Kindergarten Risk Factors, Cognitive Factors, and Teacher Judgments as Predictors of Early Reading in Dutch

Martine A. R. Gijsel, Anna M. T. Bosman, and Ludo Verhoeven

Abstract

This study focused on the predictive value of risk factors, cognitive factors, and teachers' judgments in a sample of 462 kindergartners for their early reading skills and reading failure at the beginning of Grade 1. With respect to risk factors, enrollment in speech–language therapy, history of dyslexia or speech–language problems in the family, and the role of gender were considered. None of these risk factors were significantly related to reading performance. Cognitive factors in this study included letter knowledge, rapid naming ability, and nonword repetition skills. Of these skills, letter knowledge seemed to have the highest correlation with reading. Kindergarten teachers' judgments, including a task assignment scale and teachers' predictions, demonstrated a significant relationship with reading. Finally, to judge whether these predictors could identify reading disabilities, the discriminatory power of all predictors was assessed and appeared to be insufficient. Implications for screening purposes are discussed.

Factors that have been demonstrated to correlate with reading performance in previous research include risk factors, cognitive factors, and teachers' judgments. If a significant relationship between a particular factor and reading performance has been established, the factor is often identified as a predictor of reading ability. Predictors are referred to as *concurrent predictors* if the relationship between the factor and reading is demonstrated at the same point in time. In contrast, in *longitudinal predictors*, the relationship is demonstrated over a certain time interval. In this case, assessments are usually made in kindergarten in order to predict reading skills in higher grades. In the current study, we expand on the existing research and focus on longitudinal predictors of reading performance in Dutch first graders.

For screening purposes, the prediction of reading disabilities (RD) is more important than the prediction of reading ability in general. A teacher wants to know which child is at risk for reading problems in order to prevent difficulties and also to take appropriate preventive measures. Therefore, it is important to know the critical predictive factors for reading failure. These factors could be useful in a screening battery for subsequent reading skills. In this study, we will investigate the importance of discriminating between the prediction of reading abilities in general and the prediction of reading difficulties in particular. In short, in addition to the relationship between reading ability and risk factors, cognitive factors, and teachers' judgments, we will explore the discriminatory power of these predictors.

Predictive Factors for Reading Performance

With respect to the prediction of reading skills, we explored the relationship between reading performance and risk factors (i.e., dyslexia, speech–language disorders in relatives, specific language impairment, and gender), cognitive factors (i.e., letter knowledge, rapid naming ability, and nonword repetition), and teachers' judgments. With respect to cognitive factors, we included a letter knowledge test, because letter knowledge has been found to be the best predictor of reading performance (Blaiklock, 2004; Braams & Bosman, 2000; Catts, Fey, Zhang, & Tomblin, 2001; Gallagher, Frith, & Snowling, 2000; Hammill, 2004; Pennington & Lefly, 2001; Scarborough, 1998). In addition to letter knowledge, phonological processing skills have also been shown to correlate with subsequent reading skills. Wagner and Torgesen (1987) distinguished three types of phonological processing abilities: phonological awareness, phonological recoding in lexical access (long-term memory), and phonetic recoding to maintain information in working memory (short–term memory). Although most research has aimed at investigating phonological awareness, Scarbor-

ough (1998) pointed out that phonological awareness tests appear to be more successful in predicting future superior reading than future reading problems. Moreover, Blaiklock (2004) suggested that letter knowledge and phonological awareness show substantial overlap in the explained amount of variance in reading (for relations between letter knowledge and phonological awareness in preschoolers, see Johnston, Anderson, & Holligan, 1996). Therefore, instead of measuring phonological awareness, it was decided to assess phonological memory by means of a nonword repetition test and to test phonological naming by means of a rapid serial naming test. We decided on these tests because little is known about the relative impact of these abilities on predictions of Dutch students' reading abilities.

Risk Factors

A family history of dyslexia or speech–language disorders in relatives is often considered to negatively affect students' performance on language and reading tests. Most studies of genetic factors in reading performance have been concerned with the prevalence of dyslexia in relatives. Snowling, Gallagher, and Frith (2003) showed that 66.1% of the students with at least one family member with dyslexia experienced reading problems at the age of 8. Lewis, Freebairn, and Taylor (2000) also showed a trend for a family history of RD to predict reading failure. Moreover, a family history of reading problems significantly predicted spelling impairment. A family history of speech–language disorders as a prospective risk factor for reading has been studied less frequently. Lewis et al. (2000) showed a moderate association between a family history of RD and reading impairment at school age if this variable was assessed in a dichotomous fashion (i.e., positive vs. negative). When it is was coded as an ordinal variable (i.e., the number of nuclear family members affected) no significant effect emerged.

A second variable that possibly constitutes a risk factor for RD is speech and language characteristics. In most studies, poor language or speech characteristics have been diagnosed by means of test assessment (see, e.g., de Jong, & van der Leij, 1999; Menyuk et al., 1991; Scarborough, 1990). Usually, the results of speech–language tests are related to reading performance. An alternative way to assess information about speech and language characteristics is to investigate the history of speech–language therapy. Weiner (1985) found a high incidence of later reading problems in preschoolers who were enrolled in language therapy.

Third, differences in gender may play a causal role in lowering test performance. Petersen (2002) demonstrated that although boys showed a significantly higher score on vocabulary and nonverbal IQ at the beginning of kindergarten, they showed lower scores on reading performance in Grade 2. Badian (1999) indicated that girls were significantly better at reading comprehension than boys. No such differences were found for listening comprehension. Blonk and Bosman (2003) found that teenage girls in the first year of secondary school performed better than teenage boys on reading and language skills and, more important, that teenage girls with dyslexia scored better than teenage boys with dyslexia. Flynn and Rahbar (1994) tested 708 students in Grades 1 and 3 on reading achievement and categorized the students as having severe RD (< 10th percentile), RD (11th to 30th percentile), or typical reading abilities. In the severe RD category, the ratio of boys and girls was 1.4 to 1 in first grade and 1.3 to 1 in third grade. In the RD category, the number of boys and girls was approximately the same in each grade.

Cognitive Factors

Letter Knowledge. In the domain of cognitive abilities, letter-name knowledge is a factor that has been studied frequently in relation to reading skills. The combined results of three meta-analyses demonstrated that together with reading itself and knowledge about writing convention, letter knowledge (r = .52) was the best predictor of reading (Hammill, 2004). Other studies of both English- and non–English-speaking populations have yielded similar results. De Jong and van der Leij (1999) demonstrated the high predictive value of letter knowledge in the Dutch language. In their test, they presented five letters used relatively frequently in Dutch books. On the productive letter test, both letter names and letter sounds were considered correct. The correlation between receptive and productive letter knowledge in kindergarten and word decoding at the beginning of Grade 1 was moderate (r = .39 and r = .51, respectively). At the end of Grade 2, however, the correlations between letter knowledge in kindergarten and word decoding in Grade 2 had decreased dramatically. Braams and Bosman (2000) also used a letter naming task in kindergarten to predict the reading and spelling ability of Dutch students in Grade 1. At the beginning and at the end of kindergarten, students had to name 20 letters. Letter knowledge tested at the beginning of kindergarten revealed a moderate correlation with reading performance at the middle of the curriculum in Grade 1 (r = .44), but the correlation between letter knowledge in kindergarten and reading declined with increasing reading experience to .36 at the end of Grade 1.

