslexia is a disturbance in dealing with the code of the written language based on a deficit in the phonological system of the spoken language' (p. 20). However, the persistence of the phonological deficit and the development of childhood dyslexia during a lifetime are in need of more investigation.

The theory underlying our research is that a phonological deficit is the core factor in dyslexia. There are two main aims of the present study. Firstly, to examine the persistency of phonological functions from grades 3 to 12 in a longitudinal study. Secondly, to investigate the stability of different definitions of dyslexia.

Dyslexia involves many factors and a major problem concerns the difficulties in reading and writing words. However, these difficulties may be merely a symptom of the underlying problem and may not tell the whole truth about its origin. Frith (1999) suggested a three-level model to better understand dyslexia. The model includes a biological a cognitive and a behavioural level and each level is related to environmental factors. Frith also notes the connection between the levels. The cognitive level includes phonological aspects where both speech and reading are processed and it can be seen as a bridge between the biological and behavioural levels. The biological level that includes genes and brain function is interpreted as the deepest level and might be the origin of dyslexia.

Frith's model is general and can be applied to many developmental disorders of biological origin. Lundberg (1999) has suggested a framework that is similar but more specific to dyslexia (the hypothesis of cognitive modularity). Lundberg claimed that the core factor in dyslexia is a malfunction in phonological processing that could coexist with excellent functioning of other cognitive abilities in the brain. This phonological process must become automatic and work with speed and accuracy, enabling the individual to finally achieve fluent and effortless reading. However, Lundberg also claimed that there are several factors on a manifest (behavioural) level that influence reading and writing ability, the early environment, maturation, motivation, instruction and compensation.

The phonological deficit theory states that dyslexia is basically a problem of processing phonological information, such as problems affecting short-term memory, sound segmentation and sound blending (Rack, Snowling, & Olson, 1992; Stanovich, 1988). On the behavioural level the most obvious signs are slow, effortful and error-prone word decoding. There is now a compelling body of evidence of a strong relationship between phonological awareness and later success in reading acquisition (Borstrøm & Elbro, 1997; Lundberg, Frost, & Petersen, 1988; Lundberg, Olofsson, & Wall, 1980; Sprugevica & Høien, 2003; Wagner & Torgesen, 1987).

A Danish longitudinal investigation (Lundberg *et al.*, 1988) showed that dyslexic individuals have poorer phonological awareness compared with age-matched controls at least in preschool and at the beginning of the compulsory school years. In a follow-up study Lundberg (1994) investigated children running the risk of having reading disabilities on account of their phonological awareness skills during preschool (i.e. those children who perform lowest on a phonological awareness test). The children who had no language training in preschool performed more poorly on word decoding, spelling and reading comprehension in grades 3 and 4. These findings have been replicated repeatedly (see for example Schneider, Küspert, Roth, Visé, & Marx, 1997; Kjeldsen, Niemi, &

Olofsson, 2003). However, early intervention does not predict whether or not phonological processing is important for the development of reading skills in a lifetime perspective (or at least to early adulthood). There has been some research focusing on phonological processing abilities over time and the persistence of dyslexia (Bruck, 1992; Elbro, Nielsen, & Petersen, 1994; Jacobson, 1999; Pennington, Van Orden, Smith, Green, & Haith, 1990; Shaywitz *et al*., 1999). Pennington *et al.* (1990) found a clear deficit in phoneme awareness in adult dyslexics compared to reading- and age-matched control groups. Bruck (1992) showed similar results and claimed that even if dyslexic adults show adequate word decoding skills they still have phoneme awareness deficits. In the normally achieving group there was an increasing development of phoneme awareness with age and reading level in contrast to the dyslexic group. Thus, phonemic awareness seems to be a persistent problem among adult dyslexics even if they have compensated for reading and writing disabilities on a manifest level. Elbro *et al.* (1994) investigated dyslexia in adults and confirmed the above results where there was a distinct difference in phonological coding in reading in comparison with non-dyslexics.

All of the above research concerning reading and writing difficulties in adults showed that there were persistent phonological deficiencies. However, none of the studies are longitudinal and 'the best way to learn about dyslexia in adulthood is to follow up a person who was diagnosed as dyslexic during childhood' (Elbro *et al*., 1994, p. 207). In a follow-up study of adults with childhood dyslexia, Olofsson (2002) concluded that there were clearly persistent problems in tasks involving phonological processing. Moreover, in a longitudinal study by Shaywitz *et al.* (1999) the persistence of dyslexia in adolescence was investigated. They followed a group of well-defined dyslexic children up to adolescence. The results showed that deficits in phonological coding still characterized dyslexic readers in adolescence. Those measurements of phonological processes are thus useful as discriminators between dyslexic and average readers. Furthermore, even if the dyslexic group had been exposed to remedial help they did not catch up in reading ability. Jacobson (1999) found corresponding results when analysing the persistence of reading disabilities. Jacobson used a word recognition test in measuring reading difficulties. The main findings were that poor readers in grade 2 did not catch up in reading ability by grade 9 despite extensive interventions of remedial instruction. The present study is a prolongation of Jacobson's study (1999), which followed the children up to grade 9.

Further evidence in favour of the phonological deficit theory can be seen in the genetic research that has been presented in the dyslexia domain during the last decade (Fagerheim *et al*., 1999; Fisher *et al*., 2002; Grigorenko *et al*., 2001; Petryshen *et al*., 2001). Almost all of these studies indicate relationships between phonological phenotypes and loci on chromosome regions. Furthermore, a recent study by Samuelsson and Lundberg (2003) shows that the phonological ability was relatively unaffected by environmental influences. In our study we have used a simple cut-off definition based on phonological decoding ability. This has been argued by Elbro (1998) to be both an externally valid and a reliable definition. Furthermore, Elbro (1998) reported that a simple cut-off definition in adults provided a better match to childhood histories of reading difficulties than did a discrepancy definition that took vocabulary into account.

