The Nonword-Reading Deficit of Disabled Readers:
A Developmental Interpretation

Wim Van den Broeck
Vrije Universiteit Brussel

Astrid Geudens
Lessius University College (Catholic University Leuven)

Kees P. van den Bos
University of Groningen

This article presents empirical evidence challenging the received wisdom that a nonword-reading deficit is a characteristic trait of disabled readers. On the basis of 2 large-scale empirical studies using the reading-level match design, we argue that a nonword-reading deficit is the consequence of normal developmental differences in word-specific knowledge between disabled readers and younger normal readers (both groups being matched on real-word reading). The first study shows that the nonword-reading deficit varies as a function of age and reading level and that this deficit is not typical for disabled readers. The second study demonstrates that a nonword-reading deficit crucially depends on the sensitivity of the matching word reading task to detect age-related differences in word-specific knowledge between disabled and normal readers. We clarify how these findings can be interpreted within the current framework of the phonological deficit hypothesis and discuss implications for theories of reading development.

Keywords: disabled readers, nonword-reading deficit, phonological deficit, reading development, reading-level match

The failure of some children to develop age-appropriate reading ability has been a motivation to study the underlying determinants of their reading problems. According to the currently dominant phonological deficit hypothesis of developmental dyslexia, disabled readers suffer from a specific deficit in phonological skills (for reviews, see Stanovich & Siegel, 1994; Vellutino, Fletcher, Snowling, & Scanlon, 2004; Ziegler & Goswami, 2005). The core of this hypothesis is that those aspects of reading that place heavy demands on phonological processes pose comparatively increasing difficulties for disabled readers. Probably the most frequently investigated empirical effect corroborating the phonological deficit hypothesis is the demonstration that disabled readers have a specific problem in reading nonwords (e.g., *fap, rilt, zetlop*). The idea is that nonwords can only be read by converting graphemes or grapheme clusters into their phonological counterparts. Thus, a relative disability in reading nonwords would be a clear indication of a phonological disturbance. In this article, we suggest that the nonword-reading deficit is not a characteristic trait of reading disability but rather the consequence of normal developmental differences. Before presenting our findings, we discuss the logic and the intricacies of the reading-level match design and explore the mixed empirical findings pertaining to the nonword-reading deficit.

Interpretations of a Nonword-Reading Deficit Within the Context of the Reading-Level Match Design

Researchers investigating the nonword-reading deficit have usually relied on the reading-level match (RLM) methodology (Backman, Mamen, & Ferguson, 1984; Bryant & Goswami, 1986; Jackson & Butterfield, 1989; Vellutino & Scanlon, 1989). In an RLM design, individuals with reading disabilities are matched with younger typical readers on a measure of reading ability (e.g., a real-word reading test). Subsequently, they are compared on their nonword reading. The RLM design is often preferred to the traditional chronological-age match design because many information processing differences between disabled readers and chronological-age-matched typical readers could also be the result of both groups' different reading experiences. For instance, when disabled readers are compared with normal same-age readers on a phoneme isolation task, it remains unclear whether the observed phoneme isolation disability in the poor reader group is cause or consequence of their reading disability (cf. Geudens, Sandra, & Van den Broeck, 2004). Given the experimental control on reading level, the RLM design is generally considered a more selective device in isolating critical processing differences between disabled and typical readers (cf. Stanovich & Siegel, 1994).

There are at least two ways of interpreting an observed nonword-reading deficit. First, according to the so-called quasi-experimental account of the RLM design (cf. van IJzendoorn & Bus, 1994; Ziegler & Goswami, 2005), a nonword-reading deficit is strong evidence for the existence of a phonological deficit and hence provides a causal explanation of reading disability.

According to a more descriptive interpretation, a nonword-reading deficit indicates that the older disabled readers have reached the same real-word reading level as the younger typical readers in a qualitatively different manner (cf. Olson, Wise, Conners, & Rack, 1990; Stanovich & Siegel, 1994). The absence of a nonword-reading deficit, on the other hand, indicates that both groups of readers are going through the same sequence of stages at a slower rate (suggesting a developmental lag or delay). However, just as with the quasi-experimental interpretation, the use of the design as a means to describe differential cognitive profiles of children with or without reading disabilities also implies that a relative weakness of nonword reading is a characteristic of reading disability.

The way poor readers differ in their reading processes from typical readers has been framed most often within the classic dual-route model of word recognition (Coltheart, 1978; Humphreys & Evett, 1985; for more recent versions of the dual-route model, see Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001; Perry, Ziegler, & Zorzi, 2007). According to the original dual-route model, the activation of a lexical representation (the identification of a written word) is the result of two independent processes. The primary process of direct access is thought to establish a direct association between an orthographic representation and a specific lexical entry, which contains information about the meaning and sound of a word. The secondary process transforms the orthographic representation of a word into a phonological code, using grapheme–phoneme correspondence rules. The assembled phonological code is subsequently mapped onto a lexical entry. Thus, word recognition ability is viewed as the sum of two subcomponents: a direct orthographic process and a phonologically mediated process. Consequently, because word recognition ability is held constant in an RLM study, the observation that the disabled readers perform poorly in the nonword-reading task is interpreted as direct evidence for a deficit in the phonological component of word recognition.

In addition, an orthographic compensation process is a necessary corollary to explain why the disabled readers have reached the same reading ability level as the control subjects despite their weaker phonological processing (Rack, Snowling, & Olson, 1992). This orthographic compensation process is, however, open to interpretation (cf. Stanovich & Siegel, 1994). First, it could be the result of an inherent processing ability that is less impaired in reading-disabled children. Secondly, it could be a strategic choice attempting to compensate for the phonological deficit. Finally, disabled readers could have attained the same word-reading level as their younger nondisabled controls because they have had more exposure to print (Stanovich & Cunningham, 1993; Stanovich & West, 1989).

Before continuing our discussion on the nonword-reading deficit, we briefly focus on different theories of reading development as these insights offer another approach to interpret the nonword reading deficit. Within the framework of the dual-route model, the most common interpretation is that beginning readers move gradually from the indirect phonological route to the direct lexical route (Barron, 1986).

An alternative idea is that beginning readers can store information about whole word patterns by associating distinctive graphic cues with the corresponding lexical entries, a process named "sight word learning" (Ehri, 1998) or "logographic reading" (Frith, 1985; Seymour & Elder, 1986). After gathering sufficient word-specific experiences, they progressively and implicitly abstract some rule-based knowledge. On this account, a phonological insufficiency would impair the abstraction of rule-based knowledge and consequently result in a nonword-reading deficit.

The idea that the underlying statistical structure of the orthography of a given language is implicitly extracted on the basis of learned exemplars has been bolstered by connectionist models of reading (for a recent review, see Perry, Ziegler, & Zorzi, 2007). In attempting to model reading development and reading disability (e.g., Harm & Seidenberg, 1999), most connectionist models have converged on the idea that both item-based knowledge and rule-based knowledge (explicit and implicit) play an important role in reading development (cf. Hutzler, Ziegler, Perry, Wimmer, & Zorzi, 2004).

Variable Findings of a Nonword-Reading Deficit

In a review article of 16 studies carried out in English-speaking countries, Rack et al. (1992) found evidence for a nonword-reading deficit in disabled readers in 10 out of the16 reviewed studies, whereas six studies did not report such effect. What interests us most is how these variable effects can be accounted for. One possibility is that these null findings are the result of sampling variability—that is, they could have been Type II errors. When the real effect is modest, the chance of observing Type II errors increases. In a meta-analysis of all studies reviewed by Rack et al. (1992), van IJzendoorn and Bus (1994) found that only 6% of the variance in nonword reading was explained by group membership (reading disabled or normal). Rack et al. attributed the null findings to specific characteristics of the matching task and the nonword-reading test, to the age of the control subjects, to strategic effects, and to the teaching experiences of the reading disabled. Although Rack et al. considered the positive studies in their review as paradigmatic cases (real effects) and the null findings as methodological departures, their interpretation can be viewed as the first attempt to make a systematic study of the conditions in which a nonword-reading deficit can be observed and the conditions in which no effects are to be expected.

Since both review articles have been published, the number of RLM studies has doubled (for a recent meta-analysis, see Herrmann, Matyas, & Pratt, 2006). Again, the majority of studies have reported a nonword-reading deficit, yet in a number of studies null effects have been observed (cf. Ellis, McDougall, & Monk, 1996; Foorman, Francis, Fletcher, & Lynn, 1996; Snowling, Goulandris, & Defty, 1996; Vellutino, Scanlon, & Tanzman, 1994).

Our interest in an analysis of the factors determining a nonword-reading deficit was further stimulated by some demonstrations of nonword-reading deficit in consistent orthographies like German, Dutch, and Spanish (Jiménez & Hernandez, 2000; Landerl, Wimmer, & Frith, 1997; van der Leij & van Daal, 1999; van der Leij, van Daal, & De Jong, 2002; Wimmer, 1996). These studies have

shown that the nonword-reading deficit is not less prevalent in orthographies with rather consistent relationships between spelling and phonology. In languages like Greek, German, Dutch, Spanish, or Italian, graphemes are mostly pronounced in the same way; whereas in languages with inconsistent orthographies like English or Danish, graphemes often have multiple pronunciations (e.g., the grapheme *a* in *man*, *war*, *take*, *about*, and *bar*). On the basis of these cross-linguistic effects, Ziegler and Goswami (2005, 2006) have formulated their currently influential psycholinguistic grain size (PGS) theory. According to PGS theory, the phonological representations, which are activated in the reading process, are the reflection of the grain size that is required in each language to attain orthographic–phonological consistency. In English, larger orthographic units such as rimes are generally more consistent than single graphemes (e.g., the grapheme *a* in *–ave* is pronounced more consistently, as in *gave*, *save*, and *rave*). Therefore, Ziegler and Goswami proposed that reading in consistent orthographies involves small linguistic units, whereas reading in inconsistent orthographies requires the use of larger units as well.

The demonstration of a nonword-reading deficit in consistent orthographies is remarkable. If reading words and nonwords in a consistent orthography relies on the same small linguistic units (grapheme–phoneme correspondences), as hypothesized by PGS theory, then no—or at least a less sizeable—nonword-reading deficit should be expected. However, in a direct comparison of developmental dyslexia in English and German, Ziegler, Perry, Ma-Wyatt, Ladner, and Schulte-Körne (2003) demonstrated nonword-reading deficits of equal magnitude (see also Paulesu et al., 2001).

Another source of variability in a nonword-reading deficit is the supposed existence of subtypes of reading disability. Although individuals with a reading disability as a group are generally characterized as having a phonological deficit, some people with "surface developmental dyslexia" have been reported to experience specific problems in the processing of orthographic information without any phonological impairments (cf. Castles & Coltheart, 1996; Coltheart, Masterson, Byng, Prior, & Riddoch, 1983). Castles and Coltheart (1993), who studied a sample of 53 dyslexic children, identified eight (15%) children with a phonological subtype (their nonword reading fell outside the normal range, and their exception-word reading was within the normal range), and 10 (19%) children with a surface subtype (who showed the opposite pattern). The rest of the children showed a mixed profile.