Rapid Naming. Another factor often related to the prediction of reading performance is rapid naming (e.g., Allor, 2002; Blachman, 1984; Cornwall, 1992; Hammill, Mather, Allen, & Roberts, 2002; Kirby, Parrila, & Pfeiffer, 2003). In rapid naming tests, participants are asked to name a set of items (usually pictures, colors, letters, or digits) as quickly and accurately as possible. The strength of the relationship between rapid naming and reading performance is dependent on the age

of the participants. Kirby et al. (2003) demonstrated that naming speed in higher grades had much stronger effects on reading than in kindergarten and Grade 1, with the latter grades showing significant, albeit weak, effects. De Jong and van der Leij (1999) also reported small effects of rapid naming ability in kindergarten on reading skills in Grade 1. Van den Bos, Lutje Spelberg, and Eleveld (2004) tested kindergartners on rapid naming (colors and pictures) and visual matching, and they studied the predictive power of these tests for reading ability in Grade 1. They demonstrated a moderate but significant multiple correlation (r = .52). Note that this correlation is even higher than the one between letter knowledge in kindergarten and reading ability in Grade 1 as reported by Braams and Bosman (2000). Hammill et al. (2002) also investigated the importance of rapid naming ability in predicting word identification. In their study, the correlation between rapid naming (letters and words) and word identification was .52. In a study by Cornwall (1992), the correlation between rapid naming and the identification of individual words was .49 for letters and .19 for colors. It thus seems that the strength of the relationship between rapid naming and reading also depends on the kind of stimuli used in the test.

Nonword Repetition. Another factor that is often associated with reading ability is nonword repetition. In nonword repetition tests, participants are asked to repeat a set of nonwords. The test items generally obey the phonological structure of the language. A number of studies have demonstrated that students with RD perform relatively badly on a test for repeating nonwords (Gallagher et al., 2000; Lewis et al., 2000; Muter & Snowling, 1998; Snowling, 1981; Snowling, Goulandris, Bowlby, & Howell, 1986; Snowling et al., 2003). These findings are often explained in terms of deficits in phonological working memory, integrity of phonological representations, or phonological decoding. In the study of Lewis et al. (2000), all participants were enrolled in speech–language therapy. Reading impairment in this group of students was associated with their scores on nonword repetition in kindergarten. Bishop (2001) also concluded that in children with SLI, deficits in nonword repetition and poor literacy skills were correlated, suggesting that the same genes were responsible for both deficits. Muter and Snowling (1998) demonstrated that the strength of the correlation between nonword repetition as a longitudinal predictor and reading performance at the end of Grade 5 was significant in a typical population—that is, in children without SLI (rs ranging from .31 at the age of 4 to .53 at the age of 6).

Teachers' Perceptions and Predictions

Finally, we were interested in the predictive value of the perceptions and predictions of kindergarten teachers. In a study by Kenny and Chekaluk (1993), teachers' questionnaire scores and teachers' category ratings (advanced, average, or poor reading) showed good predictive value for auditory conceptualization scores and reading achievement. Flynn and Rahbar (1998) demonstrated that teachers correctly predicted 64% of poor readers and missed 36% of those students who failed in reading. In their study, the correspondence between teacher ratings and students' scores was low, except for scores in the area of letter-sound knowledge, and it appeared that a combination of test performance scores and teacher information was most effective with respect to the prediction of reading failure.

Discriminatory Power of Predictive Factors

From a scientific point of view, it is interesting not only to acquire knowledge about predictors of reading performance in general, but also to obtain an accurate prediction of the likelihood of developing RD. Therefore, we were interested in the discriminatory ability of risk factors (e.g., dyslexia or speech–language disorders in relatives, enrollment in speech and language therapy, gender), cognitive factors, and teachers' predictions and perceptions. In other words, we wanted to know whether these variables could predict a child's likelihood of developing RD in Grade 1. A moderate or high correlation demonstrating the predictive validity of a factor does not imply that this factor will discriminate well between students who will develop RD and those who will not.

In the current study, children who performed below the cutoff score—that is, below the 25th percentile—in reading after 2 months of formal reading instruction were considered children with RD. Both practical and theoretical evidence demonstrate the validity of early assessment and stability of reading scores for Dutch students throughout all grades. Verhoeven and van Leeuwe (2003) categorized 2,873 students into five groups at different reading levels, based on their reading performance after 3 months of formal reading instruction in Grade 1. These students showed quasi-stable mean levels of reading skills through Grade 6 for one-syllable CVC words (C = consonant; V = vowel). Bast and Reitsma (1998) reported a similar outcome in 280 Dutch students—that is, children who were diagnosed as poor readers after 3 months of reading instruction remained poor readers during the first three grades.

Moreover, the reading-acquisition rate in consistent alphabetic orthographies such as Finnish (see, e.g., Aro, 2004; Lyytinen et al., 2004) is different from the rates reported for English—a language with a highly inconsistent orthography. In Finnish, one of the most consistently spelled alphabetic languages, the majority of children have acquired accurate decoding skills after the first semester in Grade 1. Thus, RD can be identified much earlier than in English. The same point can be made

for Dutch; the Dutch orthography is also relatively consistent in its spelling–sound relationships. After a very short time (within 4 months, before Christmas), children in first grade have learned all the prototypical grapheme–phoneme correspondences. Thus, after a relatively short time, children should be fully aware of the alphabet principle. Moreover, Dutch reading curricula strongly stimulate the awareness of the alphabet principle, because they *all* adhere to the phonics principle.

Moreover, Wentink and Verhoeven (2001) implemented a protocol for the early assessment and intervention of RD, which is now used widely in Dutch primary schools. They recommended assessing reading skills after 6 weeks of formal reading instruction.

Risk Factors

With respect to risk factors, the number of children with speech–language disorders or affected relatives (i.e., with dyslexia or speech–language disorders) or the number of boys in an RD sample may indicate the discriminatory power of these variables to some extent. If a significant proportion of these children at risk is represented in the lower tail of the reading distribution and a very small proportion is represented in the highest tail, the particular risk factor might discriminate quite well between typical children and children with RD. As mentioned earlier, Snowling et al. (2003) showed that 66.1% of the students with at least one family member with dyslexia experienced reading problems at the age of 8. Scarborough (1989, 1991) also demonstrated that family incidence of RD was an accurate predictor of reading ability; outcomes were correctly predicted for 72.6% to 80.6% of the participants, depending on the way in which family incidence was identified (Scarborough, 1989).

Cognitive Factors

Letter knowledge seemed to be a useful factor in predicting reading ability. But the question is, will letter knowledge discriminate between poor readers and typical readers? To answer this question, the number of valid positives and valid negatives versus false positives and false negatives can be calculated. *Valid positive rate* (hits) refers to the number of students who were predicted to have RD who did turn out to be poor readers. *False positive rate* (false alarm) refers to the number of students who were predicted to have RD but who turned out to be typical readers. *Valid negative rate* (correct rejection) refers to the group of students in which RD were not predicted and not observed. Finally, *false negative rate* (misses) refers to the number of students who were predicted to be typical readers but who turned out to have RD. In this way, the correctness of classification (poor vs. typical readers) based on performance on the predictor tests can be evaluated. Muter and Snowling (1998) considered this issue and selected a group of poor readers (reading accuracy scores below 25th percentile) and typical readers (reading scores above the 75th percentile) in their sample. They used rhyme detection, phoneme deletion, nonword repetition, and letter-name knowledge as predictors. Scores were obtained at the ages of 4, 5, and 6 years. These four variables did not classify the students well at all test moments. At the age of 4, no significant predictor set was found. When test scores were obtained at the age of 5 or 6, 80% of the students were classified correctly when rhyme detection was eliminated. Analyses revealed that phoneme deletion and nonword repetition were the best predictors. Letter knowledge was a significant predictor as well, but its discriminatory power was relatively low. These results should be interpreted carefully—first, because their sample size was quite small (only 20 children), and second, because high correlations do not imply high discriminatory power, and the latter is desirable for a screening battery.