Irrespective of which theory is used to describe dyslexia there is not yet any clear criterion for defining it i.e. it is not an either/or issue (Stanovich, 1988). The long-standing apprehension concerning the definition of dyslexia is due to the discrepancy hypothesis, in other words, reading and spelling difficulties cannot be explained as the results of poor intelligence. One form of discrepancy definition for identifying dyslexia is the regression-based definition. Fletcher *et al.* (1989) compared this definition with a cut-off definition. The results revealed clear differences between the two definitions regarding which children were identified as reading-disabled (RD). However, the researchers argued that there is little evidence for any specificity of reading problems depending on the definition. The assumed discrepancy between reading ability and general intelligence has been hotly discussed in research during the last decade (D'Angiulli & Siegel, 2003; Gustafson & Samuelsson, 1999; Gundersen & Siegel, 2001; Lyon, 1995; Lyon, Shaywitz, & Shaywitz, 2003; Siegel, 1989, 1992, 1999; Siegel & Himel, 1998; Stanovich, 1994). However, Raskind (2001) argued for the use of verbal IQ in genetic analyses of dyslexia. Reading disability defined by the IQ-discrepancy criterion may have a stronger genetic load than reading disability defined by low performance criterion. However, this might be a weak argument since verbal IQ partly depends on the development of literacy skills. Siegel and Himel (1998) claimed that the use of IQ might even exclude some individuals that would otherwise have been regarded as dyslexic. In their study it was demonstrated that the IQ of dyslexics decreased with age. Thus, an individual that was regarded as dyslexic in childhood could therefore, as an adult, be reclassified as a poor reader when the discrepancy definition was used. Another argument put forward by Siegel (1999) was that an IQ test measures what a person has learned, not what he or she is capable of learning.

Several studies have shown that non-word-reading and tests measuring phonemic awareness, spelling, pseudo-homophone words and word decoding, are very reliable tests of dyslexia (Elbro *et al.*, 1994; Hatcher, Snowling, & Giffiths, 2002; Rack *et al.*, 1992; Shaywitz *et al.*, 1999; Snowling, Nation, Moxham, Gallagher, & Frith, 1997). Grigorenko (2001) argued for the importance of longitudinally defined dyslexia phenotypes, that is tests that are resistant over the lifespan. It is a fact that the performance on reading-related tests changes over time. Word-decoding can be a reliable test of reading difficulties in the early years of reading, but in adult readers spelling and non-word-reading might be more predictable of reading and writing difficulties (DeFries, Alarcón, & Olson, 1997; Castles, Datta, Gayán, & Olson, 1999). Non-word-reading speed seems to be one of the most reliable tests for discerning dyslexia in both children and adults even in languages with a simple grapheme-phoneme correspondence (Wimmer, 1996). Snowling *et al.* (1997) claim that non-word-reading and a phonemic awareness test are best at identifying students with dyslexia. Elbro *et al.* (1994) and Hatcher *et al.* (2002) share this view of non-word-reading.

There are two main aims of the present study. Firstly, to examine the persistency of phonological ability from grades 3 to 12 in an RD group and a comparison group. We used two phonological tests, non-word-reading and phonological choice. Secondly, we investigated whether the RD children identified as dyslexic in grade 3 are still dyslexic in grade 12, using different cut-off definitions (phonological, decoding and multiple cut-off definitions based on IQ).

The hypothesis of the present study was that early phonological difficulties persist in adulthood. Consequently, the assessment of phonological processes would be a forceful instrument for the definition of the concept of dyslexia and furthermore for diagnosing dyslexia both in children and adults.

METHOD

Participants

A group of poor readers ($n = 103$) and a comparison group ($n = 90$) were selected from the population ($N = 2167$) of children attending the second grade (age 8–9) in the county of Kronoberg, Sweden. The basis for the selection of poor readers was low scores (below the fifth percentile based on compound score) in two tests of word decoding (Wordchains test and OS 400) and the teachers' ratings of the children's reading ability. The purpose of the Kronoberg study was not primarily to identify children with dyslexia. Thus, we did not use either a regression-based or a discrepancy definition but only a simple cut-off definition. Mentally retarded children as well as children with an immigrant background were excluded from the study. The participants in the RD group were then carefully matched to a comparison group on a non-verbal IQ test (Ravens matrices) in grade 2 as well as by gender and school. In the present study a subset was available, consisting of 40 poor readers (35 boys and five girls) and 30 participants (27 boys and three girls) of the original comparison group (age 19–20). For a more detailed description of the sample, of test battery-administration and inclusion criteria, see Jacobson and Lundberg (1995) and Jacobson (1999). No systematic differences between those who participated in the present study and those who were excluded were found. The attrition rate was mainly due to practical, economic or geographical circumstances.

Assessments

Selection instrument

Two tests were used as screening instruments for the participants' word recognition ability. One of the tests was the Wordchains test (Jacobson, 1995) and the other was *OS400* (Sögaard & Bording Petersen, 1974), which is a word–picture matching test. The task was to select from four pictures the one which corresponded to a target word. The measure was the number of correctly marked words in 10 min.

Wordchains test

Grades 2 and 12. In the word recognition test, the *Wordchains* test (Jacobson, 1995; Miller Guron, 1999), the task was to discriminate three words in a chain without space by marking with a pencil where the divisions should be (e.g. sandcoffeblue = sand/coffee/blue). A large number of wordchains (120) were presented. The performance is expressed by the number of correctly marked word chains in 3 min. Standardized norms from grade 1 to adults were available (Jacobson, 1995).

Non-word-reading

Grade 3. A list of non-words was presented to the participants. The participants' task was to read aloud as many non-words (80 items) correctly as possible in 1 min. The non-words presented in grade 3 consisted of one or two syllables. The norms for non-word-reading in grade 3 were based on the 90 subjects in the original comparison group of the Kronoberg project (Jacobson, 1995).

Grade 12. The task was to read the non-words (20 items) aloud as fast and accurately as possible. The non-words varied in length from two to five syllables and followed Swedish phonotactic rules (they were pronounceable). The time needed to read non-words correctly was recorded. The score was then transformed to the number of correctly read non-words per minute in order to adjust the score to grade 3. The norms for non-word-reading in grade 12 are based on 92 subjects (Jacobson & Svensson, 2003).

Phonological choice

Grade 12. In order to validate the non-word-reading test we employed a phonological choice test in grade 12. A large number of pseudo-homophones were presented (this test was not given in grade 3). The pseudo-homophones were constructed in such a way that their pronunciation but not their spelling was identical to that of a real word. The task was to read and underline the words that sounded exactly like a real word out of two pseudo-words presented in rows. The performance was expressed as the number of correctly chosen homophones in 2 min. Swedish norms were available (Olofsson, 1994).