These findings were, however, criticized on methodological grounds. It has been argued (Snowling, Bryant, & Hulme, 1996; Stanovich, Siegel, & Gottardo, 1997; Ziegler & Goswami, 2005) that comparing disabled readers to a normative sample of children of the same age, as in Castles and Coltheart (1993), may yield processing trade-offs between sublexical and lexical paths that depend on the overall level of word-reading skill. To this end, RLM controls were used in a new series of subtyping studies (Manis, Seidenberg, Doi, McBride-Chang, & Petersen, 1996; Murphy & Pollatsek, 1994; Stanovich et al., 1997). Generally, once reading age was taken into account, 20% to 25% of the children with reading disability were identified as phonological dyslexics, whereas the surface subtype almost disappeared (1% or 2%). The dominant interpretation of these results is that surface dyslexia arises from a milder form of the phonological deficit leading to a delayed pattern of reading development, whereas phonological

dyslexia represents a more severe reading disorder (Griffiths & Snowling, 2002; Stanovich et al., 1997; but see Jackson & Coltheart, 2001, for an alternative view). Thus, although there is an unquestionable awareness in the literature that the expression of reading disability is a diverse matter, the hypothesis of a phonological deficit as a common underlying disorder together with the prediction of a nonword-reading deficit remains uncontested.

Methodological Pitfalls of the RLM Design

Despite the popularity of the RLM design and its quasi-experimental interpretation, this design is not without methodological problems and interpretational ambiguities (Jackson & Butterfield, 1989). Goswami and Bryant (1989) claimed that inferior performance of disabled readers on a nonword-reading task, using an RLM design, is a strong indication of a causal link between this variable and reading development. However, according to these authors, a null result in an RLM study (no difference on nonword reading between disabled and normal readers) would be impossible to interpret, because the superior cognitive and metacognitive skills of the older group of disabled readers might have masked their real processing deficit. Goswami and Bryant's analysis revealed the essential weakness of the RLM design. Although the reading disabled and normal reading groups are matched on reading level, they are inevitably unmatched on age. Consequently, both groups can differ on a multitude of variables relating to age, which could actually produce a difference on nonword reading. In the light of this analysis, the confound of diagnostic category and age in an RLM design prevents the demonstration of the causal status of the investigated processing variable (nonword reading). Given this confound, the interpretation of the nonword-reading deficit in terms of age-related variables becomes a likely candidate to explain the observed difference between disabled and reading-level-matched typical readers (cf. Jackson & Butterfield, 1989). To disentangle potential confounding variables, Backman et al. (1984) have advocated using a dual comparison, between groups matched on reading level and between groups matched on age. However, the juxtaposition of two different confounds, one between reading group and age and one between reading group and reading level, is not a desirable solution, because the first confound does not compensate for the second and vice versa.

An Alternative Explanation of the Nonword-Reading Deficit

The aim of this article is to investigate the thus far unexplored possibility that a nonword-reading deficit is the result of normal developmental processes. Specifically, we hypothesize that the group of older disabled readers reaches the same reading score in the matching task as the younger normal readers as a result of their superior experience with the words to be read in the matching task, simply because they are older. We suggest that despite their poor reading level, older disabled readers, who are often exposed to the same reading material as their normal-reading age mates, have a better grasp of the semantics, whole-word phonology, and orthography of the words that are typical for the vocabulary of their age,

compared with the younger normal readers.[1] Alternatively, younger normal readers who are matched on the word-reading task cannot depend upon the same word-specific knowledge as the older disabled readers do. Following this reasoning, the same score on the word-reading test (due to the matching procedure) is partly the result of qualitatively different knowledge. The older disabled readers may be able to exploit their better word-specific knowledge. The younger readers, however, are selected in such a way that they match the word-reading scores of the disabled readers. They may not have had the chance yet to attain the same level of word-specific knowledge as the disabled readers, but on the other hand they need to rely on phonological decoding skills to read the real words. This interpretation entails that the differences in decoding ability between both reading groups, as exemplified in the nonword-reading test, are not necessarily the expression of an intrinsic quality of both groups but may well be provoked by the age-confounded design.

This interpretation of the nonword-reading deficit, focusing on age-dependent experiences, may be questioned. One could, on the contrary, argue that disabled readers read less than normal readers do and as such have less experience with orthographically and semantically more complex texts, leading to a lower instead of a higher level of word-specific knowledge in comparison with matched normal readers. Following this reasoning, disabled readers' word-reading level in the matching task is not achieved on the basis of superior word-specific knowledge of the items but in contrast by a superior experience with simple phonological recoding. Given their age and their more extensive experience with simple recoding to sound, a coding advantage is achieved. The problem with this alternative hypothesis, however, is that the basic idea (i.e., a phonological decoding advantage for disabled readers) is in contradiction with the majority of studies that report a nonword-reading deficit. We need more direct evidence about the role of word-specific knowledge in disabled and normal reading to disentangle these alternative interpretations. This is one of the aims of the current study.

It is important to note that our interpretation of the nonword-reading deficit in terms of normal developmental processes does not imply that the orthographic and semantic knowledge of the disabled readers would be at a normal level. On the contrary, we acknowledge that the disabled readers' orthographic knowledge is seriously impaired compared with normal readers of the same age and that their semantic knowledge is probably slightly subnormal (cf. Gallagher, Frith, & Snowling, 2000; Snowling, Gallagher, & Frith, 2003). The point, however, is that the disabled readers' larger orthographic and semantic knowledge compared with normal readers of the same reading level may well be the result of a larger amount of exposure (orthographically and semantically) to specific words, which is typical for the age of the older disabled readers.

Taking into account the age-related differences between the two groups as an explanatory variable is not new. As we have discussed before, previous studies have interpreted null findings as well as superior performance on orthographic tasks as the result of age-related experiences. What is new in our proposal is the recognition of the possibility that the inferior performance of the disabled readers on reading nonwords may be the result of the age differences of both groups, which entail qualitatively different experiences. It is important to note that our developmental interpretation is also clearly different from the aforementioned descriptive interpretation of the findings in an RLM design. According to the descriptive account, disabled readers' familiarity with the words in the matching task compensates for lower decoding ability in reading real words; however, the poor performance on the nonword-reading test indicates a genuine decoding deficit. On the contrary, our proposal implies that the matching procedure automatically selects a group of younger normal readers who have better decoding abilities, not by nature but by design. One could say that the matching procedure, which partly determines the reading ability components of the younger control group, is overrated by the quasi-experimental interpretation as a causal design and underrated by the descriptive interpretation as a kind of pure observational design.

If our hypothesis—that a nonword-reading deficit is the consequence of normal developmental differences in word-specific knowledge—were correct, then a nonword-reading deficit should not be restricted to disabled readers only. On the basis of this hypothesis we predict that average readers would also show a nonword-reading deficit when compared with excellent younger readers of the same reading level, at least when the word-reading test is sensitive to their different word-specific knowledge.

The goal of the two experiments that are presented in the following sections is to investigate the role of age-related word-specific knowledge in the determination of the nonword-reading deficit and more generally to explore the conditions in which a nonword-reading deficit can be detected and the conditions in which no effect shows up. In the first experiment, we examine whether a nonword-reading deficit varies as a function of age and reading level by measuring the nonword-reading ability of disabled and normal readers at a wide range of reading levels and ages. If the difference in word-specific knowledge between the reading-level-matched groups is a major determinant of the nonword-reading deficit, this deficit will change as a function of age and reading level. Moreover, on the basis of our developmental interpretation of the nonword-reading deficit, we make the unique prediction that average readers will show a nonword-reading deficit when compared with excellent younger readers of the same reading level. In the second study we further examine the role of word-specific knowledge by varying the required word-specific knowledge of the word-reading task.

Experiment 1

Method

Design and analytic strategy. In this study, we used an improved continuous regression-based approach of the RLM design instead of the classic pairwise matching procedure to compare (a) the performance of disabled and younger typical readers and (b) the performance of average and excellent younger readers on a nonword-reading task. After performing a word-reading and a nonword-reading test, we statistically matched both groups on the

---

[1] We refer to knowledge at the word level of the specific meaning, sound, and orthography of words as *word-specific knowledge*. Although our use of this term encompasses the orthographic knowledge at the word level, it is clearly more comprehensive.

word-reading test, using a regression analysis. This regression methodology was proposed by Jackson and Butterfield (1989) and first used in a study by Stanovich and Siegel (1994). This method is preferred because of its statistical advantages compared to the classic matching procedure. Yet, it doesn't provide a solution to the problem of unrepresentative sampling from extreme groups, causing bias from regression effects. When the reliability of the selection variable is not perfect, regression effects arise, because under the null hypothesis (no difference in nonword reading) the expected mean difference on the dependent variable (nonword reading) deviates from zero if the matching variable (word reading) and the selection variable are not the same.

In many nonword-reading deficit studies the selection of disabled readers is based on unverifiable clinical criteria. Despite being matched on the matching variable, participants regress toward the means of their original populations, being defined by the selection variable. Consequently, the effect of the independent variable (being reading disabled or not) cannot be estimated independently from the regression effect. However, when the matching variable is the same as the selection variable (a specific word-reading test), either using the classic matching technique or the regression procedure, the expected mean difference on the nonword-reading test is zero when the null hypothesis is true, so no regression artifact emerges. The logic behind this procedure is that in the case of an identical selection and matching variable, the biasing effect of the selection variable becomes neutralized by the experimental (as in the classic RLM design) or statistical (as in the regression procedure) match on the same variable (cf. Huitema, 1980). Thus, although regression artifacts are not typical for one or the other design, they are less frequent in a regression-based RLM design, because in that case the same word-recognition test is often used, both as the selection variable and as the statistical matching variable. Furthermore, a classic RLM design often suffers from logistical problems in finding enough matched subjects. Moreover, the statistical power of the regression-based design is superior because it allows data from many typical participants to be entered in the analysis. In the regression-based design, participants are sampled from a large continuous population varying in age and ability.

The analysis based on the regression-methodology proceeds as follows. In the first step of the hierarchical regression-analysis, the dependent variable (nonword reading) is regressed onto the control variable (word reading) and all its significant higher order trends. After thus partialing out the variance due to word-reading ability, a variable designating group membership (reading disabled vs. normal) is entered into the equation. So, a significant regression coefficient (beta) of the vector coding group membership reflects a nonword-reading deficit (or advantage, depending on the sign of beta). Because we are interested not only in an eventual overall nonword-reading deficit but also in detecting variability in this deficit as a function of reading level and grade, we also introduced interaction vectors into our analysis. An interaction of group by reading level, as revealed by nonparallel regression lines for both groups, points to variation in the nonword-reading deficit as a function of reading level. In the presence of heterogeneous regression for normal and reading-disabled subjects, it is possible to identify ranges of reading ability in which both groups differ in nonword-reading ability and ranges in which both groups are similar (the Johnson-Neyman technique). If normal and disabled readers are to be equated on reading level for the purpose of comparing profiles of phonological processing, a common metric

of reading performance regardless of age or grade is required (see also Foorman et al., 1996, who recommend Rash scores). In our analyses we employed simple raw scores (number of words read correctly), because they provide a good approximation of the equal-interval measurement requirement. A potential difficulty of the regression technique when comparing groups of poor and good readers is the existence of partially nonoverlapping distributions. In this case, participants in one group are compared with statistically extrapolated but nonexistent participants in another group, resulting in meaningless conclusions (cf. Foorman et al., 1996). For example, when no disabled readers are found in the upper tail of the distribution of word-reading scores, the normal readers falling in that region are compared with statistically predicted scores based on the performance of disabled readers of lower reading ability. To avoid this problem, the statistical analyses are exclusively based on the overlapping part of the distributions.