Hammill et al. (2002) also assessed the discriminatory power of tests. For this purpose, participants were divided into two groups (scoring above or below the 25th percentile) based on performance variables. Levels of agreement (i.e., the score on a test or subgroup of tests correctly identifies the level of word identification) were calculated. Results indicated 26 out of 56 false positives (46%). The percentage of false negatives was 15%; none of the test variables in the study (semantics, grammar, phonology, rapid naming, and rapid marking) were effective predictors of poor reading. With respect to nonword repetition, Muter and Snowling (1998) showed that scores on nonword repetition and phoneme deletion obtained at the age of 5 or 6 significantly discriminated between good and poor readers.

Teachers' Perceptions and Predictions

Satz and Fletcher (1988) reviewed a set of major studies regarding teachers' predictions. A comparison between kindergarten teachers' predictions and test results in Grades 1 and 2 revealed that overall hit rates were almost identical for test results and teacher predictions. Kenny and Chekaluk (1993) studied the utility of teachers' perceptions and test outcomes from kindergarten through Grade 2. Teachers completed a questionnaire and had to categorize the students into advanced, average, and poor readers. In teachers' classifications, false positive rates (ranging from .30 in kindergarten to .13 in Grade 2) were higher than false negative rates (.23, and .06, respectively). This means that the number of students who were predicted to have RD but who turned out to be typical readers was higher than the number of students who were predicted to be typical readers but who turned out to have RD.

Flynn and Rahbar (1998) presented teachers with a rating scale with behavioral descriptions reflecting 10 kindergarten skills and assessed their predictive value for reading skills in third grade. Teachers correctly classified 64% of the poor readers, in contrast with a screening test, which

yielded 80% valid positives. A combination of teachers' ratings and the screening test resulted in 88% valid positives. However, the number of false positives increased from .23 (teachers' ratings alone) to .39, which indicated an overidentification of students at risk.

In addition to level of agreement, the sensitivity, specificity, and positive predictive value can be calculated to explore the practical value of a factor. The *sensitivity* index reflects the ability of a test to correctly identify individuals who have the disorder. The *specificity* index reflects the ability of a test to correctly identify individuals who do *not* have the disorder. The *positive predictive value* reflects the proportion of valid positives among all individuals whom the screening measure identifies as at risk. Petersen (2002) demonstrated a better sensitivity of predictors for advanced reading than for poor word reading.

To sum up, in the present study, we seek answers to the following two questions:

1. What is the relationship between risk factors, cognitive factors, and teachers' predictions and perceptions on the one hand and reading performance on the other hand?
2. How well do these factors identify students who will develop RD?

Method

The data collection in this study comprised a kindergarten test battery, two questionnaires designed for the parents and teachers of the kindergartners, and a Grade 1 test to assess reading skills.

Participants

Of the 462 students participating in this study, 241 were boys (52.2%) and 221 were girls (47.8%) from 20 general education primary schools in the Netherlands. All students were first tested when attending kindergarten, and their mean age was 70.8 months (SD = 4.4). In Grade 1, when they were tested a second time, the mean age was 79.1 months (SD = 4.4).

Questionnaires

Parents' Questionnaire. This questionnaire consisted of four questions concerning the native language of the child, the child's enrollment in speech or language therapy, and the presence or absence of dyslexia or speech–language difficulties in relatives. These questions had to be answered by marking the option *yes* or *no*. If the answer was yes, further information was required (e.g., which of the first-degree relatives had dyslexia or speech–language problems, or what the child's native language was if not Dutch). The parents of all kindergartners received a questionnaire.

Teacher Questionnaire. This questionnaire consisted of two parts. In Part 1, teachers were asked to write down the names of students whom they believed would develop reading or spelling difficulties in Grade 1. With respect to the analyses, this variable was dichotomous (*yes* vs. *no*). Part 2 was a subscale of the *Task Assignment Scale* (*Werkhoudingslijst*) of van Doorn (1996) and consisted of 10 items concerning concentration, motivation, and attitude of the child. The teacher was asked to make judgments for each of the students in their classroom about the frequency of the described behavior on a 5-point scale, ranging from *never* to *always*. Examples are "the child tries to solve problems by himself or herself," "the child needs encouragement during work," and "the child starts working immediately after instruction." With respect to the analyses, scores on task assignment were assigned to the answers and categorized either in the lowest quartile (scores below the 25th percentile) or in the highest quartile (scores above the 75th percentile).

Kindergarten Test Battery

Nonword Repetition Test. The *Nonword Repetition Test* (*Nonwoord Repetitietaak*; Irausquin, 1999) required the repetition of 22 pronounceable nonwords. To control for possible articulation errors or hearing problems of the child, the test was preceded by 15 real words. Thus, errors in nonword repetition (e.g., substitutions due to certain articulation errors) that already occurred in real word repetition were counted as correct. In case of hearing disorders, test assessment was stopped. Performance on real words was not included in the test score; only repetition of the nonwords determined performance. The length of the items increased, ranging from one to four syllables. Nonword examples are *kloda* and *bledistot*. Three practice items preceded the nonwords. The experimenter pronounced each word and nonword with a piece of paper in front of the mouth in order to prevent the use of visual information. The child was asked to repeat the words and nonwords after the experimenter. With respect to the analyses, the test score was the number of nonwords correctly repeated. The maximum score on the test was 22.

Rapid Naming Picture Test. The *Rapid Naming Picture Test* (*Benoemtaak Plaatjes*; van den Bos, 2004) consisted of 5 different pictures (tree, duck, chair, scissors, and bicycle) depicted in five columns of 10 pictures each, yielding a total number of 50 pictures to be named. Each picture was presented 10 times at random positions. The child was asked to name all the pictures as fast as possible and as correctly as possible in a vertical direction (i.e., column by column). The test was preceded by a short training, in which the child had to name the pictures in the final column. The time to complete the task and the number of errors were registered. With respect to the analyses, the mean time to name one picture (correctly or incorrectly) was computed.

The number of errors was not included as a variable because few errors were made.

Rapid Naming Colors Test. The *Rapid Naming Colors Test* (*Benoemtaak Kleuren*; van den Bos, 2004) consisted of five columns of 10 colored squares (black, yellow, red, green, and blue). The test procedure and the scoring were exactly the same as with the *Rapid Naming Picture Test*. Note that we did not use alphanumeric stimuli because the students had not received formal reading and writing instruction yet.