Verbal ability

Grade 3. The Illinois test of psycholinguistic abilities (ITPA, Holmgren, 1984) for children was administered as a test of verbal communication ability. Two subtests from the ITPA test were used, *Auditory analogy* (for example; grass is green, milk is…) and *Auditory reception* (for example; can birds paint? yes or no), as a measure of verbal ability. The ITPA tests were only administered to the RD children. Standardized Swedish norms were available (Holmgren, 1984).

Grade 12. To measure vocabulary, a list of 40 target words was presented in grade 12 (Johansson, 1992). The participant had to choose which of five alternative words was the best synonym of the target word. The number of correctly marked words was recorded. Standardized Swedish norms were available (Johansson, 1992).

Non-verbal ability

Grade 2. Raven's Coloured Progressive Matrices (Raven, 1965) was given as a group test. This test was used as a matching variable in grade 2. The maximum score was 33.

Grade 3. The participants carried out a non-verbal logical reasoning test, the *Figure Relation Test* (Westerlund & Ullstadius, 1991). The maximum score was 26.

Grade 12. A short form of *Raven's Standard Matrices* (Raven, 1995) composed of scales C and D, was given to the participants in grade 12. The maximum score was 24.

As a measure of IQ ability we have used a compound *z*-score including two verbal (*Auditory analogy* and *Auditory reception*) and one non-verbal test (*Figure Relation Test)* in grade 3 and one verbal (*Vocabulary*) and one non-verbal test (*Raven*) in grade 12. We have compared four definitions of dyslexia to study the stability of the concept.

1. A phonologically based definition (one S.D. below the mean on the non-word-reading test).
2. A decoding-based definition (one S.D. below mean on the Wordchains test).
3. A phonological approach based on IQ (at least one S.D. below the mean on the non-word-reading test and above −1.0 S.D. on the compounded IQ score).
4. A decoding approach based on IQ (at least one S.D. below the mean on Wordchains test and above −1.0 S.D. on the compounded IQ score).

Procedure

Data for the selected groups were collected in grades 3 and 12 (except for *Wordchains* and *Raven* which were administered as group tests in grade 2). Details for the data collection procedure in grade 3 have been reported elsewhere (Jacobson & Lundberg, 1995). The complete test battery in grades 2 and 3 was administered on several occasions in contrast to the test procedure in grade 12, which was performed in one session lasting from 1.5 to 2 h. Data for grade 12 was collected during the participants' last semester in high school and no later than half a year after they had finished high school. The tests were administered at the various schools of the participants', at their homes or at our clinic.

RESULTS

Firstly, we show the mean, standard deviation (S.D.) and *t*-value for all tests included concerning the disabled children and the comparison group. Secondly, we examine the persistence of the phonological ability expressed as non-word-reading in grades 3 and 12. Thirdly, the correlations for word recognition in grades 2 and 12 are shown. In addition we have carried out a cross-lag correlation between non-word-reading and word recognition. Finally, the results concerning the persistency across different definitions of dyslexia are presented.

Table 1 shows significant differences on all reading-related tests with almost no overlap where the controls outperformed the RD. There was no significant difference on the matching variable in non-verbal IQ ability (Raven) in grade 2. However, in grade 3 the *Figure Relation Test* showed significant differences between the groups. Furthermore, there was a significant difference in vocabulary as well as in Raven's Matrices in grade 12.

Figure 1 shows the stability of phonological ability expressed as non-word-reading in grade 3 and 12. We present separate correlations for the two groups as well as the combined correlation. The figures reveal that there was a high combined correlation (Pearson, $r = 0.81$, $p < 0.01$) between non-word-reading in grades 3 and 12. When we separated the groups, the correlations for the comparison group were lower ($r = 0.46$) but still significant ($p < 0.01$). In the RD group the high correlation was maintained ($r = 0.73$, $p < 0.01$). Even when

Table 1. Results for the RD and the comparison group on all tests

		RD ($n = 40$)		Controls ($n = 30$)		
	Grade	M	S.D.	M	S.D.	t (df)
Word decoding	2	11.3	3.4	33.4	7.7	−16.29 (68)***
Word decoding	12	57.4	14.2	81.1	9.7	−7.88 (68)***
Non-word-reading	3	22.6	8.8	49.9	12.9	−10.54 (68)***
Non-word-reading	12	19.2	7.5	35.5	10.0	−7.84 (68)***
Ravens matrices	2	20.1	5.2	20.6	4.9	−0.41 (68)
Figure Relations	3	14.7	5.1	18.4	3.4	−3.45 (69)***
Ravens matrices	12	16.1	4.0	19.6	2.0	−4.21 (68)***
Vocabulary	12	25.8	5.9	32.1	4.5	−4.91 (68)***

***$p < 0.001$.

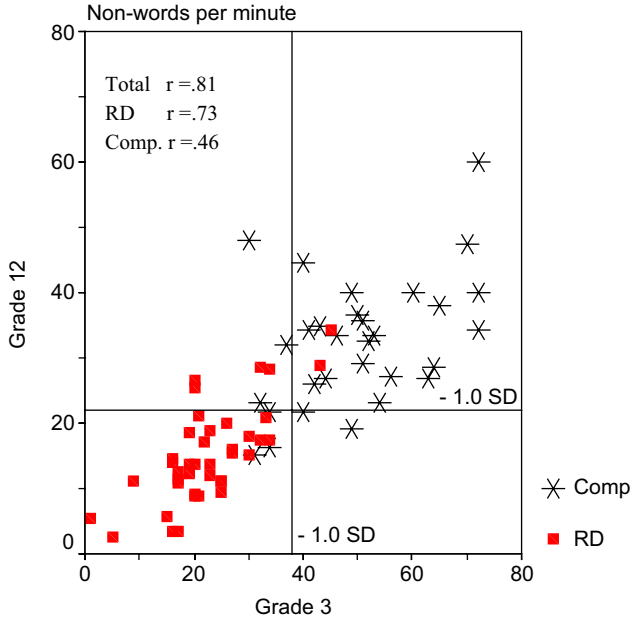

Figure 1. Scores for number of correctly read non-words per minute in grades 3 and 12 plotted for the RD and the comparison group.

controlling through partial correlation for IQ (a combined score of non-verbal and verbal IQ) in the RD group the correlation was still powerful ($r = 0.72$, $p < 0.01$). For the comparison group the correlation was higher when controlling for IQ ($r = 0.62$, $p < 0.01$). From the figure we can also see that two poor readers consistently performed above −1.0 S.D. on both occasions and two individuals from the comparison group performed below this boundary line.