Participants. In the analysis, the data from children ($n = 1,418$) who participated in a previously unpublished norm study were amalgamated with the data from participants in some recent unpublished studies ($n = 660$). All children from Grades 1 to 6 ($n = 1,520$) attended regular elementary schools, mostly located in northern and eastern parts of the Netherlands. Furthermore, 302 thirteen-year-old students in Grade 7 (in the Netherlands, Grade 7 is equivalent to the first year of secondary education), 174 sixteen-year-old students, and 82 adults were also administered a word-reading and a nonword-reading test. The total sample thus consisted of 2,078 participants. The schools were located in middle-class to upper-middle-class neighborhoods with mainly Caucasian families whose first language was Dutch. Participants were recruited for testing from entire classes.

The participants were classified into one of two groups. The disabled readers scored on the widely employed Dutch One Minute Test (OMT; Brus & Voeten, 1973) $\leq$ the 10th percentile compared with their age group. This criterion is much stricter than the criteria used in many published nonword-reading deficit studies, in which the 20th or 25th percentile represents the upper limit. None of the individuals participating in this study were identified as having language, emotional, or behavioral disorders. The percentile scores of the nondisabled readers on the OMT were $\geq$ 30. This procedure resulted in 211 disabled readers and 1,438 nondisabled readers. For each of the two groups at each of the nine grade/age levels, Table 1 displays the means and standard deviations on the OMT, for the nonword-reading test, and for age in months. The table shows that the disabled readers are seriously retarded in comparison with the normal readers: The retardation varies from a delay of at least 1 year for Grade 2 disabled readers to more than 4 years for Grade 7 and older disabled readers. For instance, adult disabled readers are reading at a level normal for Grade 3–4 children. Furthermore, it is clear from the table that the chronological age of the disabled readers was very similar to the age of the nondisabled readers for each grade. The classic discrepancy criterion of developmental dyslexia (reading level significantly below IQ level) was not considered, but it should be noted that due to recent developments in the research on dyslexia, this criterion has been largely abandoned (cf. Lyon, Shaywitz, & Shaywitz, 2003).

Measures.

Word identification. All participants were administered the OMT. The test consists of 116 real words (nouns, verbs, adjectives, etc.) that are printed in four columns. These words are ordered from lower to higher degree of reading difficulty. Most

Table 1
Mean Word-Reading Scores (OMT), Nonword-Reading Scores (Klepel) and Age as a Function of Subject Classification and Grade

	Nondisabled readers			Disabled readers		
Variable	M	SD	n	M	SD	n
OMT raw score						
Grade 1	25.8	10.0	196	7.9	1.3	29
Grade 2	49.2	12.0	222	19.7	8.0	31
Grade 3	65.7	7.9	114	34.4	5.0	16
Grade 4	72.3	8.6	188	45.5	6.9	28
Grade 5	79.6	9.3	160	49.9	6.4	25
Grade 6	86.1	8.5	166	59.0	4.4	27
Grade 7	88.1	9.3	214	60.0	5.5	28
Secondary	97.2	7.8	120	66.1	7.8	19
Adult	105.4	7.2	58	68.7	4.8	8
Klepel raw score						
Grade 1	23.0	10.1	196	9.2	2.6	29
Grade 2	43.7	14.4	222	16.2	4.6	31
Grade 3	52.1	11.9	114	23.6	6.9	16
Grade 4	62.4	13.0	188	33.9	10.7	28
Grade 5	66.1	13.4	160	32.8	11.2	25
Grade 6	74.3	14.1	166	45.7	14.5	27
Grade 7	76.8	15.8	214	49.8	13.8	28
Secondary	87.0	12.8	120	47.4	20.7	19
Adult	94.0	11.6	58	58.0	14.1	8
Age (in months)						
Grade 1	85.8	4.6	196	87.4	5.0	29
Grade 2	95.6	5.0	222	95.8	5.4	31
Grade 3	110.4	4.8	114	114.8	5.9	16
Grade 4	120.8	6.5	188	122.8	6.8	28
Grade 5	134.5	5.8	160	140.3	6.6	25
Grade 6	146.0	4.9	166	147.5	6.0	27
Grade 7	157.7	6.5	214	160.7	7.4	28
Secondary	190.9	7.4	120	195.8	9.7	19
Adult	547.1	52.8	58	570.0	112.8	8

Note. OMT = One Minute Test; Klepel is the nonword reading test.

words (112) are multisyllabic. All words in the test are typically regular words, which is customary in Dutch reading tests. The mean log frequency of the words, based on the Dutch corpus of the CELEX database (Baayen, Piepenbrock, & Gulikers, 1995), is 0.711 (see Table 2).

We also calculated the overall orthographic similarity of the words by means of the recently developed measure—termed *orthographic Levenshtein distance 20* (OLD20; Yarkoni, Balota, & Yap, 2008)—which has been shown to be an improvement on the commonly used Coltheart's *N* (Coltheart, Davelaar, Jonasson, & Besner, 1977).

Table 2
Mean Frequency and Orthographic Similarity of the Items in the Reading Tests of Experiments 1 and 2

Test characteristic	OMT	Klepel	TMT	WRT	Exception words
Log frequency	0.711		0.898	0.600	1.037
OLD20	2.452	3.377	1.900	3.667	2.546

Note. OMT = One Minute Test; Klepel = nonword-reading test; TMT = Two Minute Test; WRT = Word Recognition Test; OLD20 is a measure of overall orthographic similarity. Low OLD20 scores indicate high orthographic similarity and vice versa.

OLD20 is based on Levenshtein distance (LD; Levenshtein, 1966). The LD between two words is the minimum number of substitution, insertion, or deletion operations needed to turn one word into the other. First, LD from each word to every other word in the Dutch corpus of the CELEX database was calculated. Then, to obtain OLD20, the mean LD from a word to its 20 closest neighbors was computed with the Levenshtein Distance Calculator (Yarkoni, 2009). We assume that words of low frequency and low orthographic similarity invoke word-specific knowledge that is typical for higher reading levels. The combination of the rather low frequency and the medium orthographic similarity of the words (OLD20 = 2.45; see Table 2) ensures that this kind of word-specific knowledge will be helpful to read the words of this test.[2] The participants were instructed to correctly read aloud as many words as possible. The raw score is the number of words read correctly after 1 min of reading. In many English language studies, word-reading tests are pure accuracy measures. However, word-reading skill is determined by both accuracy

[2] Although orthographic similarity scores are probably language specific, from Figure 1 in Yarkoni et al. (2008) we can infer that for the length of the words presented in our tests, low OLD20 scores range approximately from 1.0 to 2.0, medium OLD20 scores range from 2.0 to 3.0, and high OLD20 scores exceed 3.0.

and speed (cf. Landerl & Wimmer, 2008; Share, 2008). Therefore, a word-reading fluency test such as the OMT is arguably the better choice. By fixing the allowed time to read (1 min), interpretation difficulties due to accuracy–speed trade-offs are largely avoided (cf. Salthouse & Hedden, 2002). van den Bos, Lutje Spelberg, Scheepstra, and De Vries (1994) reported parallel test reliabilities with a median of .90 over Grades 1 through 7, ranging from .94 in Grade 1 to .87 in Grade 7. Separate conservative estimates of parallel test reliabilities for the 16-year-old students and the adults are .81 and .85, respectively. Thus, the test is highly reliable and has a high discriminative power at all reading levels because it is a combination of a power test and a speed test.

Nonword reading. The administered nonword-reading test (Klepel) was constructed by manipulating the words in the OMT (van den Bos et al., 1994). Pronounceability, word length, and consonant and vowel distributions are preserved in the Klepel, but meaningful and morphemic structures are avoided (see Appendix A for a sample of the words and nonwords). This procedure accomplishes a structural match between word- and nonword-reading tests, a requirement that is usually not fulfilled in RLM studies. One of the factors mentioned by Rack et al. (1992) that could be responsible for not finding a nonword-reading deficit is the similarity of the nonsense words to real words. The overall orthographic similarity of the nonwords to words in the Dutch corpus is low, indicated by a high OLD20 of 3.38 (see Table 2). Of course, as in any nonword-reading test, subjects now and then make an analogy to a real word. However, extensive experience with this test has shown that such instances are rare. The instructions were the same as for the OMT, except that the time limit was 2 min. Again, the raw score is the number of nonwords read correctly. Median parallel test reliability over Grades 1 to 7 is .92, ranging from .93 in Grade 1 to .91 in Grade 7 (van den Bos et al., 1994). Separate conservative estimates of parallel test reliabilities for the 16-year-old students and the adults are .82 and .85, respectively.

Results

The overall correlation through all age levels between word-reading ability and nonword-reading ability was very high (r = .91). The magnitude of this correlation, in comparison with somewhat less extreme correlation coefficients in English studies (e.g., Stanovich & Siegel, 1994, who reported correlations of .73 to .80), is probably the consequence of not only the consistency of the Dutch orthography but also the speed factor, which was common for both tests. The correlations for each grade level are presented in Table 3.

In the regression analysis, the dependent variable (nonword reading) was regressed on word-reading ability (and all significant power polynomials), thus removing all of the variance in nonword reading that is associated with word-reading ability.[3] To test for a quadratic trend in the relationship between word reading and nonword reading, we compared two instances of R^2. The first R^2 includes all nonquadratic terms, whereas the second R^2 also includes the quadratic terms (see Table 4). The difference between the two R^2s indicates the increment in the proportion of variance accounted for by the quadratic terms. This increment was .003, which demonstrates a small but highly reliable quadratic trend, $F(2, 939) = 4.79$, $p = .009$. The same procedure, comparing all

Table 3

Pearson Correlation Coefficients Between OMT and Klepel by Grade Level

Grade level	r	n
1	.89	293
2	.82	322
3	.80	163
4	.78	273
5	.78	229
6	.78	241
7	.67	301
Secondary	.74	174
Adult	.76	82

Note. OMT = One Minute Test; Klepel = nonword-reading test.

terms up to the cubic trend with all terms up to and including the cubic terms, did not reveal a significant cubic trend, $F(2, 937) = 0.48$, $p = .616$.

Next, we examined whether the quadratic regression curves of the disabled and nondisabled readers were parallel. To test this interaction between group and word-reading ability, all singular terms up to and including the quadratic trend were compared with the same terms including the interaction terms. The increment in explained variance due to the group interaction was .003, $F(2, 939) = 5.14$, $p = .006$. This significant interaction indicates that the regression of nonword reading on real-word reading for disabled readers differed from that of nondisabled readers. In other words, an eventual nonword-reading deficit varies in magnitude depending on the level of word-reading ability. Before determining the regions of word-reading level where a nonword-reading difference exists and the regions where no difference between the groups is observed, the overall difference in nonword reading is calculated (see Table 4). The overall nonword-reading deficit amounts to a small 0.9% of explained variance, $F(1, 941) = 30.47$, $p = .000$. Expressed in raw score units, this entails that disabled readers read four nonwords less than nondisabled readers when equated on word-reading level. Finally, using the Johnson-Neyman technique (see Rogosa, 1980, 1981), we determined a simultaneous region of significance. According to Pothoff (1964), this is "a region such that, with confidence ≥ 95 percent (for α = .05) we can state that the two groups are different simultaneously for all points contained in it" (p. 244). Employing Rogosa's (1981, p. 82) formula for two predictors (OMT and OMT^2), the simultaneous region of significance lies within the range of 23 and 67 on the OMT. Thus for all OMT scores between 23 and 67, there is a significant difference between disabled and nondisabled readers in reading nonwords (see Figure 1).