Letter Knowledge Test. The *Letter Knowledge Test* considered productive letter knowledge as well as receptive letter knowledge. First, productive letter knowledge was tested. This test consisted of all 26 letters of the alphabet, presented on a card in six rows of 4 letters and one row of 2 letters. The letters were typed in Helvetica regular font, size 24. If the capital letter had a different shape than the lowercase letter (which is the case for most letters), the capital letter was presented just below the lowercase letter. Thus, all of the letters had two presentation forms, except the letter *a*, which had three different representations (a, A, and ɑ). Three different orders of letters were created. The child was presented one row of letters at a time; a sheet of paper covered the other letters. The child had to name the letters in a horizontal direction. Responses were considered correct if the child named either the lowercase letter or the capital letter correctly; both letter names and letter sounds were considered correct. In case of doubt about the production of a letter, the child was asked to name a word with that letter in the initial position. Uncertainty only emerged in the distinction between voiced sounds and their voiceless counterparts like *s*/*z* and *f*/*v*. When the child did not know a certain letter, he or she was allowed to guess and continue with the next letter.

Subsequently, receptive letter knowledge was tested. The child was presented with the same card. The experimenter named a letter sound (in random order), and the child was asked to point to that letter on the card. Guessing was permitted. With respect to the analyses, two variables were computed, namely, the score on the productive and the score on the receptive letter knowledge test. The test score was the total number of correct responses. The maximum score on each test was 26.

Grade 1 Reading Test

The *Word Reading Test* (*Toets Woorden Lezen*; Wentink & Verhoeven, 2001) consisted of three parts, each of them including 10 (C)VC words. The first part consisted of well-known words that had been taught in the classroom. The second part consisted of new words that differed by one letter from words that had been taught in the classroom. In this way, the items resembled the words that had already been taught. The third part consisted of new words that differed by two letters from the words taught in the classroom. In this condition, there was hardly any resemblance to the words that had already been taught. The child was asked to read aloud the words as fast and as accurately as possible. For each part, the test was stopped when the child failed on three consecutive items. With respect to the analyses, the number of items read correctly and the time needed to perform each part of the test were converted into the number of items read correctly in one minute. The first part of the test (decoding well-known words) was not considered in the analyses. Therefore, the results of the second and third part are henceforth referred to as Word Reading 1 and 2, representing the ability to decode new words.

Procedure

Students were tested in kindergarten and at the beginning of Grade 1. In kindergarten, no structural (reading) program was used and, in most schools, there was no structural remediation in kindergarten concerning reading or reading-related skills. In Grade 1, the majority of the children were instructed with *Veilig Leren Lezen* (*Learning to Read Safely*), the most widely used reading program in Dutch schools (Mommers, Verhoeven, & van der Linden, 1979, 1994). The emphasis in this method is on the structure of the orthographic system and the relationship between letters and sounds (i.e., phonics). Initially, only consistently spelled words are used. After 4 months of instruction, the children are familiar with the main grapheme–phoneme correspondence rules. It is a fairly rigid, preprogrammed curriculum, which imposes a strict day-by-day and week-by-week progression.

In kindergarten, the students were tested halfway through the curriculum (January–March, 2002; henceforth Time 1; see Note 1). This took about 20 min for each child. The order of presentation of the tests varied among students, such that each test was administered equally often first, second, third, and so forth. At the end of kindergarten (May, 2002), the parents' questionnaires were provided to the teachers with the request to present them to the parents. Parents were asked to complete the list and return it in 2 weeks. Furthermore, the teachers' questionnaires were provided. In Grade 1, all students were tested 8 weeks after formal reading instruction had started (Fall 2002; henceforth Time 2). All students completed the *Word Reading Test*. In Grade 1, additional instruction or intervention was implemented at the earliest after the first test assessment—that is, after 2 months of formal reading instruction. This coincided (not accidentally) with the time that RD were diagnosed.

Analysis

To evaluate the predictive value of cognitive factors for reading performance, a structural analysis was performed.

Furthermore, a multiple regression was carried out to determine the best combination of variables to predict reading. To investigate the relationship between risk factors and teachers' judgments, *t* tests were performed on scores on the *Word Reading Test* in Grade 1. To investigate the discriminatory power of all variables, percentages of valid and false positive and negative outcomes were calculated. Moreover, the sensitivity and specificity indexes were computed. For these analyses, the outcome had to be labeled dichotomously (i.e., developing RD or not). We defined students with RD as students with reading scores below the 25th percentile, and good readers as students demonstrating reading scores above the 75th percentile. The reading score was the mean number of items read correctly in one minute on *Word Reading Test* 1 and 2. We chose the 25th percentile to represent the group of students that performed below standard, because this criterion is used in standardized Dutch. Finally, we performed a discriminant function analysis in order to establish which combination of variables discriminated best between poor and typical readers.

Results

Sample Characteristics

The majority of students (95%) had Dutch as their native language; 22 students (5%) were originally from other countries (both in and outside Europe) and had Dutch as their second language or were bilingually educated. The number of reported speech and language problems was 71 (15.4%), whereas 6.3% of the parents of all participants had coped with speech or language problems. Finally, 6.9% of the parents reported a history of dyslexia in relatives—that is, a first- or second-degree family member with dyslexia. The lack of response on the questionnaires was very low: Of the parents' questionnaires, 400 copies (86.6%) were returned, and of the teachers' questionnaires, 439 copies (95%) were returned.

Statistics

Because 22 students in our sample were native speakers of a language other than Dutch, we first established whether the performance of these students differed from that of their peers whose mother tongue was Dutch. A one-way ANOVA on the mean scores of all tests in kindergarten and on the mean scores of the *Word Reading Test* in Grade 1 indicated that students whose native tongue was not Dutch did not differ significantly from students with Dutch as native language (all $ps > .05$). Therefore, it was decided to include in subsequent analyses all students who attended kindergarten and participated in the *Word Reading Test* at the beginning of Grade 1. Students who had to repeat kindergarten or who repeated Grade 1 were excluded. Despite this exclusion criterion, a few missing values were inevitable. To resolve this issue, we made use of the "nearest neighborhood" method to estimate the missing scores.

Predictive Factors for Reading Performance

Table 1 shows the mean scores and standard deviations on all tests in kindergarten and Grade 1. Scores were computed for all students together and for poor readers and typical readers separately. Table 2 shows the characteristics for the poor reading group and the typical reading group. To explore the predictive value of the reported tests, a structural analysis was performed. In this analysis, all variables were included. Figure 1 depicts all variables and standardized regression weights. The fit of the model was good, $\chi^2(9, N = 462) = 16.2$, $p = .06$, $GFI = .99$, $AGFI = .97$, $NFI = .99$, $RMSEA = .04$.

Next, a stepwise multiple regression was carried out in order to find out whether a combination of factors increased the strength of the correlation. The dependent variable was the mean number of words read correctly on *Word Reading Test* 1 and 2. All risk factors, cognitive factors, and teachers' predictions and perceptions were included as independent variables. Receptive letter knowledge was excluded from the analysis because of its strong correlation ($r = .96$) with productive letter knowledge. The results showed that productive letter knowledge was the best predictor of reading performance, $r = .53$, $p < .001$. A second factor that contributed significantly to the variance was the rapid naming of colors. If this variable was added to letter knowledge, 29% of the variance was explained ($r = .54$, $p < .001$). Despite these two factors, no other predictor turned out to be significant. These results support the outcome of the structural analysis. Furthermore, to investigate the relationship between risk factors and teachers' predictions and perceptions and reading ability, *t* tests were performed. Table 3 shows that only the teachers' predictions and perceptions significantly discriminated students' scores on word reading.