Figure 2 presents non-word-reading in grade 3 plotted against the performance of the phonological choice test in grade 12. Figure 2 confirms the high correlation between the phonological measurements in Figure 1 ($r = 0.79$, $p < 0.01$). However, in the RD group the correlation was lower ($r = 0.55$, $p < 0.01$) and for the

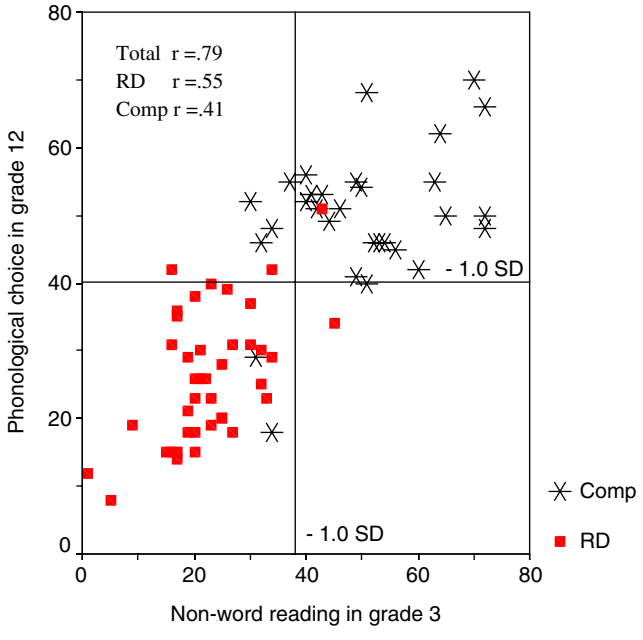

Figure 2. Scores for phonological choice and non-word-reading in grades 3 and 12 plotted for the RD and the comparison group.

comparison group the correlation was $r = 0.41$ ($p < 0.05$). The correlation was almost the same when controlling for non-verbal and verbal IQ ($r = 0.54$, $p < 0.01$) for the RD group. For the comparison group the correlation was slightly higher ($r = 0.49$, $p < 0.05$). The relation between the two phonological tests in grade 12, non-word-reading and phonological choice, was $r = 0.77$, $p < 0.01$).

In Figure 3, the correlations for word decoding in grades 2 and 12 are shown. There was a high correlation (Pearson) between the grades in the Wordchains test ($r = 0.72$, $p < 0.01$) but when we separated the groups the correlation was fairly moderate ($r = 0.48$, $p < 0.01$) in the RD group and quite low ($r = 0.31$, ns) in the comparison group. When controlling for a combined IQ score the correlation analysis revealed a lower relation for the RD group ($r = 0.39$, $p < 0.05$) and for the comparison group ($r = 0.21$, ns). From Figure 3 it is possible to observe that there is no overlap on the Wordchains test between the groups in grade 2. However in grade 12 the overlap is substantial. It is obvious that all participants have improved their word recognition skills. When we examined the relation between non-word-reading in grade 3 and the Wordchains test in grade 12, for the RD group, and controlled for IQ, the correlation was fairly high ($r = 0.54$, $p < 0.01$). Nevertheless, 10 out of 16 participants in the RD group that had reached at least average in word recognition in grade 12 (>60 word chains) were still more than 1 S.D. below norms in non-word-reading in grade 12.

The main reading tests in this study are based on decoding ability and they show that the difference between the groups is clearly pronounced. However, the ultimate goal of reading is to comprehend a text. Therefore we have also looked at a standardised achievement test in reading comprehension for grade 9, distributed by the Swedish Board of Education. Furthermore, to validate the

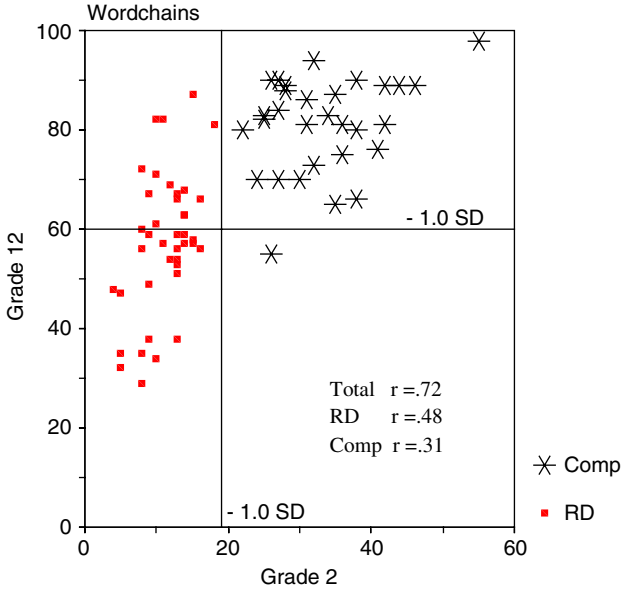

Figure 3. Scores for word decoding in grades 3 and 12 plotted for the RD and the comparison group.

decoding ability we studied marks in the Swedish language in grade 10. There were significant differences in both the achievement test $t(63) = -7{,}54$, $p < 0.001$ and the marks $t(54) = -4.56$, $p < 0.001$, where the comparison group outperformed the RD group.

It is clear that there are high correlations between phonological and word recognition tests between grades 3 and 12. However, it is less easy to see what is the relation between those tests and whether it is possible to see any causal direction between them. In Figure 4 we show a cross-lag correlation for Wordchains and non-word-reading in the RD group.

The correlations are overall quite high and significant. However, the correlations between non-word-reading in grade 3 and Wordchains in grade 12 are higher ($r = 0.60$) than the correlations between Wordchains in grade 2 and non-word-reading in grade 12 ($r = 0.48$). This result might indicate that non-word-reading in lower grades explains the greater variance (36%) in word decoding in higher grades than vice versa (23%). The results on the cross-lag were similar when controlling for IQ. In the control group the cross-lag correlations were lower and only the correlation between non-word-reading in grade 3 and Wordchains in grade 2 ($r = 0.45$) was significant. The cross-lag correlation in the total group was high (in the range of $r = 0.62-0.83$).