This means that disabled readers reading at the level of Grade 1 (cf. Table 1) do not display a nonword-reading deficit. However, at the levels of Grades 2 to 4, a clear nonword-reading deficit is discernable. Disabled readers reading at still higher levels do not display a nonword-reading deficit. The rising and declining nature of the nonword-reading deficit is obvious from the overall effect calculated for several raw score regions of the word-reading test:

---

[3] We followed the specific analytic procedure for curvilinear regression analysis that was recommended by Pedhazur (1982).

Table 4
Summary of Hierarchical Regression Analysis for Variables Predicting Nonword Reading (N = 1,325)

Variable	R^2	ΔR^2	F_{change}	p
Step 1: Quadratic trend				
Block 1: OMT, Group, Group × OMT	.735			
Block 2: OMT, Group, Group × OMT, OMT^2, Group × OMT^2	.738	.003	4.79	.009
Step 2: Cubic trend				
Block 1: OMT, Group, Group × OMT, OMT^2, Group × OMT^2	.738			
Block 2: OMT, Group, Group × OMT, OMT^2, Group × OMT^2, OMT^3, Group × OMT^3	.738	.000	0.48	.616
Step 3: Interaction				
Block 1: OMT, Group, OMT^2	.735			
Block 2: OMT, Group, OMT^2, Group × OMT, Group × OMT^2	.738	.003	5.14	.006
Step 4: Overall effect				
Block 1: OMT, OMT^2	.726			
Block 2: OMT, OMT^2, Group	.735	.009	30.47	.000

Note. The 95% confidence interval for the overall group effect of ΔR^2 is [.002, .015], calculated using the procedure described by Alf & Graf (1999).

For OMT scores between 0 and 20, ΔR^2 is 0%; for OMT scores between 20 and 40, ΔR^2 is 11.7%; for OMT scores between 40 and 60, ΔR^2 is 3.8%; and for OMT scores larger than 60, ΔR^2 is 0.7%.

A second goal of the study was to test the hypothesis that a nonword-reading deficit is not restricted to disabled readers only. If true, average readers should also show a nonword-reading deficit when compared to excellent younger readers of the same reading level. As the OMT proved to be sensitive to a broad range

of differences in word-specific knowledge, the available data allowed us to test this second prediction as well. We defined average readers as reading on the OMT between the 40th and 60th percentile ($n = 354$) and younger excellent readers as reading on the OMT above the 80th percentile in their age group ($n = 335$). As in the previous analysis, both groups were statistically matched on the OMT. For these groups, we replicated a quadratic trend: The increment in explained variance was .005, $F(2, 683) = 5.65$, $p =$
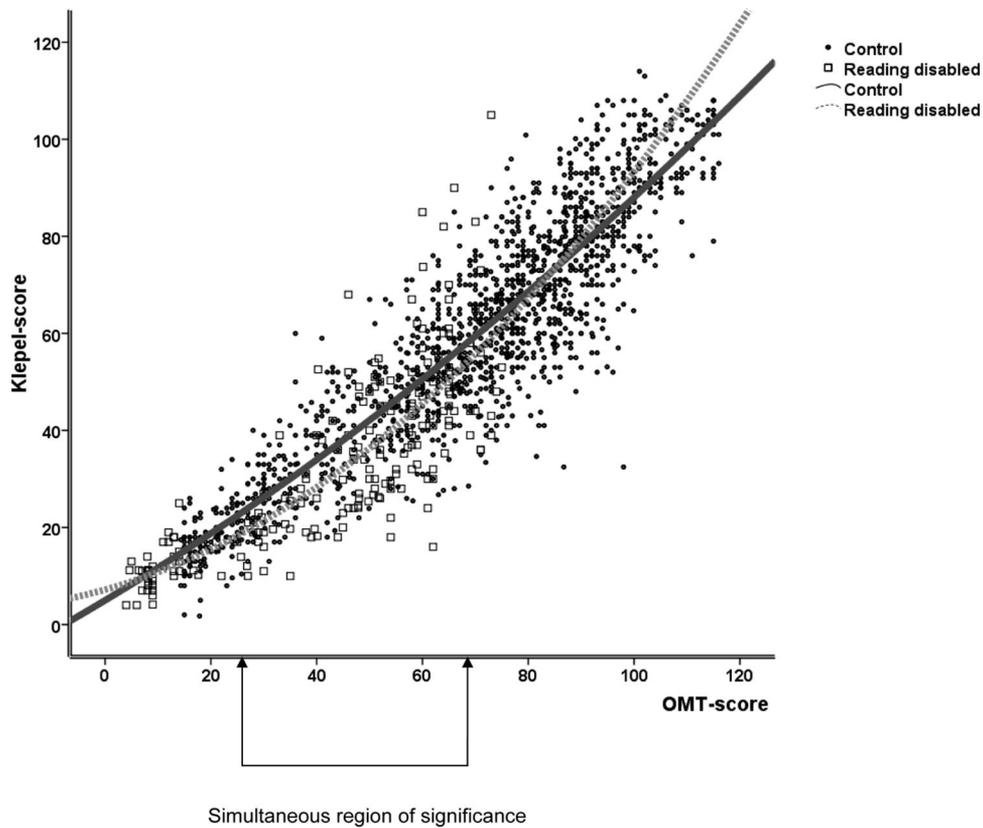

Figure 1. Scatter plot of nonword-reading scores on the Klepel test as a function of word-reading scores on the One Minute Test (OMT) for disabled and nondisabled readers. The simultaneous region of significance indicates the region of OMT scores in which the two groups differ in reading nonwords.

.004. Again, a significant interaction was detected; ΔR^2 is 0.3%, $F(2, 683) = 3.61$, $p = .028$, indicating that the quadratic regression of nonword reading on real-word reading for average readers differed from that of excellent readers. Although the overall reading deficit was clearly nonsignificant, $F(1, 685) = .12$, $p = .73$, we found a substantial nonword-reading deficit for the average readers compared with the excellent younger readers for OMT scores between 45 and 70; $\Delta R^2 = 3.1\%$, $F(1, 180) = 7.40$, $p = .007$.

Discussion

The results essentially show that in a relatively consistent orthography like Dutch, it is possible to observe a large nonword-reading deficit, which in this study was only detectable in a well-specified region of word-reading ability. For the specific tests used in our study, this region corresponds approximately to Grade 2 up to Grade 4. This result is in accordance with the already published demonstrations of a nonword-reading deficit in consistent orthographies. Moreover, the existence of a large nonword-reading deficit for the specified region (in fact, it is almost twice as large as the mean effect found in English studies, viz., 11.7% vs. 6% of explained variance) demonstrates that the exclusive use of regular words in the word-reading task does not diminish the nonword-reading deficit, at least when reading speed is taken into account. The finding that the nonword-reading problem of disabled readers depends on the word-reading ability level is difficult to explain in the light of the hypothesis that a nonword-reading deficit is an essential characteristic of dyslexia. Why would such a deficit be absent for the lower and higher reading levels? Furthermore, an interpretation in terms of floor or ceiling effects can safely be excluded, given the high discriminative power of both tests at all levels of reading ability.

An alternative explanation may be found in the developmental hypothesis that assumes different levels of word-specific knowledge between the matched groups, offering a more parsimonious account than the already discussed ad hoc explanations in terms of the inexperience of the youngest normal readers with the reading of nonwords or the superior cognitive and metacognitive skills of the older disabled readers. If the word recognition test were sensitive to the different degrees of word-specific knowledge of the two groups, a nonword-reading deficit should emerge. On the other hand, if the word-specific knowledge of both groups were under or above the range of word-specific knowledge tapped by the word recognition test, no nonword-reading deficit should be detected. The absence of a nonword-reading deficit for the lower reading levels is comprehensible when we assume that their word-specific knowledge is not yet sufficiently developed to be detected by the specific word-reading task we used. Likewise, the absence of a nonword-reading deficit for the higher reading levels can be understood by assuming that the matched normal readers have gained sufficient word-specific knowledge so that they are not disadvantaged on the word-reading test in comparison with the disabled readers.

Parsimony aside, the developmental hypothesis makes three unique predictions. First, if the older disabled readers are aided by their developmentally higher level of word-specific knowledge when reading the words of the OMT, they should make fewer errors on a subset of high frequency words that is part of the OMT. In an independent study using the OMT as the selection and matching variable and the Klepel as the dependent variable, Riethoven and Hooijmans (2003) replicated a nonword-reading deficit in Dutch disabled readers of Grades 3 and 4 when compared to reading-level-matched younger readers. They found that the disabled readers made statistically fewer errors on 10 high frequency words of the OMT (mean log frequency = 2.62) compared with the normal readers. Second, according to the developmental hypothesis, this higher level of word-specific knowledge in the older disabled readers merely reflects a normal developmental trend. To test this prediction, we performed an additional analysis and regressed the word reading scores of the entire sample on their nonword-reading scores and on the interaction of reading age by nonword reading. The significant interaction effect ($t = -21.26$; $p = .000$; $b = -0.066$; $SE = 0.003$) indicates that the influence of decoding ability on word reading diminishes with reading age. Plotting the simple effects for three distinct reading ages makes clear that with increasing age, word reading is decreasingly determined by decoding ability (see Figure 2). For novice readers, both words and nonwords are practically new. Hence, they have to rely mainly on their word decoding skills. Older readers, however, can also depend upon their growing word-specific knowledge. Furthermore, this implies that the word-reading task (OMT) measures different knowledge components of the word-reading process at different ages.

Third, a nonword-reading deficit also applied to average readers for a rather broad range of reading levels when they were compared with excellent younger readers of the same reading level. This result can unambiguously be attributed to age-related factors, because bias from regression effects was excluded in this study. Hence, we conclude that the nonword-reading deficit was not typical for disabled readers but was determined by the different levels of word-specific knowledge of both groups.

Researchers who are familiar with less extreme correlation coefficients between word-reading and nonword-reading tests in English language studies may wonder how it is possible to obtain a nonword-reading deficit, given the high correlation in this study (.91). It can be proven mathematically (Van den Broeck & Geudens, 2009) that when reading disability or developmental dyslexia is operationally defined as the cutoff score on a continuous variable (as is customary), a nonword-reading deficit is entirely determined by characteristics of the total continuous population and an arbitrary cutoff score. The ultimate implication is that the only empirical question a RLM design is able to answer is how the empirical distributions of the word- and nonword-reading tests change with age. Specifically, the magnitude of the nonword-reading deficit depends entirely on the correlation coefficients between both measures at the ages of both groups, on the ratio of the standard deviations of the measures and on the slopes of both measures with age. The precise prediction depends on the specific relationship between these three components. In general, the difference in slopes and the ratio of the standard deviations are more important in determining a nonword-reading deficit than the height of the correlations.