To further investigate the role of the teacher, scores on all tests were compared between students with positive and negative predictions (see Table 4) and between students with scores within the lowest quartile on task assignment and those with scores within the highest quartile (see Table 5). For screening purposes, it would be useful to know whether the students with negative ratings represented the lower tail of the distribution of reading scores. Analyses revealed that this was indeed the case. About half of the students who were assigned to the lowest task assignment group and who were predicted to develop RD were represented in the 25th reading percentile. The scores of the students without predicted RD were more widely distributed.

Discriminatory Power of Predictive Factors

To answer our second question concerning the discriminatory power of

TABLE 1
Mean Scores and Standard Deviations in Kindergarten and Grade 1

Grade/test	All students (N = 462)		Lowest reading quartile (n = 118)		Highest reading quartile (n = 123)		
	M	*SD*	*M*	*SD*	*M*	*SD*	*d*
Kindergarten							
Rapid Naming							
Colors	1.6	0.5	1.9	.6	1.4	.4	.97
Pictures	1.6	0.4	1.8	.4	1.4	.4	.96
Letter Knowledge							
Productive	10.9	7.8	5.4	5.3	17.5	6.8	1.99
Receptive	10.5	7.4	5.4	4.8	16.7	6.8	1.92
Nonword Repetition	16.8	2.8	16.2	2.9	17.2	3.0	.35
Grade 1							
Word Reading 1							
Score	7.5	2.4	4.5	2.3	9.3	1.0	
Time (sec)	54.7	30.7	92.2	25.6	22.4	8.6	
wpm	13.8	14.1	3.2	2.1	30.4	17.7	2.16
Word Reading 2							
Score	7.3	2.9	3.4	2.5	9.1	1.4	
Time (sec)	59.9	31.5	94.9	29.6	27.9	11.7	
wpm	11.9	13.0	2.1	1.7	26.0	17.6	1.91
Word Reading 1/2							
Score	7.4	2.5	3.9	2.1	9.2	1.1	
Time (sec)	57.3	29.4	93.6	22.6	25.1	9.4	
wpm	12.8	13.1	2.7	1.6	28.2	16.7	2.15

Note. The scores on Rapid Naming are the number of seconds to name an object (colors/pictures). The scores on Letter knowledge (maximum score = 26) and Nonword repetition (maximum score is 22) are the number of items responded correctly. Word Reading score is the number of items read correctly (maximum score = 10). Wpm is the number of words read per minute.

the variables, we calculated the number of false positive and valid positive outcomes and false negative and valid negative outcomes. Recall that the false positive rate reflects the number of students who were predicted to have RD but who turned out to be typical readers. The valid positive rate refers to the number of students who were predicted to have RD who indeed turned out to be poor readers. The valid negative rate refers to the group of students for whom RD were not predicted and not observed. Finally, the false negative rate refers to the number of students who were predicted to be typical readers but who turned out to have RD. These percentages were computed for all variables. In this way, we were able to evaluate different predictors of RD. Concerning the cognitive factors, difficulties were predicted if the score on the test was below the 25th percentile. Recall that RD were defined by scores below the 25th percentile on decoding new words (*Word Reading Test* 1 and 2). Next, for all predictors, we calculated the sensitivity and specificity. The sensitivity index is the number of valid positives/(valid positives + false negatives), and the specificity index is the number of valid negatives/(valid negatives + false positives). These data are presented in Table 6. The results

TABLE 2
Frequency Distribution of Poor-Reading (Lowest Quartile) and Superior-Reading (Highest Quartile) Groups With Respect to Risk Factors and Teachers' Predictions

Factor/judgment	All students (N = 462)	Lowest reading quartile (n = 118)	Highest reading quartile (n = 123)
Risk factors			
Speech/language therapy	71	30	11
Dyslexia	32	13	7
SLI	29	11	4
Boy	241	67	62
Teachers' judgments			
Negative prediction	91	50	5
Low task assignment [a]	127	59	15

[a]Lowest quartile task-assignment group.

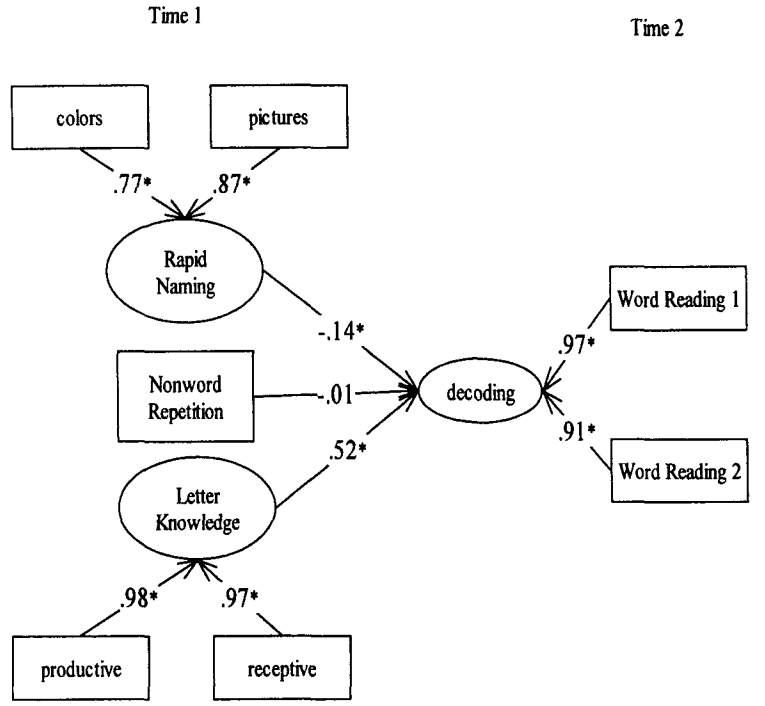

FIGURE 1. Structural model of the relationship between kindergarten test scores and decoding skills in Grade 1 ($N = 462$). *$p < .005$.

TABLE 3
Results of *t* Tests for All Students Concerning Risk Factors and Teachers' Predictions

	Word-Reading Test 1 and 2		
Factor/judgment	***M***	***SD***	***t test***
Risk factors			
Speech/language therapy			
Yes	10.3	14.0	$t(394) = 1.87$, $p = .06$
No	13.5	12.5	
Dyslexia			
Yes	9.17	7.07	$t(366) = 1.81$, $p = .07$
No	13.50	13.36	
SLI			
Yes	8.90	7.27	$t(390) = 1.73$, $p = .08$
No	13.19	13.17	
Gender			
Boy	12.49	13.78	$t(460) = -.59$, $p = .56$
Girl	13.20	12.38	
Teachers' judgments			
Prediction			
Negative	6.3	6.7	$t(306) = 8.21$, $p < .001$
Positive	14.7	13.5	
Task assignment			
Below 25th percentile	8.3	10.6	$t(201) = -6.0$, $p < .001$
Above 75th percentile	18.9	16.7	

demonstrated that teachers' predictions have more predictive power than some cognitive factors (i.e., rapid naming and nonword repetition).

Finally, we submitted all variables to a stepwise discriminant function analysis in order to establish which combination of variables discriminated best between poor readers (reading scores below the 25th percentile) and typical readers. All risk factors, cognitive factors, and teachers' perceptions were included in the analysis. Productive letter knowledge, teachers' predictions, and rapid naming of colors contributed significantly to the classification ($p < .001$). The canonical correlation between these variables and group membership (above or below the 25th percentile) was .49. Standardized discriminant function coefficients reflecting the contribution of each variable showed that productive letter knowledge contributed most to correct classification; rapid naming of colors contributed .36, teachers' predictions .47, and productive letter knowledge –.64. The discriminant analysis had an overall accuracy rate of 70.8%. The number of valid positives was 46.3%, whereas 12% of the children were misclassified (i.e., false negatives). The number of valid negatives was 88%, and 53.7% of the children were predicted to develop RD but turned out to be typical readers (false positives).