Finally, the results concerning the persistency of different definitions of dyslexia are presented.

Table 2 reveals that a phonological definition shows the highest stability through the grades. Thirty-eight subjects in the RD group were regarded as dyslexic in grade 3 according to the phonological definition, and 34 of them were still regarded as dyslexic in grade 12. When using only word decoding ability in defining dyslexia, 38 children were regarded as dyslexics in grade 3 and 23

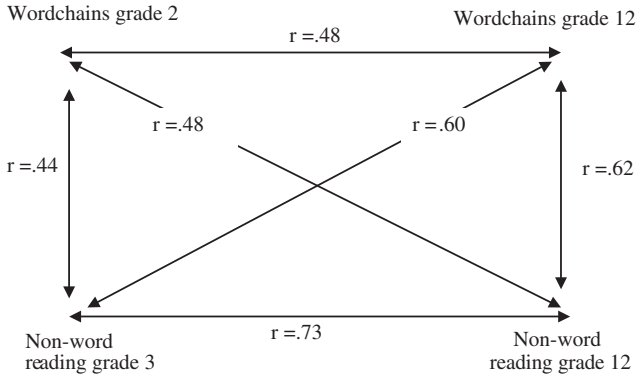

Figure 4. A cross-lag correlation for the RD group.

Table 2. Stability of dyslexia for 40 RD children on the basis of different definitions in grades 3 and 12

	Stable classification		Unstable classification	
Type of definition	Dyslexic in 3 and 12	Not dyslexic in 3 and 12	Dyslexic in 3 not in 12	Dyslexic in 12 not in 3
Phonological ability	34	2	4	0
Decoding ability[a]	23	1	15	1
Phonological ability and IQ	11	13	10	6
Decoding ability[a] and IQ	6	15	15	4

[a] Presented as a group test in grade 2.

subjects were still regarded as dyslexics in grade 12. When phonology and IQ was included in the dyslexia definition, 21 subjects were regarded as dyslexics in grade 3 and 11 were still regarded as dyslexics in grade 12. When word-decoding and IQ was included in the dyslexia definition, 21 subjects were regarded as dyslexics in grade 3 and only six were still regarded as dyslexics in grade 12. With the definitions based on IQ, many subjects were never regarded as dyslexics. Furthermore, quite a few of those subjects that were regarded as dyslexics in grade 3 were considered not to be dyslexic in grade 12. Some of those that were not dyslexic in grade 3 were found to be dyslexics in grade 12.

DISCUSSION

In the present study we could confirm that phonological deficiencies shown in early school years still persist in adulthood. Furthermore, a phonologically based definition of dyslexia exhibited more stability over time compared to definitions based on word-decoding tests and/or definitions based on IQ.

The correlation of non-word-reading between grades 3 and 12 was high notwithstanding the short test time and the fact that the version of the tests was

not the same in grades 3 and 12, both of which factors influence the reliability negatively. In addition, the current sample belongs to the 5% constituted by the poorest readers in a cohort of 2200 students (the large-scale study, see Jacobson, 1999), which makes the high correlation even more remarkable in consideration of the restriction of range. The current findings are consistent with several other studies indicating that deficits in phonological coding persist into adulthood (Bruck, 1998; Downey, Snyder, & Hill, 2000; Olofsson, 2002; Shaywitz *et al*., 1999). Bruck (1998) compared non-word-reading among dyslexic children and dyslexic adults and found that they were equally impaired relative to their age group. In the present study we followed the same participants longitudinally, which provides stronger evidence of a core phonological deficit.

Furthermore, even if RD students in the present study increased their word recognition ability over the years, their phonological ability seems to show less progress, despite quite intensive remediation during the school years (Jacobson, 1999). This fact is very similar to the conclusion drawn by Bruck (1998). Bruck concluded that even if the dyslexics have been rather successful in their educational career and have increased their decoding skills they still have the same primary phonological deficiencies as they had as children.

With our cut-off criteria on the Wordchains and the non-word-reading test a majority (10 out of 16) of the RD persons that had reached average in word recognition in grade 12 were still below −1.0 S.D. in non-word-reading. It is reasonable to assert that those with average reading skills as adults but still having phonological deficiencies could be compensated dyslexics. The other pupils (6 out of 16) might be slow starters in reading acquisition.

The cross-lag procedure indicated that non-word-reading in grade 3 explained more of the variance (36%) of the Wordchains test in grade 12 than Wordchains in lower grades explained Wordchains in higher grades (23%). Thus, an interesting finding in our study was that a non-word-reading test actually seems to predict word recognition better than a word recognition test. However, the causal direction was moderate. Nevertheless, an abundance of research confirms this statement. For example, van der Leij *et al*. (2001) maintain that a deficit in non-word-reading is a core factor in dyslexia through the lifespan (see also Elbro *et al*., 1994; Grigorenko, 2001; Hatcher *et al*., 2002; Rack *et al*., 1992; Snowling *et al*., 1997; Snowling, 1998; Stanovich & Siegel, 1994).

Moreover, in this investigation we have also shown that a simple cut-off phonological definition of dyslexia is more stable than definitions that include discrepancy criteria using verbal and non-verbal IQ factors. There is now substantial evidence indicating that IQ does not contribute to clarify the definition of dyslexia (Gustafson & Samuelsson, 1999; Siegel, 1992; Siegel & Himel, 1998; Stanovich, 1998). In line with this, the present study demonstrates that using a definition based on IQ seems to confuse the picture of who is dyslexic and who is not. There was a lower stability when using IQ in the definition, that is, fewer participants were regarded as dyslexic in grade 3 as well as in grade 12. At the same time there were some individuals who were regarded as dyslexic in grade 12 even when they were not considered as dyslexic in grade 3. However, had our RD group been selected with reference to IQ, this definition might have been more stable. In the Kronoberg study we decided not to utilize a discrepancy definition since at the beginning of the project the aim was not specifically to identify dyslexic children. To use IQ in the definition may confuse

the task of finding the core deficits in dyslexia (Gustafson & Samuelsson, 1999) and, even worse, exclude some of those individuals that might have dyslexia.