To conclude, Experiment 1 shows that the nonword-reading deficit clearly varies as a function of age and reading level. It can be plausibly construed as the consequence of a normal developmental process in which the role of word-specific knowledge in word recognition increases with age.
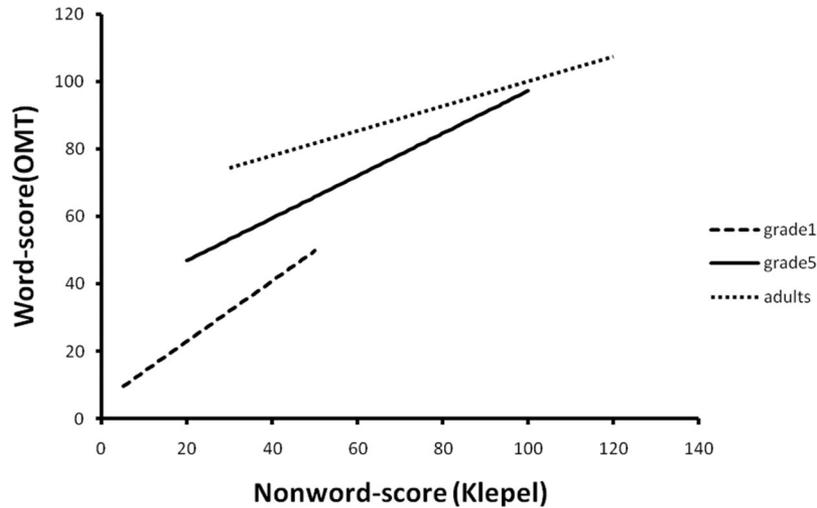
Figure 2. Interaction effect of nonword reading and grade level on word-reading ability.

Experiment 2

If a nonword-reading deficit is really the consequence of normal developmental differences in word-specific knowledge, then the invoked word-specific knowledge in the word-reading task should be a crucial variable in the determination of a nonword-reading deficit. According to the meta-analytical position (cf. van IJzendoorn & Bus, 1994), varying task characteristics are only responsible for Type II errors (the real nonword-reading deficit is too small to be statistically detectable). In contrast, the developmental hypothesis predicts that there will be a dramatic change from a minimal or absent nonword-reading deficit when the required word-specific knowledge in the word-reading task is largely in the reading ability range of all participating children to a substantial nonword-reading deficit when the required word-specific knowledge is at a considerably higher level. To this end, in Experiment 2, we varied the required word-specific knowledge by employing two word recognition tests as matching variables. The first word recognition test is the Two-Minute Test (TMT; Wassenaar, 1989), containing only words of relatively high frequency (see Table 2). Because high-frequency words are typically encountered much earlier in development than low-frequency words, we assume that the younger normal readers' specific knowledge of these high-frequency words is approximately the same as that of the older disabled readers. Hence, the TMT is a test of low word-specificity requirements. In the first analysis, the nonword reading of disabled readers is compared with the nonword reading of normal readers after matching both groups on the TMT. In this analysis, the matching variable is identical to the selection variable, avoiding unrepresentative sampling and the ensuing bias from regression effects.

However, in many English-speaking studies that show a nonword-reading deficit, participants were clinically selected and administered a word-reading test as the matching variable, requiring a high level of word-specific knowledge (e.g., the reading subtest of the Wide Range Achievement Test–Revised, Jastak & Wilkinson, 1984). As the clinical selection of disabled readers is most probably based on the reading of books and texts appropriate to the children's age, this selection does not involve a high level of word-specific knowledge. To simulate this situation as closely as possible in our second analysis, we selected our participants using the TMT (with low word specificity requirements), and administered another word reading test (Word Recognition Test) as a matching variable involving a high level of word-specific knowledge (mostly low-frequency multisyllabic words). Because in this analysis the matching variable differed from the selection variable, bias from regression effects is possible.

On the basis of our developmental interpretation of the nonword-reading deficit, a second aim in our study was to test the prediction again that average readers show a nonword-reading deficit when compared with excellent younger readers of the same reading level.

A final goal in this study was to gain additional independent empirical evidence to justify our contention that the word-specific knowledge of the older disabled readers differs from that of the younger normal readers. To this end we administered several orthographic knowledge tasks.

Method

Design and analytic strategy. In this study, we used the same regression-based logic to the RLM design, as explained in Experiment 1. To avoid the statistical problem of nonoverlapping distributions, we present all analyses, using only the overlapping part of the distributions.

Participants. Participants in the study were 593 Dutch-speaking children of Grades 2, 3, and 4, all attending regular elementary schools. Boys constituted 57.5% of the sample, and girls made up 42.5%. All children attended regular elementary schools, located in the region of Leiden (a smaller provincial town) or in the region of The Hague (the third largest city in the Netherlands). The schools were located in middle-class neighborhoods. Among the participants, 74% were native Dutch and 26% were of North African (Morocco), Turkish, Indonesian, or Surinamese origin. All participants had a sufficient command of the

Dutch language to be able to attend the Dutch curricula. Participants were recruited for testing from entire classes.

The participants were classified into two groups: disabled and nondisabled readers. The disabled readers' percentile scores on a norm-referenced word recognition test (i.e., the TMT), were ≤ 10th percentile compared with their age group, and their estimated verbal IQs (based on the Vocabulary subtest of the Wechsler Intelligence Scale for Children–Revised) were > 90. The group of nondisabled readers consisted of the children scoring in the ≥ 30th percentile on the TMT. No children displayed serious language, emotional, or behavioral disorders. This procedure resulted in 62 disabled readers and 414 nondisabled readers. For each of the two groups in each of the three grade levels, Table 5 displays means and SDs for the TMT, for the Word Recognition Test, for the nonword-reading test (Klepel), and for the ages in months.

Measures.

Word identification. Two word identification tests in our battery were used in the analyses as matching variables.

TMT. As in a typical Dutch norm-referenced word recognition test, the TMT requires that the participants read the words aloud as accurately and as quickly as possible during a prescribed length of time (2 min). The score is the number of words read correctly. The test consists of 112 real regular words, listed in four columns and ordered from lower to higher orthographic complexity (see Appendix B for a sample of the items).

Most words are monosyllabic (92 words), and 20 words are disyllabic. All the words that are of relatively high frequency and high orthographic similarity (see Table 2) are in the reading ability range and vocabulary of normal Grade 2 to Grade 4 children. Therefore, the required word-specific knowledge is not beyond the typical reading ability range of the participating children. Parallel reliability of this test is .90. The correlations of this test with all other tests in this experiment are reported in Table 6.

Word Recognition Test (WRT). To maximize the chance to observe a nonword-reading deficit, we imitated as closely as possible the procedure of many English language word recognition tests. Thus, the Word Recognition Test consists of 40 words that increase very rapidly in degree of difficulty (see Appendix B). After 10 monosyllabic words, the test continues with multisyllabic words (up to seven syllables). Except for the first 10 words, the words are of very low frequency (see Table 2). Also, the high orthographic uniqueness of these words (see OLD-20 index in Table 2) indicates that specific knowledge of these words is typically acquired later in reading development. The children were instructed to read the words as accurately and as quickly as possible. The score is the number of words read correctly in 1 min. It is notable that although this test does not contain exception words, the invoked word-specific knowledge is at a higher level than the typical knowledge of Grade 2 to Grade 4 children. Reliability was estimated using the test–retest method on half of the subjects. The reliability coefficient is .85. Thus, despite the enhanced difficulty, the discriminative power of this test is excellent.

Nonword reading. The nonword-reading test (Klepel) is the same as described in Experiment 1.

Orthographic knowledge tasks.

Orthographic choice task. In this task, children are asked to choose the correct spelling from two homophonic alternatives (e.g., *zout*, *zaut*). Because the two alternatives sound the same when decoded, knowledge of the specific orthography of each word is necessary to solve this task. Thus, the task may be considered a relatively pure measure of orthographic knowledge. Subsequent to two practice trials, each child receives 35 experimental trials. The children are instructed to underline the correct spelling and are asked to work as fast as possible. The score is the number of correctly underlined words divided by the time used. A conservative estimate of the reliability, based on the correlations with the other tests, is .80.

Exception words. The reading of exception words is traditionally regarded as the cornerstone of the assumption of a visual

Table 5

Mean Reading Scores (TMT, WRT, and Klepel) and Age as a Function of Subject Classification and Grade

Variable	Nondisabled readers			Disabled readers		
	M	*SD*	*n*	*M*	*SD*	*n*
TMT raw score						
Grade 2	90.6	15.5	111	26.5	7.7	19
Grade 3	105.2	6.2	160	60.0	10.6	22
Grade 4	108.0	4.9	143	76.5	13.6	21
WRT raw score						
Grade 2	22.8	5.6	111	12.1	6.9	19
Grade 3	28.4	5.0	160	19.7	5.1	22
Grade 4	31.9	4.0	143	23.2	7.4	21
Klepel raw score						
Grade 2	41.9	15.1	111	12.1	6.0	19
Grade 3	55.8	15.7	160	25.2	8.8	22
Grade 4	65.3	16.5	143	30.4	10.5	21
Age (in months)						
Grade 2	94.2	5.5	111	97.1	6.1	19
Grade 3	106.6	5.8	160	110.4	7.1	22
Grade 4	118.4	5.2	143	121.2	6.6	21

Note. TMT = Two Minute Test; WRT = Word Recognition Test; Klepel = nonword-reading test.

Table 6
Correlations of the Tasks in Our Battery (N = 593)

Tasks	1	2	3	4	5
1. Two Minute Test	—				
2. Word Recognition Test	.74	—			
3. Nonword reading (Klepel)	.87	.75	—		
4. Orthographic choice	.75	.65	.67	—	
5. Exception words	.84	.78	.77	.80	—

lexical route in the dual-route theory (Coltheart, 1978). Although not all grapheme–phoneme correspondences in these words are irregular, knowledge of the specific orthographic patterns at least facilitates the reading of exception words. Because exception words are rather uncommon in Dutch, most words are loan words from English or French (e.g., *goal*, *atelier*). All 50 words in this test (see Appendix B) are irregular at the grapheme–phoneme level, and most words are also inconsistent at the rime level. Although the words are highly frequent, their medium orthographic uniqueness invokes a higher level of word-specific knowledge compared with the selection variable (see Table 2). The words are to be read aloud as accurately and as quickly as possible during 1 min. The score is the number of words read correctly. The estimated reliability of this task is .84. All tasks were administered in two sessions in a fixed order for every participant.

Results

As in the first experiment, hierarchical regression analyses were conducted. All analyses followed the same logic (see also Experiment 1). First, we examined whether the relationship between the matching variable and the dependent variable fitted a nonlinear trend (quadratic and/or cubic). Second, a test of the interaction of group (reading disabled or not) by matching variable determined whether the observed relationship in Step 1 differed for both groups. Finally, regardless of the result of the interaction test, the overall group effect between disabled and nondisabled readers was calculated.

All analyses are presented for the overlapping distributions between disabled and nondisabled readers (see Tables 7 and 8). As can be inferred from the sample sizes indicated in Tables 7 and 8, the overlap between disabled and nondisabled readers is much smaller with the TMT as the matching variable, resulting in a considerable reduction of sample size compared with the Word Recognition Test as the matching variable. This is a consequence of the use of the TMT as both the matching variable and the selection variable for dividing the participants into the groups of disabled and nondisabled readers. However, a sufficient number of disabled ($n = 33$) and nondisabled readers ($n = 68$) remained in this analysis to allow statistically meaningful conclusions.

When the Word Recognition Test was employed as the matching variable (see Table 7), requiring a high level of word-specific knowledge, a clear nonword-reading deficit was found. Also a significant ordinal interaction was detected, $F(1, 372) = 18.30$, $p = .000$, indicating that the nonword-reading deficit diminishes for the lower range of the matching task. Nevertheless, the overall nonword-reading deficit was significant and explained 7.9% of unique variance, $F(1, 399) = 63.25$, $p = .000$.