Discussion

Our first question concerned the relationship between risk factors, cognitive factors, and teachers' predictions and perceptions on the one hand and reading ability on the other hand. Risk factors turned out to play a minor role in the prediction of the reading performance of Dutch children in Grade 1. Students without speech–language therapy only seemingly outperformed those students enrolled in speech–language therapy in Grade 1; the difference, however, did not reach significance. In kindergarten, students with

a history of speech–language therapy performed significantly worse on all tests. The lower performance on letter knowledge and the lower (though not significant) reading performance in students with speech or language disorders correspond with the findings of earlier studies (e.g., Carroll & Snowling, 2004; Scarborough, 1990). Note that most studies include formal speech or language tests, and few have used enrollment in speech–language therapy as an indicative variable for reading problems. Further research is required to establish the value of this variable in identifying students at risk. We suggest gathering detailed information about the treatment (kinds of problems, duration of the treatment) to increase the accuracy of the identification of students at risk.

Furthermore, and surprising enough, the current study provided no evidence for a hereditary factor in reading or speech–language problems: The test performance of students with and students without a family history of dyslexia or speech–language difficulties was statistically equal for all measures. This result is in contrast with a number of studies performed with English-speaking children (e.g., Pennington & Lefly, 2001; Scarborough, 1989, 1991; Snowling et al., 2003). Scarborough (1998) reviewed several studies and concluded that the family incidence of dyslexia certainly increases the risk for RD in a child. However, the degree of risk varies among studies and may be caused by the way in which information was gathered. In the current study, parents were interviewed by means of a questionnaire. If we had tested the reading ability of the parents and their speech and language performance, our results might have been different. Scarborough (1989) showed that although self-reported RD in parents significantly predicted children's reading performance, predictions were more accurate when test results were used. With respect to gender, the difference between boys and girls did not reach significance for reading performance at the beginning of Grade 1. Scarborough (1998) also concluded in another review that boys are only slightly more at risk for RD than girls.

With respect to cognitive factors, the relative importance of kindergarten test scores for reading performance was analyzed with the use of structural equation modeling. Letter knowledge turned out to be the strongest predictor, followed (remotely) by rapid naming. The important role of letter knowledge in predicting reading per-

TABLE 4
Mean Scores and Standard Deviations of Students With Predicted Reading Difficulties (Negative Prediction) and Students Without Predicted Difficulties (Positive Prediction) by their Kindergarten Teacher

	Students (N = 414)				
	Negative (n = 91)		Positive (n = 323)		
Prediction	*M*	*SD*	*M*	*SD*	*t* test
Kindergarten					
Rapid Naming					
Pictures	1.7	0.5	1.6	.4	$t(412) = -3.30$, $p = .001$
Colors	1.8	0.6	1.6	.5	$t(412) = -3.45$, $p = .001$
Letter Knowledge					
Productive	5.1	5.3	12.4	7.6	$t(206) = 10.48$, $p < .001$
Receptive	5.2	4.9	11.9	7.2	$t(213) = 10.18$, $p < .001$
Nonword Repetition	15.4	3.0	17.1	2.7	$t(412) = 5.38$, $p < .001$
Grade 1					
Word-Reading 2	6.8	6.8	15.8	14.6	$t(324) = 8.27$, $p < .001$
Word Reading 3	5.7	6.8	13.7	13.4	$t(295) = 7.68$, $p < .001$
Word Reading 2/3	6.3	6.7	14.7	13.5	$t(306) = 8.21$, $p < .001$

TABLE 5
Mean Scores and Standard Deviations on All Tests for Both Task-Assignment Groups

	Task assignment				
	Lowest quartile (n = 127)		Highest quartile (n = 121)		
Test	*M*	*SD*	*M*	*SD*	*t* test
Kindergarten					
Rapid Naming					
Pictures	1.74	.51	1.46	.37	$t(231) = -5.02$, $p < .001$
Colors	1.85	.65	1.42	.36	$t(198) = 6.50$, $p < .001$
Letter Knowledge					
Productive	7.1	6.9	14.5	7.5	$t(246) = -8.13$, $p < .001$
Receptive	7.1	6.2	14.0	7.3	$t(236) = -8.05$, $p < .001$
Nonword Repetition	16.0	2.9	17.1	2.5	$t(246) = -3.23$, $p = .001$
Grade 1					
Word-Reading 2	8.9	11.5	19.9	17.6	$t(205) = -5.81$, $p < .001$
Word Reading 3	7.7	10.7	17.9	16.8	$t(202) = -5.71$, $p < .001$
Word Reading 2/3	8.3	10.6	18.9	16.7	$t(201) = -6.00$, $p < .001$

TABLE 6
Percentages of Valid and False Positives and Negatives, Sensitivity and Specificity of Risk Factors, Cognitive Factors, and Kindergarten Predictors

Factor/judgment	Valid positives[a]	False positives	Valid negatives	False negatives	Sensitivity	Specificity
Risk factors						
Speech/language therapy	42.2%	57.8%	78.2%	21.8%	.30	.86
Dyslexia	40.6%	59.4%	77.7%	22.3%	.15	.93
SLI	37.9%	62.1%	75.5%	24.5%	.11	.94
Gender	27.8%	72.2%	76.9%	23.1%	.57	.49
Cognitive factors						
Rapid Naming						
Pictures	36.5%	63.5%	78.1%	21.9%	.36	.79
Colors	38.5%	61.5%	79.1%	20.9%	.40	.78
Letter Knowledge						
Productive	49.2%	50.8%	84.1%	15.9%	.56	.80
Receptive	50.8%	49.2%	83.9%	16.1%	.54	.82
Nonword Repetition	33.3%	66.7%	77.8%	22.2%	.39	.73
Teachers' judgments						
Prediction	55.0%	45.0%	83.0%	17.0%	.48	.87
Task assignment	46.5%	53.5%	82.0%	18.0%	.51	.79

Note. Sensitivity index is the number of valid positives/(valid positives + false negatives), and specificity index is the number of valid negatives/(valid negatives + false positives).
[a]*The percentage of valid positives is equal to positive predictive value.*

formance has been found in previous studies (e.g., Braams & Bosman, 2000; Catts et al., 2001; de Jong & van der Leij, 1999; Gallagher et al., 2000; Hammill, 2004; Pennington & Lefly, 2001). However, the adequacy of this predictor might be restricted to the beginning stage of reading. Walsh, Price, and Gillingham (1988) demonstrated that letter-naming speed correlated much more strongly with reading achievement among kindergartners than among students in Grade 2. Other studies yielded similar results, demonstrating declining correlations between letter knowledge and reading ability in higher grades (Braams & Bosman, 2000; de Jong & van der Leij, 1999; Wesseling & Reitsma, 2000).