In the current study it is obvious that tests (such as non-word-reading) aiming at measuring skills beneath the surface level seem to be more stable than those on the manifest level. Thus, word recognition would be more sensitive to environmental factors such as print exposure, instruction and motivation than non-word-reading (Lundberg, 1999; Samuelsson & Lundberg, 2003). However, the lower stability of the Wordchains test (on the manifest level) might be explained by the fact that the test was presented as a group test in grade 2 compared to the other tests that were presented as individual tests in grades 3 and 12.

For good readers Wordchains could have a 'soft' ceiling effect due to a restricted manual ability (quickly drawing lines between the words in the chains). This effect might explain the low correlation in the comparison group.

However, it is important to be careful and not be too reductionistic in the struggle to find the core factor in dyslexia, as there is a risk of excluding other important signs. Nevertheless, until we have a greater knowledge of the biological issues, the cognitive level can serve as a bridge between the biological and behavioural levels in understanding and deriving indices of dyslexia (Frith, 1999).

The focus in the present study has been on the phonological module on the cognitive level, that is the hypothesis of cognitive modularity (Lundberg, 1999). Lundberg has argued for the module as a specific part of the cognitive system, which is impaired in dyslexics and presented as a hump at the lower end of the distribution of reading skill (that is, a sub-distribution of dyslexic individuals). The findings in the current study confirm the deficit hypothesis since we found high correlations and stability in phonological skills over a 10-year period. Notwithstanding, many factors on the manifest level had probably influenced the subjects' reading skills during the years between the measurement occasions. Frederickson and Frith (1998) found that 80% of potential dyslexics showed phonological deficits on testing. In our study we found an even higher percentage (89%).

The outcome of our findings may help professionals who work with reading and writing disabilities in the diagnostic procedure and in deciding on the appropriate remediation. Students at the university level can be especially hard to diagnose since they have often compensated for their low word recognition ability. However, when university students with a compensated word-decoding ability are required to read a large amount of text they soon apprehend this as too overwhelming and hence run the risk of failing on tests and examinations. With reliable measurements, such as non-word-reading, the diagnostic procedure is made more stringent, enabling earlier remediation and breaking a vicious circle. Furthermore, a stable test, such as non-word-reading, would also enhance the certainty when diagnosing bilingual children and adults with reading and writing disabilities, since the linguistic surface level can be quickly acquired in a new language among non-dyslexic individuals (for a review see Frederickson & Frith, 1998; Miller Guron & Lundberg, 2003).

One restriction in this study is that there are few tests that allow for a comparison of reading ability in grades 3 and 12. There was an extensive assessment battery in grade 3 but unfortunately some of these tests could not be

used in grade 12 or in the comparison group. Another limitation was that we did not use the same non-word-reading test in grades 3 and 12. In 1989, which was the starting point of the large-scale research project, the knowledge of phonological impact in the reading process was less widely accepted and therefore the use of phonological tests was not common.

In addition, the cut-off level we have used in the present study concerning the definition issue (below −1 S.D. in non-word-reading and word recognition and above −1 S.D. in IQ) might be too inclusive. Thus, some of the participants in the RD group might have been incorrectly regarded as dyslexic. However, there is no clear cut-off in diagnosing dyslexia. Siegel (1999) has discussed this issue and claims, 'there is no way of knowing what a valid cut-off score is' (p. 310). A score below the 25th percentile is a cut-off value that researchers in the learning disability area have often used (see Siegel, 1999, for a review) and therefore 1 S.D. seemed reasonable for our study.

A limitation of this study is that we have only used word-decoding tests as a measurement of reading at the behavioural level. We have access to reading comprehension tests, but unfortunately not from grade 12. However, when we looked at a standardized achievement test from grade 9 and marks in Swedish from grade 10 we found that the comparison group clearly outperformed the RD group. This finding confirms that there are undoubtedly large differences in reading ability between the groups.

Another limitation of this study is that the samples are rather small. However, since the attrition rate was due to practical, economic or geographical circumstances we believe that there is no systematic bias in the results.

The fact that there was a significant difference between the RD and the comparison group concerning non-verbal ability in grades 3 and 12 must be taken into consideration when interpreting the results (the differences in vocabulary were as expected). The differences in non-verbal ability could be due to several factors such as a real difference in IQ, where the comparison group performed at a higher level than the RD.

Another explanation could be that the performance of the RD declined in respect of IQ just because of their disability in literacy, due to, for example, a lack of print exposure or motivation (Siegel & Himel, 1998; Siegel, 1999). The Matthew effect might be another cause of the decline in IQ for the RD (Stanovich, 1986; Gundersen & Siegel, 2001). Even non-verbal IQ tasks seem to be dependent on proficiency in language (Oller Jr, 1997). However, when we controlled for verbal and non-verbal ability a fairly high correlation still remained between non-word-reading in grades 3 and 12.

CONCLUSIONS

The main finding in the current study is that phonological deficiencies, measured by a non-word-reading test, are a strong indicator of dyslexia. Many participants with a low performance on non-word-reading in grade 3 also did poorly on the equivalent tests in grade 12, even if the subjects had reached the average level in word recognition. Furthermore, the dyslexia definition that includes only phonological skill was the most consistent, that is, most stable over time compared to the word recognition definition or those definitions that include IQ

variables. If we assume that the measure of phonological skills is a short cut towards a sharper definition of dyslexia it would make the diagnosis effective and by that the remediation more specific. However, research is needed both on the biological, the cognitive and the manifest levels to find a clearer definition of dyslexia.

REFERENCES

Borstrøm, I., & Elbro, C. (1997). Prevention of dyslexia in kindergarten: Effects of phonological training. In C. Hulme, & M. Snowling (Eds.), *Dyslexia*: *Biology, cognition*, *and intervention* (pp. 235–253). London: Whurr.

Bruck, M. (1992). Persistence of dyslexics' phonological awareness deficits. *Developmental Psychology, 28*, 874–886.

Bruck, M. (1998). Outcomes of adults with childhood histories of dyslexia. In C. Hulme, & R. M. Joshi (Eds.), *Reading and spelling*: *Development and disorders* (pp. 179–200). Mahwah: Lawerence Erlbaum Associates Publishers.

Castles, A., Datta, H., Gayan, J., & Olson, R. K. (1999). Varieties of developmental reading disorder: Genetic and environmental influences. *Journal of Experimental Child Psychology, 72*, 73–94.