Table 7
Summary of Hierarchical Regression Analysis for the WRT as the Matching Variable

Task	Dependent variables		
	Nonword reading	Orthographic choice	Exception words
	Unique variance explained (ΔR^2)		
Quadratic fit	.005	.001	.000
Interaction	.021***	.020***	.026***
Group effect	.079***	.013**	.019***
95% CI	[.041, .117]	[.000, .030]	[.002, .037]
	β in final equation		
Group effect	−.295***	−.122**	−.145***
	F ratio in final equation		
Group effect	63.25	8.96	18.22
	Sample size		
N	402	402	402

Note. A negative beta indicates a processing deficit for the disabled readers. The reported confidence limits are for the group effects.
* $p < .05$. ** $p < .01$. *** $p < .001$.

As the Word Recognition Test is also sensitive to differences in word-specific knowledge at higher reading levels, we were able to test the prediction that average readers also show a nonword-reading deficit when compared with excellent younger readers of the same reading level. We defined average readers as reading on the TMT between the 40th and 60th percentiles and younger excellent readers as reading on the TMT above the 80th percentile

Table 8
Summary of Hierarchical Regression Analysis for the TMT as the Matching Variable

Task	Dependent variable		
	Nonword reading	Orthographic choice	Exception words
	Unique variance explained (ΔR^2)		
Quadratic fit	.002	.014	.012
Interaction	.007	.001	.000
Group effect	.000	.159***	.193***
95% CI		[.033, .285]	[.065, .322]
	β in final equation		
Group effect	−.016	.413***	.455***
	F ratio in final equation		
Group effect	0.03	20.42	28.46
	Sample size		
N	101	101	101

Note. A negative beta indicates a processing deficit for the disabled readers, and a positive beta indicates a processing advantage for the disabled readers. The confidence limits are for the group effects and are reported if significant.
* $p < .05$. ** $p < .01$. *** $p < .001$.

in their age group and matched both groups on the Word Recognition Test; we demonstrated a very substantial nonword-reading deficit for the average readers ($N = 101$) compared with the excellent younger readers ($N = 118$); $\Delta R^2 = 15.8\%$, $F(1, 216) = 71.81$, $p = .000$. In contrast to the first study, this huge effect is probably inflated due to bias from regression effects.

In addition, the results show that a nonword-reading deficit could not be detected when the children were statistically matched on the TMT (see Table 8). Disabled readers did not differ at all from normal readers on the nonword-reading test, $\Delta R^2 = 0\%$; $F(1, 98) = 0.03$, $p = .86$.

The results of the tasks measuring orthographic knowledge can be summarized as follows. Taking the TMT as the matching variable (see Table 8), the results of the two orthographic knowledge tasks display a consistent pattern. The disabled readers' performance is superior to the performance of the RLM normal readers on the orthographic choice task, $F(1, 98) = 20.42$, $p = .000$, and the exception words, $F(1, 98) = 28.46$, $p = .000$. Note that the magnitude of this orthographic effect is, respectively, 16% and 30% of explained variance. When we used the Word Recognition Test as the matching variable, the orthographic effects were small and negative (see Table 7).

Discussion

The major goal of Experiment 2 was to examine whether the nonword-reading deficit depends on the required word-specific knowledge of the word-reading task. The results demonstrated that when children were matched on a word-reading task demanding word-specific knowledge that is typical for higher reading levels, a substantial nonword-reading deficit emerged. If the participants, however, were matched on a word-reading test that invoked word-specific knowledge in the reading ability range of all participating children, no nonword-reading deficit was found. These results strengthen the hypothesis formulated in the first experiment and suggest that a nonword-reading deficit should appear when the matching variable is a word-reading test that is sensitive to the different degrees of word-specific knowledge of the different groups.

A second purpose of this study was to investigate whether older disabled readers would display superior orthographic knowledge compared with the younger control subjects. Remarkably, in the situation where no nonword-reading deficit was observed (when using the TMT as the matching variable), an orthographic advantage for the disabled readers emerged. Obviously, as there was no need to compensate any deficit, the orthographic compensation hypothesis (see Introduction) can be discarded. Alternatively, this orthographic advantage may be explained in terms of a higher word-specific knowledge of the older disabled readers. As the older disabled readers may have been able to rely on a more developmentally advanced level of word-specific knowledge in the orthographic tasks, an orthographic advantage may have emerged. Future research should be designed to test this hypothesis more directly by measuring the actual word-specific knowledge of each child. The small and negative orthographic effects observed when the WRT was used as the matching variable indicate that the word-specific knowledge required in this test, containing many low frequency words, is still at a higher level than the word-specific knowledge needed in the orthographic knowledge tasks.

Conclusions and General Discussion

In this article, we have formulated and tested an alternative interpretation for the nonword-reading deficit by rigorously analyzing the methodological restrictions of the RLM design. In both experiments we observed a nonword-reading deficit in a relatively consistent orthography like Dutch (Experiments 1 and 2). This effect, however, was only detected in a well-specified region of word-reading ability (Experiment 1) and depended crucially on the required word-specific knowledge of the word reading task (Experiment 2). Importantly, the nonword-reading deficit also applied to average readers for a rather broad range of reading levels when they were compared with excellent younger readers of the same reading level (Experiments 1 and 2). Finally, Experiment 2 showed that the word-specific knowledge of the older disabled readers differed from that of the younger normal readers.

Taken together, these findings suggest that the coming and going of a nonword-reading deficit crucially depends on the sensitivity of the word-reading task to detect differences in word-specific knowledge between disabled and younger normal readers. We argued that these differences reflect a normal developmental trend: With increasing age, word reading is determined by developmentally higher levels of word-specific knowledge. Our finding that a nonword-reading deficit could also be observed in normal readers when compared with excellent younger readers with the same word-reading level supported our interpretation. This hypothesis offers a novel explanation for the nonword-reading deficit and stands in sharp contrast to the established view that the nonword-reading deficit is a distinctive feature of disabled reading and represents a robust effect.

Implications for the Phonological Deficit Hypothesis

What are the implications of our findings for theories pertaining to the underlying causes of poor or disabled reading? As researchers generally regard a nonword-reading deficit as important evidence for the phonological deficit hypothesis, one might argue that the interpretation of a nonword-reading deficit as a consequence of normal developmental processes potentially weakens the phonological deficit hypothesis. However, the following arguments show that this reasoning does not hold; our study cannot be interpreted as evidence against the established phonological deficit hypothesis. First, apart from the nonword-reading deficit, a wealth of evidence shows that reading-disabled children have a specific impairment in representing, storing, and retrieving phonological information (for reviews, see Ramus, 2003; Vellutino et al., 2004). For instance, research demonstrates that disabled readers have difficulty detecting, comparing, or manipulating phonemic segments in words. They are slower at naming objects and symbols, and they show problems with the coding of phonological information in short-term memory (presumably due to poor quality of the underlying phonological representations). Moreover, they reveal subtle difficulties with speech perception and speech production tasks.

Second, the prediction that disabled readers have a nonword-reading deficit has been based on the conventional assumption that nonwords specifically invoke phonological processing, whereas the reading of real words (regular or exception) would depend less on phonological processes. However, all recent theoretical ac-

counts of written word recognition agree that apart from a sequential, more laborious grapheme–phoneme mode of processing, a faster interactive mode of processing is needed to read real words fluently and almost automatically (cf. Coltheart et al., 2001; Perfetti, Liu, & Tan, 2005; Perry et al., 2007). It is important to note that this interactive mode of processing written words is also phonological in nature. A series of empirical findings suggests a mandatory or at least crucial role for phonological processing in word recognition (Perfetti & Bell, 1991; Perfetti, Bell, & Delaney, 1988; Perfetti, Zhang, & Berent, 1992; Van Orden, 1987; see also Frost, 1998, for an excellent review and Ehri, 1998, for a developmental account of these ideas). For example, in the phonologic coherence theory of Van Orden and Goldinger (1994), orthographic and phonologic codes become fully and irreducibly interdependent in the course of the reading process. Furthermore, the lexical quality hypothesis (Perfetti, 1992; Perfetti & Hart, 2002) postulates that reading skill is the extent to which a reader's lexicon is characterized by fully specified lexical representations (orthographic form, phonological form, and meaning). Now, if phonology is seen as an essential constituent of word recognition (rather than a by-product), as the lexical quality hypothesis advocates, and reading disability is indeed the result of a phonological deficit, as many researchers argue, then a specific deficit in the reading of nonwords would not be a natural prediction. Given the available evidence, phonology is crucial in both word and nonword reading. Why then should a phonological deficit pertain to nonwords only? In addition, we assume that the problems of disabled readers in becoming aware of the explicit grapheme–phoneme correspondences originate from the same underlying phonological disturbance as the difficulty of implicitly processing phonological information during the fast identification of written words. This assumption predicts that disabled readers do not show a specific nonword-reading deficit. Clearly, the presented empirical data are consistent with this hypothesis.

The absence of a genuine nonword-reading deficit applies to the disabled readers as a group. Nevertheless, on the basis of the subtype literature (see the Introduction) one might argue that within this population a subgroup does show a genuine nonword-reading deficit. Although this subgroup of phonological dyslexics would probably be small, the observed variability in the difference scores between word reading and nonword reading certainly leaves some room for the existence of subtypes. It should be taken into account, however, that the extremely skewed proportions of phonological dyslexics vs. surface dyslexics in the studies using an RLM design are demonstrably the consequence of the age confound discussed here (Van den Broeck & Geudens, 2009).

Implications for Theories of Reading Development

Our developmental interpretation of the nonword-reading deficit assumes that disabled readers are capable of acquiring orthographic knowledge of the specific words they encounter during reading. It should be noted that the word-specific knowledge of the disabled readers is not regarded as "normal," yet it is better than that of the younger normal readers. Our assumption in terms of developmental differences is also in agreement with some widely cited studies in Dutch (Reitsma, 1983) and in English (Ehri & Saltmarsh, 1995; Hogaboam & Perfetti, 1978), demonstrating that orthographic learning does occur in disabled readers but requires

more exposures than in normal readers (see also Barca, Burani, Di Filippo, & Zoccolotti, 2006; Moll & Landerl, 2009). According to the self-teaching model (Jorm & Share, 1983; Share, 1995, 2004), phonological recoding (print-to-sound translation) functions as a self-teaching mechanism that enables the learner to acquire the detailed orthographic representations necessary for fluent, efficient visual word recognition. Each successful identification (decoding) of an encountered word provides an opportunity to acquire this word-specific orthographic information. Moreover, the self-teaching model hypothesizes that orthographic learning is severely constrained by decoding proficiency and hence predicts that problems in orthographic learning are commensurate with the decoding problems. Empirical support for this prediction was obtained in a study by Share and Shalev (2004), who demonstrated a close link between levels of target word decoding and the acquisition of orthographic information among two groups of poor readers (IQ discrepant and IQ nondiscrepant) and a group of chronological-age-matched normal readers (fourth grade). These results are also congruent with our contention that the underlying phonological deficit of disabled readers equally affects the acquisition of grapheme–phoneme correspondences and the attainment of the fast interactive mode of processing.

We assume that the orthographic knowledge of both disabled and normal readers applies to all word and subword levels. PGS theory, however, states that the degree of orthographic consistency of the language imposes the grain size of the processing units used in reading words: small units in consistent orthographies, and both small and larger units (e.g., consonant clusters, rimes, syllables) in inconsistent orthographies (Ziegler & Goswami, 2005). Because there are far more orthographic units to learn when the grain size is large than when the grain size is small, it takes longer to master an inconsistent orthography (Seymour, Aro, & Erskine, 2003). Research in consistent orthographies, however, has suggested that both beginning and fluent reading in all orthographies entails sublexical knowledge beyond the grapheme–phoneme level (cf. Geudens & Sandra, 2002; Wolters, 2004). To account for this empirical evidence, several researchers have made suggestions to amend PGS theory (De Jong, 2006; Frost, 2006; Paulesu, 2006; Wimmer, 2006). In their response, Goswami and Ziegler (2006) agreed with the proposal that large grain size couplings between orthography, phonology, and semantics will sooner or later emerge in all orthographies.