The relationship between rapid naming and reading performance was much weaker, although it also reached significance. The predictive value of rapid naming has been demonstrated in several studies. However, the present study only provides weak evidence for the predictive value of rapid naming in kindergarten. The strength of the relationship in the current study would probably have been larger if we had used letters in the test instead of pictures or colors. However, because of the limited letter knowledge of children in kindergarten, this was not possible at the time of testing. Some researchers have suggested that rapid naming is a better predictor of poor reading than it is of typical reading (Hammill et al., 2002; Meyer, Wood, Hart, & Felton, 1998; Petersen, 2002).

The ability to correctly repeat a set of nonwords was not significantly related to reading skills and thus played a minor role in the prediction of decoding skills. The role of nonword repetition in identifying students at risk for RD and the interpretation of the test outcome are not yet quite clear. In English, Snowling and her colleagues (Gallagher et al., 2000; Muter & Snowling, 1998; Snowling, 1981; Snowling et al., 1986; Snowling et al., 2003) demonstrated that poor readers scored significantly worse on the repetition of nonwords than controls. However, effects were influenced by phonological complexity (group differences were most apparent when four-syllable nonwords had to be repeated). It might be that our selection of items was not appropriate, because we included relatively many nonwords with a low level of phonological complexity.

Next, we were interested in whether kindergarten teachers' predictions and perceptions (ratings of task assignment) would be predictive for reading outcome. Several studies (e.g., Coleman & Dover, 1993; Flynn & Rahbar, 1998; Kenny & Chekaluk, 1993; Taylor, Anselmo, Foreman, Schatschneider, & Angelopoulos, 2000; Teisl, Mazzocco, & Myers, 2001) have demonstrated significant relations between teachers' ratings and future school performance. The results of the current study supported these findings and demonstrated that test performance supported the predictions of teachers quite well. Both in kindergarten and in Grade 1, the performance of students who were expected to develop RD was lower than the performance of students with positive predictions of reading ability. This pattern was clear in all tests. The role of kindergarten teachers' predictions may be under-

estimated because students who repeated kindergarten were excluded from the analyses. Those students were probably expected to develop RD in Grade 1 with great certainty. The power of teachers' judgments was further demonstrated by the results of the task assignment scale. Kindergartners who were assigned to high levels of task assignment performed significantly better (both in kindergarten and Grade 1) than students with moderate or low task assignment. This result corresponds with that of Kenny and Chekaluk (1993), who reported a significant relation between teachers' perceptions and reading achievement. Flynn and Rahbar (1998) reported high correspondence between teacher ratings and letter-sound knowledge, but low correspondences between teacher ratings and all other measures (vocabulary, syntax, visual discrimination, form copy). In sum, the present findings demonstrated the usefulness of teachers' predictions and ratings for classifying children at risk for RD.

The second question concerned the discriminatory ability of predictive factors. Rather than looking at the ability to predict actual scores, we examined the ability to predict whether a child will fail in reading or not. This should be an import property of a screening battery. To evaluate this capacity, calculations of the sensitivity index, specificity index, and positive predictive value were performed. For practical use in a screening battery, these indexes need to reach a minimum value of .75 (Hammill et al., 2002). In our study, only the specificity index matched this criterion. This means that most of our predictors were only able to correctly identify those students who did *not* develop RD. For all predictors, the sensitivity was much lower than the specificity, which is a general trend in prediction studies (e.g., Hammill et al., 2002; Pennington & Lefly, 2001; Petersen, 2002; Schneider & Näslund, 1993; Teisl et al., 2001; see Scarborough, 1998, for an overview). Thus, in general, the number of false positives and false negatives was too high. In particular, the large number of false positives—that is, the percentage of students who were predicted to have RD but who turned out to be typical readers (ranging from 45% to 72.2%)—was troublesome. Several other studies also demonstrated the occurrence of too many false positives (e.g., Catts et al., 2001; Flynn & Rahbar, 1998; O'Connor & Jenkins, 1999; Taylor et al., 2000) or false negatives (e.g., Coleman & Dover, 1993; Hammill et al., 2002; Mantzicopoulos & Morrison, 1994). In contrast to the findings of other studies, RD were identified very early in Grade 1—that is, after 2 months of formal reading instruction. However, as noted in the introduction, results from other Dutch studies demonstrated strong stability of reading scores during primary school (Bast & Reitsma, 1998; Verhoeven & van Leeuwe, 2003). Thus, the negative results in our study are unlikely to be due to early assessment. In sum, risk factors, cognitive factors, and teachers' perceptions were not sufficiently adequate on their own to correctly identify those students who exhibited RD.

The results of a discriminant function analysis, however, demonstrated that a combination of productive letter knowledge, rapid naming of colors, and teachers' predictions increased the accuracy of prediction to an overall accuracy rate of 70.8%. These results are consistent with the results of Pennington and Lefly (2001), who performed a discriminant function analysis and demonstrated that letter-name knowledge and rapid serial naming of colors and objects were most important in predicting RD.

In conclusion, the results of this study suggest that group membership (RD or not) in Dutch students at the beginning of Grade 1 can be moderately predicted in kindergarten. Although letter knowledge seems to be a strong correlate of word reading, it cannot on its own correctly identify students who will develop RD. To improve the accuracy of classification, a combination of variables is needed (see also Scarborough, 1998). Surprising enough, teachers' predictions and perceptions seem at least as effective as cognitive factors and might, therefore, contribute to accurate prediction. A combination of productive letter knowledge, teachers' predictions, and rapid naming of colors accurately classified 71% of the children. Unfortunately, however, too many students who were predicted to have RD turned out to be typical readers, and too many poor readers were not identified in kindergarten. These results suggest that kindergarten measures moderately predict subsequent reading skills in Dutch children. However, the results might have been different if, in addition to letter knowledge, the battery had included other measures, such as one or more tests of phonemic awareness and a test of verbal memory. In the Netherlands, a great deal of attention is already focused on the early screening and prediction of RD, especially since the implementation of the *Protocol Leesproblemen en Dyslexie* (*Protocol Reading Problems and Dyslexia*; Wentink & Verhoeven, 2001), which includes a checklist for kindergartners. Letter knowledge, phonological awareness, and phonemic awareness are part of this checklist. In addition to the assessment of letter knowledge, we recommend the use of kindergarten teachers' predictions. These two measures are low cost and accurate and require only a limited amount of time. Finally, reading skills in Grade 1 have to be assessed as early as possible–that is, after 2 months of formal reading instruction, as recommended by Wentink and Verhoeven (2001). After all, the best predictor of future reading is reading itself (see, e.g., Hammill, 2004).

ABOUT THE AUTHORS

Martine A. R. Gijsel *is a certified speech and language pathologist and a PhD student in the Department of Special Education, Radboud University Nijmegen. Her research interests include prediction and prevention of reading difficulties, and deficiencies in speech and language development.* ***Anna M. T. Bosman*** *is associate professor in the Department of Special Education, Radboud University Nijmegen. Her*

research interests include reading, spelling, dyslexia, intervention and dynamic systems. ***Ludo Verhoeven*** *is professor in the Department of Special Education, Radboud University Nijmegen. His research interests include reading and language development in L1 and L2 learners. Address:*

AUTHORS' NOTE

We are greatly indebted to all children and teachers of the schools that participated in our study, and we want to express our appreciation to all students who assisted us in collecting the data. We are also very grateful to Jan van Leeuwe for his statistical support. Finally, we want to thank three anonymous reviewers for their helpful comments on an earlier version of the article.

NOTE

Near the end of kindergarten, children were tested on letter knowledge for a second time. Because correlations were as high as .92 for both productive and receptive letter knowledge, however, this measurement was not included in the analysis.