D'Angiulli, A., & Siegel, L. S. (2003). Do children with learning disabilities have distinctive patterns of performance? *Journal of Learning Disabilities*, *36*, 48–58.

DeFries, J. C., Alarcon, M., & Olson, R. K. (1997). Genetic etiologies of reading and spelling difficulties: Developmental differences. In C. Hulme & M. J. Snowling (Eds.), *Dyslexia: Biology, cognition, and intervention* (pp. 20–37). London: Whurr.

Downey, D. M., Snyder, L. E., & Hill, B. (2000). College students with dyslexia: Persistent linguistic deficits and foreign language learning. *Dyslexia*, *6*, 101–111.

Elbro, C. (1998). Reading-listening discrepancy definitions of dyslexia. In P. Reitsma, & L. Verhoeven (Eds.), *Problems and interventions in literacy development* (pp. 129–146). Dordrecht: Kluwer.

Elbro, C., Nielsen, I., & Petersen, D. K. (1994). Dyslexia in adults: Evidence for deficits in non-word reading and the phonological representation of lexical items. *Annals of Dyslexia*, *44*, 205–226.

Fagerheim, T., Raeymaekers P., Tønnessen, F. E., Pedersen, M., Tranebjærg, L., & Lubs, H. A. (1999). A new gene (DYX3) for dyslexia is located on chromosome 2. *Journal of Medical Genetics*, *36*, 664–669.

Fisher, S. E., Francks, C., Marlow, A. J., MacPhie, L., Newbury, D. F., Cardon, L. R., Ishikawa-Brush, Y., Richardson, A. J., Talcott, J. B., Gayán, J., Olson, R. K., Pennington, B. F., Smith, S. D., DeFries, J. C., Stein, J. F., & Monaco, A. P. (2002). Independent genome-wide scans identify a chromosome 18 quantitative-trait locus influencing dyslexia. *Nature Genetics*, *30*, 86–91.

Fletcher, J. M., Espy, K. A., Francis, D. J., Davidson, K. C., Rourke, B. P., & Shaywitz, S. E. (1989). Comparisons of cutoff and regression-based definition of reading disabilities. *Journal of Learning Disabilities*, *22*, 334–338.

Frederickson, N., & Frith, U. (1998). Identifying dyslexia in bilingual children: A phonological approach with inner London Sylheti speakers. *Dyslexia*, *4*, 119–131.

Frith, U. (1999). Paradoxes in the definition of dyslexia. *Dyslexia*, *5*, 192–214.

Grigorenko, E. L. (2001). Developmental dyslexia: An update on genes, brains, and environments. *Journal of Child Psychology and Psychiatry, 42*, 91–125.

Grigorenko, E. L., Wood, F. B., Meyer, M. S., Pauls, J. E. D., Hart, L. A., & Pauls, D. L. (2001). Linkage studies suggest a possible locus for developmental dyslexia on chromosome 1p. *American Journal of Medical Genetics*, *105*, 120–129.

Gundersen, L., & Siegel, L. S. (2001). The evils of the use of IQ tests to define learning disabilities in first and second-language learners. *The Reading Teacher*, *55*, 48–55.

Gustafson, S., & Samuelsson, S. (1999). Intelligence and dyslexia: Implications for diagnosis and intervention. *Scandinavian Journal of Psychology*, *40*, 127–134.

Hatcher, J., Snowling, M. J., & Giffiths, Y. M. (2002). Cognitive assessment of dyslexic students in higher education. *British Journal of Higher Education*, *72*, 119–133.

Høien, T., & Lundberg, I. (2000). *Dyslexia*: *From theory to intervention.* Dordrecht, NL: Kluwer Academic Publishers.

Holmgren, B. (1984). *ITPA Manual* (Illinois Test of Psychlinguistic Abilities). Stockholm: Psykologiförlaget.

Jacobson, C. (1995). Hur utvecklas läsförmågan? Resultat från Ordkedjetestet (How does reading ability develop? Results of the Wordchains test). In C. Jacobson, & I. Lundberg (Eds.), *Läsutveckling och dyslexi* (pp. 196–204). Stockholm: Liber utbildning.

Jacobson, C. (1999). How persistent is reading disability? Individual growth curves in reading. *Dyslexia*, *5*, 78–93.

Jacobson, C., & Lundberg, I. (1995). *Läsutveckling och dyslexi* (Reading development and dyslexia). Stockholm: Liber utbildning.

Jacobson, C., & Svensson, I. (2003). *O-ord* (Pseudo-word). Norms for grades 8–12. Unpublished data.

Johansson, M. G. (1992). *Klassdiagnoser i Läsning och Skrivning för högstadiet och gymnasiet* (Group diagnosis in reading and writing for middle school pupils). Handledning. Stockholm: Psykologiförlaget.

Kjeldsen, A. C., Niemi, P., & Olofsson, Å. (2003). Training phonological awareness in kindergarten level children: Consistency is more important than quantity. *Learning and Instruction*, *13*, 349–365.

Lundberg, I. (1994). Reading difficulties can be predicted and prevented: A scandinavian perspective on phonological awareness and reading. In C. Hulme, & M. Snowling (Eds.), *Reading development and dyslexia* (pp. 180–199). London: Whurr.

Lundberg, I. (1999). Towards a sharper definition of dyslexia. In I. Lundberg, F. E. Tønnesson, & I. Austad (Eds.), *Dyslexia*: *Advances in theory and practice* (pp. 9–29). Dordrecht, The Netherlands: Kluwer Academic Publishers.

Lundberg, I., Frost, J., & Petersen, O. P. (1988). Effects of an extensive program for stimulating phonological awareness in preschool children. *Reading Research Quarterly*, *33*, 263–284.

Lundberg, I., Olofsson, Å., & Wall, S. (1980). Reading and spelling skills in the first school years predicted from phonemic awareness skills in kindergarten. *Scandinavian Journal of Psychology*, *21*, 159–173.

Lyon, G. R. (1995). Toward a definition of dyslexia. *Annals of Dyslexia*, *45*, 3–27.

Lyon, G. R., Shaywitz S. E., & Shaywitz, B. A. (2003). A definition of dyslexia. *Annals of Dyslexia*, *53*, 1–14.