We agree with PGS theory that orthographic consistency determines the size of processing units in the beginning phase of reading acquisition, but we adapt the theory further to accommodate the developmental changes we have discussed. We assume that all readers (also disabled readers) in an alphabetic orthography progress from the slow serial mode of processing to the fast interactive mode of processing, which is driven by the development of word-specific knowledge. Furthermore, we suggest that the pace of this shift depends on the orthographic–phonological consistency of the language and on the individual capability to store all the relevant orthography–phonology mappings. Readers in inconsistent orthographies need more time before they can make this shift, just as is the case for disabled readers. Regrettably, disabled readers never completely catch up due to their underlying deficit, whereas the initial developmental difference between normal readers in an inconsistent versus a consistent orthography

gradually disappears with time (Landerl, 2000). The role of word-specific knowledge becomes dominant in that phase.

This developmental dual process hypothesis is congenial to Ehri's (1992) amalgamation theory, Perfetti's (1992) lexical quality hypothesis, and Share's (1995) self-teaching model, and it intends to integrate all of these within the context of PGS theory. Common to all these theoretical accounts is the idea that decoding ability underpins the development of word-specific knowledge, which in its turn facilitates grapheme–phoneme couplings to become increasingly context-sensitive or "lexicalized" (Share, 1995). Presumably, decoding ability is required for the growth of word-specific knowledge. Word-specific knowledge, on the other hand, substantially contributes to the fluent reading of nonwords. In line with this view, a growing body of empirical evidence (Cohen & Dehaene, 2004; Cohen et al., 2000, 2002; Dehaene, Le Clec'H, Poline, Le Bihan, & Cohen, 2002) reveals a region in the left ventral occipitotemporal cortex that functions as a visual word form area (VWFA). On the basis of fMRI data, Kronbichler et al. (2004, 2007) interpreted the function of VWFA as an orthographic input lexicon, where incoming letter strings are matched against orthographic lexicon entries. However, other interpretations (e.g., prelexical recognition of recurring letter sequences) of the role of VWFA are currently debated (cf. Cohen et al., 2002; Devlin, Jamison, Gonnerman, & Matthews, 2006).

It appears that reading development depends on two intricately related processes: an explicit rule-based system and a developmental item-based system, where the first rule-based system functions as a scaffold for the second item-based system. These systems both lie at the foundation of word and nonword reading.

References

Alf, E. F., & Graf, R. G. (1999). Asymptotic confidence limits for the difference between two squared multiple correlations: A simplified approach. *Psychological Methods, 4,* 70–75.

Baayen, R. H., Piepenbrock, R., & Gulikers, L. (1995). *The CELEX lexical database* (CD-ROM). Philadelphia, PA: University of Pennsylvania, Linguistic Data Consortium.

Backman, J. E., Mamen, M., & Ferguson, H. (1984). Reading-level design: Conceptual and methodological issues in reading research. *Psychological Bulletin, 96,* 560–568.

Barca, L., Burani, C., Di Filippo, G., & Zoccolotti, P. (2006). Italian developmental dyslexic and proficient readers: Where are the differences? *Brain & Language, 98,* 347–351.

Barron, R. W. (1986). Word recognition in early reading: A review of the direct and indirect access hypotheses. *Cognition, 24,* 93–119.

Brus, B. T., & Voeten, M. J. M. (1973). *Eén minuut-test* [One Minute Test]. Nijmegen, the Netherlands: Berkhout.

Bryant, P. E., & Goswami, U. (1986). Strengths and weaknesses of the reading level design: A comment on Backman, Mamen, and Ferguson. *Psychological Bulletin, 100,* 101–103.

Castles, A., & Coltheart, M. (1993). Varieties of developmental dyslexia. *Cognition, 47,* 149–180.

Castles, A., & Coltheart, M. (1996). Cognitive correlates of developmental surface dyslexia: A single case study. *Cognitive Neuropsychology, 13,* 25–50.

Cohen, L., & Dehaene, S. (2004). Specialization within the ventral stream: The case for the visual word form area. *NeuroImage, 22,* 466–476.

Cohen, L., Dehaene, S., Naccache, L., Lehericy, S., Dehaene-Lambertz, G., Hennaf, M. A., & Michel, F. (2000). The visual word form area: Spatial and temporal characterization of an initial stage of reading in normal subjects and posterior split-brain patients. *Brain, 123,* 291–307.

Cohen, L., Lehericy, S., Chochon, F., Lemer, C., Rivaud, S., & Dehaene, S. (2002). Language-specific tuning of visual cortex? Functional properties of the visual word form area. *Brain, 125,* 1054–1069.

Coltheart, M. (1978). Lexical access in simple reading tasks. In G. Underwood (Ed.), *Strategies of information processing* (pp. 151–216). New York, NY: Academic Press.

Coltheart, M., Davelaar, E., Jonasson, J. T., & Besner, D. (1977). Access to the internal lexicon. In S. Dormic (Ed.), *Attention and performance VI* (pp. 535–555). Hillsdale, NJ: Erlbaum.

Coltheart, M., Masterson, J., Byng, S., Prior, M., & Riddoch, J. (1983). Surface dyslexia. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 35*(A), 469–495.

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review, 108,* 204–256.

Dehaene, S., Le Clec'H, G., Poline, J. B., Le Bihan, D., & Cohen, L. (2002). The visual word form area: A prelexical representation of visual words in the fusiform gyrus. *NeuroReport, 13,* 321–325.

De Jong, P. F. (2006). Units and routes of reading in Dutch. *Developmental Science, 9,* 441–442.

Devlin, J. T., Jamison, H. L., Gonnerman, L. M., & Matthews, P. M. (2006). The role of the left posterior fusiform gyrus in reading. *Journal of Cognitive Neuroscience, 18,* 911–922.

Ehri, L. C. (1992). Reconceptualizing the development of sight word reading and its relationship to recoding. In P. B. Gough, L. C. Ehri, & R. Treiman (Eds.), *Reading acquisition* (pp. 107–143). Hillsdale, NJ: Erlbaum.

Ehri, L. C. (1998). Grapheme–phoneme knowledge is essential for learning to read words in English. In J. L. Metsala & L. C. Ehri (Eds.), *Word recognition in beginning literacy* (pp. 3–40). London, England: Erlbaum.

Ehri, L. C., & Saltmarsh, J. (1995). Beginning readers outperform older disabled readers in learning to read words by sight. *Reading and Writing: An Interdisciplinary Journal, 7,* 295–326.

Ellis, A. W., McDougall, S., & Monk, A. F. (1996). Are dyslexics different? I. A comparison between dyslexics, reading age controls, poor readers, and precocious readers. *Dyslexia, 2,* 31–58.

Foorman, B. R., Francis, D. J., Fletcher, J. M., & Lynn, A. (1996). Relation of phonological and orthographic processing to early reading: Comparing two approaches to regression-based, reading-level-match designs. *Journal of Educational Psychology, 88,* 639–652.

Frith, U. (1985). Beneath the surface of developmental dyslexia. In K. E. Patterson, J. C. Marshall, & M. Coltheart (Eds.), *Surface dyslexia* (pp. 301–322). London, England: Erlbaum.

Frost, R. (1998). Toward a strong phonological theory of visual word recognition: True issues and false trails. *Psychological Bulletin, 123,* 71–99.

Frost, R. (2006). Becoming literate in Hebrew: The grain size hypothesis and Semitic orthographic systems. *Developmental Science, 9,* 439–440.

Gallagher, A., Frith, U., & Snowling, M. J. (2000). Precursors of literacy delay among children at genetic risk of dyslexia. *Journal of Child Psychology and Psychiatry, 41,* 203–213.

Geudens, A., & Sandra, D. (2002). The role of orthographic onset-rime units in Dutch beginning readers. In L. Verhoeven, P. Reitsma, & C. Elbro (Eds.), *Precursors of functional literacy* (pp. 49–67). Amsterdam, the Netherlands: Benjamins.

Geudens, A., Sandra, D., & Van den Broeck, W. (2004). Segmenting two-phoneme syllables: Developmental differences in relation to early reading skills. *Brain and Language, 90,* 338–352.

Goswami, U., & Bryant, P. E. (1989). The interpretation of studies using the reading level design. *Journal of Reading Behavior, 21,* 413–424.

Goswami, U., & Ziegler, J. C. (2006). Fluency, phonology and morphology: A response to the commentaries on becoming literate in different languages. *Developmental Science, 9,* 451–453.

Griffiths, Y. M., & Snowling, M. J. (2002). Predictors of exception word and nonword reading in dyslexic children: The severity hypothesis. *Journal of Educational Psychology, 94,* 34–43.

Harm, M. W., & Seidenberg, M. S. (1999). Phonology, reading acquisition, and dyslexia: Insights from connectionist models. *Psychological Review, 106,* 491–528.

Herrmann, J. A., Matyas, T., & Pratt, C. (2006). Meta-analysis of the nonword-reading deficit in specific reading disorder. *Dyslexia, 12,* 195–221.

Hogaboam, T. W., & Perfetti, C. A. (1978). Reading skill and the role of verbal experience in decoding. *Journal of Educational Psychology, 70,* 717–729.

Huitema, B. (1980). *Analysis of covariance and alternatives.* New York, NY: Wiley.

Humphreys, G. W., & Evett, L. J. (1985). Are there independent lexical and nonlexical routes in word processing? An evaluation of the dual-route theory of reading. *The Behavioral and Brain Sciences, 8,* 689–740.

Hutzler, F., Ziegler, J. C., Perry, C., Wimmer, H., & Zorzi, M. (2004). Do current connectionist models account for reading development in different languages? *Cognition, 91,* 273–296.

Jackson, N. E., & Butterfield, E. C. (1989). Reading-level-matched designs: Myths and realities. *Journal of Reading Behavior, 21,* 387–412.

Jackson, N. E., & Coltheart, M. (2001). *Routes to reading success and failure. Toward an integrated cognitive psychology of atypical reading.* New York, NY: Taylor & Francis.

Jastak, S., & Wilkinson, G. S. (1984). *The Wide Range Achievement Test-Revised.* Wilmington, DE: Jastak.

Jiménez, J. E., & Hernandez, I. (2000). Word identification and reading disorders in the Spanish language. *Journal of Learning Disabilities, 33,* 44–60.

Jorm, A. F., & Share, D. L. (1983). Phonological recoding and reading acquisition. *Applied Psycholinguistics, 4,* 103–147.

Kronbichler, M., Bergmann, J., Hutzler, F., Staffen, W., Mair, A., Ladurner, G., & Wimmer, H. (2007). Taxi vs. taksi: On orthographic word recognition in the left ventral occipitotemporal cortex. *Journal of Cognitive Neuroscience, 19,* 1584–1594.

Kronbichler, M., Hutzler, F., Wimmer, H., Mair, A., Staffen, W., & Ladurner, G. (2004). The visual word form area and the frequency with which words are encountered: Evidence from a parametric fMRI study. *NeuroImage, 21,* 946–953.

Landerl, K. (2000). Influences of orthographic consistency and reading instruction on the development of nonword reading skills. *European Journal of Psychology of Education, 15,* 239–257.

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

Landerl, K., Wimmer, H., & Frith, U. (1997). The impact of orthographic consistency on dyslexia: A German–English comparison. *Cognition, 63,* 315–334.

Levenshtein, V. I. (1966). Binary codes capable of correcting deletions, insertions and reversals. *Soviet Physics Doklady, 10,* 707.

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

Manis, F. R., Seidenberg, M. S., Doi, M. L., McBride-Chang, C., & Petersen, A. (1996). On the bases of two subtypes of developmental dyslexia. *Cognition, 58,* 157–195.

Moll, K., & Landerl, K. (2009). Double dissociation between reading and spelling deficits. *Scientific Studies of Reading, 13,* 359–382.

Murphy, L., & Pollatsek, A. (1994). Developmental dyslexia: Heterogeneity without discrete subgroups. *Annals of Dyslexia, 44,* 120–146.

Olson, R. K., Wise, B., Conners, F., & Rack, J. P. (1990). Organization, heritability, and remediation of component word recognition and language skills in disabled readers. In T. H. Carr & B. A. Levy (Eds.), *Reading and its development: Component skills approaches* (pp. 261–322). San Diego, CA: Academic Press.

Paulesu, E. (2006). On the advantage of "shallow" orthographies: Number and grain size of the orthographic units or consistency per se? *Developmental Science, 9,* 443–444.

Paulesu, E., Demonet, J., Fazio, F., McCrory, E., Chanoine, V., Brunswick, N., . . . Frith, U. (2001, March 16). Dyslexia: Cultural diversity and biological unity. *Science, 291,* 2165–2167.

Pedhazur, E. J. (1982). *Multiple regression in behavioral research: Explanation and prediction.* Orlando, FL: Holt, Rinehart, & Winston.

Perfetti, C. A. (1992). The representation problem in reading acquisition. In P. B. Gough, L. C. Ehri, & R. Treiman (Eds.), *Reading acquisition* (pp. 145–174). Hillsdale, NJ: Erlbaum.

Perfetti, C. A., & Bell, L. (1991). Phonemic activation during the first 40 ms of word identification: Evidence from backward masking and masked priming. *Journal of Memory and Language, 30,* 473–485.

Perfetti, C. A., Bell, L., & Delaney, S. (1988). Automatic phonetic activation in silent word reading: Evidence from backward masking. *Journal of Memory and Language, 27,* 59–70.

Perfetti, C. A., & Hart, L. (2002). The lexical quality hypothesis. In L. Vehoeven, C. Elbro, & P. Reitsma (Eds.), *Precursors of functional literacy* (pp. 189–213). Amsterdam, the Netherlands: Benjamins.

Perfetti, C. A., Liu, Y., & Tan, L. H. (2005). The lexical constituency model: Some implications of research on Chinese for general theories of reading. *Psychological Review, 112,* 43–59.

Perfetti, C. A., Zhang, S., & Berent, I. (1992). Reading in English and Chinese: Evidence for a universal phonological principle. In R. Frost & L. Katz (Eds.), *Orthography, phonology, morphology, and meaning* (pp. 227–248). Amsterdam, the Netherlands: North-Holland.

Perry, C., Ziegler, J. C., & Zorzi, M. (2007). Nested incremental modeling in the development of computational theories: The CDP+ model of reading aloud. *Psychological Review, 114,* 273–315.

Pothoff, R. F. (1964). On the Johnson-Neyman technique and some extensions thereof. *Psychometrika, 29,* 241–256.

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.

Ramus, F. (2003). Developmental dyslexia: Specific phonological deficit or general sensorimotor dysfunction. *Current Opinion in Neurobiology, 13,* 212–218.

Reitsma, P. (1983). Word-specific knowledge in beginning reading. *Journal of Research in Reading, 6,* 41–56.

Riethoven, M. K., & Hooijmans, N. C. P. (2003). *De rol van woordspecifieke kennis in het leesproces van zwakke lezers* [The role of word-specific knowledge in the reading process of poor readers]. (Unpublished masters thesis). Leiden University, Leiden, the Netherlands.

Rogosa, D. R. (1980). Comparing nonparallel regression lines. *Psychological Bulletin, 88,* 307–321.

Rogosa, D. R. (1981). On the relationship between the Johnson-Neyman region of significance and statistical tests of parallel within-group regressions. *Educational and Psychological Measurement, 41,* 73–84.

Salthouse, T. A., & Hedden, T. (2002). Interpreting reaction time measures in between-group comparisons. *Journal of Clinical and Experimental Neuropsychology, 24,* 858–872.

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

Seymour, P. H. K., & Elder, L. (1986). Beginning reading without phonology. *Cognitive Neuropsychology, 3,* 1–36.

Share, D. L. (1995). Phonological recoding and self-teaching: Sine qua non of reading acquisition. *Cognition, 55,* 151–218.

Share, D. L. (2004). Orthographic learning at a glance: On the time course and developmental onset of self-teaching. *Journal of Experimental Child Psychology, 87,* 267–298.

Share, D. L. (2008). On the Anglocentricities of current reading research and practice: The perils of overreliance on an "outlier" orthography. *Psychological Bulletin, 134,* 584–615.

Share, D. L., & Shalev, C. (2004). Self-teaching in normal and disabled readers. *Reading and Writing: An Interdisciplinary Journal, 17,* 769–800.

Snowling, M. J., Bryant, P. E., & Hulme, C. (1996). Theoretical and methodological pitfalls in making comparisons between developmental and acquired dyslexia: Some comments on A. Castles and M. Coltheart (1993). *Reading and Writing, 8,* 443–451.

Snowling, M. J., Gallagher, A., & Frith, U. (2003). Family risk of dyslexia is continuous: Individual differences in the precursors of reading skill. *Child Development, 74,* 358–373.

Snowling, M. J., Goulandris, N., & Defty, N. (1996). A longitudinal study of reading development in dyslexic children. *Journal of Educational Psychology, 88,* 653–669.

Stanovich, K. E., & Cunningham, A. E. (1993). Where does knowledge come from? Specific associations between print exposure and information acquisition. *Journal of Educational Psychology, 85,* 211–229.

Stanovich, K. E., & Siegel, L. S. (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.

Stanovich, K. E., Siegel, L. S., & Gottardo, A. (1997). Converging evidence for phonological and surface subtypes of reading disability. *Journal of Educational Psychology, 89,* 114–127.

Stanovich, K. E., & West, R. F. (1989). Exposure to print and orthographic processing. *Reading Research Quarterly, 24,* 402–433.

van den Bos, K. P., Lutje Spelberg, H. C., Scheepstra, A. J. M., & de Vries, J. (1994). *De Klepel. Vorm A en B. Een test voor de leesvaardigheid van pseudowoorden. Verantwoording, handleiding, diagnostiek en behandeling* [The Klepel: Form A and B. A test for the ability to read pseudowords: Justification, manual, diagnostics and treatment]. Nijmegen, the Netherlands: Berkhout.

Van den Broeck, W., & Geudens, A. (2009). *An alternative method to the reading level match design in examining a nonword-reading deficit.* Manuscript in preparation.

van der Leij, A., & van Daal, V. H. P. (1999). Automatization aspects of dyslexia: Speed limitations in word identification, sensitivity to increasing task demands, and orthographic compensation. *Journal of Learning Disabilities, 32,* 417–428.

van der Leij, A., van Daal, V., & De Jong, P. F. (2002). Task-related factors in reading efficiency of dyslexic children. In L. Verhoeven, C. Elbro, & P. Reitsma (Eds.), *Precursors of functional literacy* (pp. 229–245). Amsterdam, the Netherlands: Benjamins.

van IJzendoorn, M. H., & Bus, A. G. (1994). Meta-analytic confirmation of the nonword-reading deficit in developmental dyslexia. *Reading Research Quarterly, 30,* 266–275.

Van Orden, G. C. (1987). A ROWS is a ROSE: Spelling, sound, and reading. *Memory & Cognition, 15,* 181–198.

Van Orden, G. C., & Goldinger, S. D. (1994). Interdependence of form and function in cognitive systems explains perception of printed words. *Journal of Experimental Psychology: Human Perception and Performance, 20,* 1269–1291.

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

Vellutino, F. R., & Scanlon, D. M. (1989). Some prerequisites for interpreting results from reading level matched designs. *Journal of Reading Behavior, 21,* 361–385.

Vellutino, F. R., Scanlon, D. M., & Tanzman, M. S. (1994). Components of reading ability: Issues and problems in operationalizing word identification, phonological coding, and orthographic coding. In G. R. Lyon (Ed.), *Frames of reference for the assessment of learning disabilities* (pp. 279–323). Baltimore, MD: Brooks.

Wassenaar, W. A. (1989). *Diagnostische toetsbatterij technische leesvaardigheid. Verantwoording en handleiding* [Diagnostic test battery of decoding ability: Justification and manual]. Nijmegen, the Netherlands: Berkhout.

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.

Wimmer, H. (2006). Don't neglect reading fluency! *Developmental Science, 9,* 447–448.

Wolters, G. E. (2004). *Subsyllabic processing in word reading of Dutch beginning readers.* (Unpublished doctoral dissertation). Leiden University, Leiden, the Netherlands.

Yarkoni, T. (2009). Levenshtein Distance Calculator (Version 0.22) [Computer software]. Retrieved from http://www.talyarkoni.com/materials.php

Yarkoni, T., Balota, D., & Yap, M. (2008). Moving beyond Coltheart's *N*: A new measure of orthographic similarity. *Psychonomic Bulletin & Review, 15,* 971–979.

Ziegler, J. C., & Goswami, U. (2005). Reading acquisition, developmental dyslexia, and skilled reading across languages: A psycholinguistic grain size theory. *Psychological Bulletin, 131,* 3–29.

Ziegler, J. C., & Goswami, U. (2006). Becoming literate in different languages: Similar problems, different solutions. *Developmental Science, 9,* 429–453.

Ziegler, J. C., Perry, C., Ma-Wyatt, A., Ladner, D., & Schulte-Körne, G. (2003). Developmental dyslexia in different languages: Language-specific or universal? *Journal of Experimental Child Psychology, 86,* 169–193.

(*Appendices follow*)

Appendix A

Example of One Minute Test (OMT) and Nonword-Reading Test (Klepel) Items

OMT	Klepel
1. weg	1. kes
11. zoom	11. zoof
21. beweeg	21. betees
31. vervelen	31. vorvaner
41. terugvallen	41. tadugkallem
51. stoof	51. steef
61. sabel	61. saber
71. opsieren	71. olseiret
81. pleister	81. greistir
91. vermelden	91. kormildan
101. opsjorren	101. apsjommen
111. gemoedsrust	111. kewoetspust

Appendix B

Example of Two Minute Test (TMT), Word Recognition Test (WRT), and Exception Word Items

TMT	WRT	Exception words
1. boek	1. zoek	1. jam
11. vul	5. dicht	5. Mike
21. denk	9. spuwt	9. machine
31. fruit	13. beperkt	13. vakantie
41. fles	17. verstekelingen	17. route
51. merk	21. asielzoeker	21. chauffeur
61. nieuws	25. natuurliefhebber	25. cadeau
71. sport	29. symbolisch	29. revolutie
81. schokt	33. chromosoom	33. camping
91. lijmpot	37. recreatiecentrum	37. populair
101. kwartier	40. industrialisering	41. logeren
111. knarst		45. yoghurt