Miller Guron, L. (1999). *Wordchains*: *A word reading test for all ages*. Windsor: NFER-Nelson.

Miller Guron, L., & Lundberg, I. (2003). Identifying dyslexia in multilingual students: Can phonological awareness be identified through the majority language? *Journal of Research in Reading*, 26, 69–82.

Oller, Jr J. W. (1997). Monoglottosis: What's wrong with the idea of IQ meritocracy and its racy cousins? *Applied Linguistics*, *18*, 467–507.

Olofsson, Å. (1994). *Ordavkodning*: *Mätning av fonologisk och ortografisk ordavkodningsförmåga* (Word decoding, phonological and orthographic decoding ability). Östersund: Läspedagogiskt Centrum.

Olofsson, Å. (2002). Twenty years of phonological deficits: Lundberg's sample revisited. In E. Hjelmquist, & C. von Euler (Eds.), *Dyslexia & literacy* (pp. 151–162). London: Whurr.

Pennington, B., Van Orden, G. C., Smith, S. D., Green, P. A., & Haith, M. M. (1990). Phonological processing skills and deficits in adult dyslexia. *Child Development*, *61*, 1753–1778.

Petryshen T. L., Kaplan, B. J., Liu, M. F., Schmill de French, N., Tibias, R., Hughes, M. L., & Field, L. L. (2001). Evidence for a susceptibility locus on chromosome 6q influencing phonological coding dyslexia. *American Journal of Medical Genetics*, *105*, 507–517.

Rack, J. P., Snowling, M. J., & Olson R. K. (1992). The nonword reading deficit in developmental dyslexia: A review. *Reading Research Quarterly*, *27*, 29–53.

Raskind, W. H. (2001). Current understanding of the genetic basis of reading and spelling disabilities. *Learning Disability Quarterly*, *24*, 141–157.

Raven, J. C. (1965). *Guide to using coloured progressive matrices, set A,Ab,B*. London: H K Lewis Co. Ltd.

Raven, J. C. (1995). *Standard progressive matrices sets A–E*. Oxford: Oxford Psychologists Press Ltd.

Samuelsson, S, & Lundberg, I. (2003). The impact of environmental factors on components of reading and dyslexia. *Annals of Dyslexia*, *46*, 201–217.

Schneider, W., Küspert, P., Roth, E., Visé, E., & Marx, H. (1997). Short- and long-term effects of training phonological awareness in kindergarten: Evidence from two German studies. *Journal of Experimental Child Psychology, 66*, 311–340.

Shaywitz, S. E., Fletcher, J. M., Holahan, J. M., Shneider, A. E., Marchione, K. E., Stuebing, K. K., Francis, D. J., Pugh, K. R., & Shaywitz, B. A. (1999). Persistence of dyslexia: The Connecticut longitudinal study at adolescence. *Pediatrics*, *104*, 1351–1359.

Siegel, L. (1989). Why we do not need intelligence tests score in the definition and analyses of learning disabilities. *Journal of Learning Disabilities*, *22*, 514–518.

Siegel, L. (1992). An evaluation of the discrepancy definition of dyslexia. *Journal of Learning Disabilities*, *25*, 618–629.

Siegel, L. (1999). Issues in the definition and diagnosis of learning disabilities: A perspective on Guckenberger v. Boston University. *Journal of Learning Disabilities*, *32*, 304–319.

Siegel, L., & Himel, N. (1998). Socioeconomic status, age and the classification of dyslexics and poor readers: The dangers of using IQ scores in the definition of reading disabilities. *Dyslexia*, *4*, 90–104.

Snowling, M. J. (1998). Dyslexia as a phonological deficit: Evidence and implications. *Child Psychology & Psychiatry Review, 3*, 4–11.

Snowling, M. J. (2001). From language to reading and dyslexia. *Dyslexia*, *7*, 37–46.

Snowling, M. J., Nation, K., Moxham, P., Gallagher, A., & Frith, U. (1997). Phonological processing skills of dyslexic students in higher education: A preliminary report. *Journal of Research in Reading*, *20*, 31–41.

Sögaard, A. & Bording Petersen, S. (1974). *OS 400, ordstillelaesnings-pröve* (Test of silent reading of words. Swedish translation by Margit Tornéus). Copenhagen: Dansla Psykologisk Forlag.

Sprugevica, I., & Høien, T. L. (2003). Early phonological skills as a predictor of reading acquisition: A follow-up study from kindergarten to middle of grade 2. *Reading and Writing*: *An Interdisciplinary Journal*, *16*, 159–177.

Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly, 21*, 360–407.

Stanovich, K. E. (1988). Explaining the differences between the dyslexic and the garden-variety poor reader: The phonological-core variable-difference model. *Journal of Learning Disabilities*, *21*, 590–612.

Stanovich, K. E. (1994). Annotation: Does dyslexia exist? *Journal of Child Psychology and Psychiatry, 35*, 579–595.

Stanovich, K. E. (1998). Refining the phonological core deficit model. *Child Psychology and Psychiatry Review, 3*, 17–21.

Stanovich, K. E., & Siegel, L. (1994). Phenotypic performance profile of children with reading disabilities: A regression-based test of the phonological-core variable-difference model. *Journal of Educational Psychology, 86*, 24–53.

van der Leij, A., de Jong, P. F., & Rijswijk-Prins, H. (2001). Characteristics of dyslexia in a Dutch family. *Dyslexia*, *7*, 105–124.

Vellutino, F. R., & Scanlon, D. M. (1987). Phonological coding, phonological awareness, and reading ability: Evidence from a longitudinal and experimental study. *Merrill-Palmer Quarterly, 33*, 321–363.

Wagner, R. K., & Torgesen, J. K. (1987). The nature of phonological processing and its causal role in the acquisition of reading skills. *Psychological Bulletin*, *101*, 192–212.

Westerlund, A., & Ullstadius, E. (1991). *Prov för mätning av allmänna färdigheter* (*Test for the measurement of general abilities*). Resultat från prövningen inom den nationella utvärderiingen av grundskolan. Rapporter från Institutionen från Pedagogik, Göteborgs Universitet, 1991:02.

Wimmer, H. (1996). The nonword reading deficit in developmental dyslexia: Evidence from children learning to read German. *Journal of Experimental Child Psychology, 61*, 80–90.