REFERENCES

Allor, J. H. (2002). The relationships of phonemic awareness and rapid naming to reading development. *Learning Disability Quarterly, 25,* 47–57.

Aro, M. (2004). *Learning to read; The effect of orthography.* Jyväskylä, Finland: University of Jyväskylä.

Badian, N. A. (1999). Reading disability defined as a discrepancy between listening and reading comprehension: A longitudinal study of stability, gender differences, and prevalence. *Journal of Learning Disabilities, 32,* 138–148.

Bast, J., & Reitsma, P. (1998). Analyzing the development of individual differences in terms of Matthew effects in reading: Results from a Dutch longitudinal study. *Developmental Psychology, 34,* 1373–1399.

Bishop, D. V. M. (2001). Genetic influences on language impairment and literacy problems in children: Same or different? *Journal of Child Psychology and Psychiatry and Allied Disciplines, 42,* 189–198.

Blachman, B. A. (1984). Relationship of rapid naming ability and language analysis skills to kindergarten and first-grade reading achievement. *Journal of Educational Psychology, 76,* 610–622.

Blaiklock, K. E. (2004). The importance of letter knowledge in the relationship between phonological awareness and reading. *Journal of Research in Reading, 27,* 36–57.

Blonk, A. M. T., & Bosman, A. (2003, November). *Girls are better, even girls with dyslexia?! Gender differences in reading performance of teenagers with and without dyslexia.* Paper presented at the International Conference on Dyslexia, University of Amsterdam, the Netherlands.

Braams, T., & Bosman, A. M. T. (2000). Fonologische vaardigheden, geletterdheid en lees- en spellinginstructie. [Phonological skills, literacy, and reading and spelling instruction]. *Tijdschrift voor Orthopedagogiek, 39,* 199–211.

Carroll, J. M., & Snowling, M. J. (2004). Language and phonological skills in students at high risk of reading difficulties. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 45,* 631–640.

Catts, H. W., Fey, M. E., Zhang, X., & Tomblin, J. B. (2001). Estimating the risk of future reading difficulties in kindergarten students: A research-based model and its clinical implementation. *Language, Speech, and Hearing Services in Schools, 32,* 38–50.

Coleman, J. M., & Dover, G. M. (1993). The RISK Screening Test: Using kindergarten teachers' ratings to predict future placement in resource classrooms. *Exceptional Children, 59,* 468–477.

Cornwall, A. (1992). The relationship of phonological awareness, rapid naming, and verbal memory to severe reading and spelling disability. *Journal of Learning Disabilities, 25,* 532–538.

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

Flynn, J. M., & Rahbar, M. H. (1994). Prevalence of reading failure in boys compared with girls. *Psychology in the Schools, 31,* 66–71.

Flynn, J. M., & Rahbar, M. H. (1998). Improving teacher prediction of students at risk for reading failure. *Psychology in the Schools, 35,* 163–172.

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 and Allied Disciplines, 41,* 203–213.

Hammill, D. D. (2004). What we know about correlates of reading. *Exceptional Students, 70,* 453–469.

Hammill, D. D., Mather, N., Allen, E. A., & Roberts, R. (2002). Using semantics, grammar, phonology, and rapid naming tasks to predict word identification. *Journal of Learning Disabilities, 35,* 121–136.

Irausquin, R. (1999). *Nonword repetition test.* Unpublished test materials, University of Nijmegen, the Netherlands.

Johnston, R. S., Anderson, M., & Holligan, C. (1996). Knowledge of the alphabet and explicit awareness of phonemes in prereaders: The nature of the relationship. *Reading and Writing, 8,* 217–234.

Kenny, D. T., & Chekaluk, E. (1993). Early reading performance: A comparison of teacher-based and test-based assessments. *Journal of Learning Disabilities, 26,* 227–236.

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

Lewis, B. A., Freebairn, L. A., & Taylor, H. G. (2000). Academic outcomes in students with histories of speech sound disorders. *Journal of Communication Disorders,* 33, 11–30.

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

Mantzicopoulos, P. Y., & Morrison, D. (1994). Early prediction of reading achievement: Exploring the relationship of cognitive and noncognitive measures to inaccurate classifications of at-risk status. *Remedial and Special Education, 15,* 244–251.

Menyuk, P., Chesnick, M., Liebergott, J. W., Korngold, B., D'Agostino, R., & Belanger, A. (1991). Predicting reading problems in at-risk students. *Journal of Speech and Hearing Research, 34,* 893–903.

Meyer, M. S., Wood, F. B., Hart, L. A., & Felton, R. H. (1998). Selective predictive value of rapid automatized naming in poor readers. *Journal of Learning Disabilities, 31,* 106–117.

Mommers, M. J. C., Verhoeven, L., & van der Linden, S. (1979). *Veilig leren lezen* [Learning to read safely]. Tilburg, the Netherlands: Zwijsen.

Mommers, M. J. C., Verhoeven, L., & van der Linden, S. (1994). *Veilig leren lezen* [Learning to read safely] (2nd ed.). Tilburg, the Netherlands: Zwijsen.

Muter, V., & Snowling, M. (1998). Concurrent and longitudinal predictors of reading: The role of metalinguistic and short-term memory skills. *Reading Research Quarterly, 33,* 320–337.

O'Connor, R. E., & Jenkins, J. R. (1999). Prediction of reading disabilities in kindergarten and first grade. *Scientific Studies of Reading, 3,* 159–197.

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

Petersen, D. K. (2002). Prediction of poor and superior word reading. *Working Papers, 50,* 17–32.

Satz, P., & Fletcher, J. M. (1988). Early identification of learning disabled children: An old problem revisited. *Journal of Consulting and Clinical Psychology, 56,* 824–829.

Scarborough, H. S. (1989). Prediction of reading disability from familial and individual differences. *Journal of Educational Psychology, 81,* 101–108.

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

Scarborough, H. S. (1991). Antecedents to reading disability: Preschool language development and literacy experiences of children from dyslexic families. *Reading and Writing, 3,* 219–233.

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

Schneider, W., & Näslund, J. C. (1993). The impact of early metalinguistic competencies and memory capacity on reading and spelling in elementary school: Results of the Munich Longitudinal Study on the genesis of individual competencies. *European Journal of Psychology of Education, 8,* 273–287.

Snowling, M. J. (1981). Phonemic deficits in developmental dyslexia. *Psychological Research, 43,* 219–234.

Snowling, M., Goulandris, N., Bowlby, M., & Howell, P. (1986). Segmentation and speech perception in relation to reading skill: A developmental analysis. *Journal of Experimental Child Psychology, 41,* 489–507.

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.

Taylor, H. G., Anselmo, M., Foreman, A. L., Schatschneider, C., & Angelopoulos, J. (2000). Utility of kindergarten teacher judgments in identifying early learning problems. *Journal of Learning Disabilities, 33,* 200–210.

Teisl, J. T., Mazzocco, M. M., & Myers, G. F. (2001). The utility of kindergarten teacher ratings for predicting low academic achievement in first grade. *Journal of Learning Disabilities, 34,* 286–293.

van den Bos, K. P., Lutje Spelberg, H. C., & Eleveld, M. A. (2004). Voorspelling van leessnelheid van leerlingen in Groep 3 met tests afgenomen in Groep 2 en 3. [Prediction of decoding speed in Grade 1 with test assessment in kindergarten and Grade 1]. *Tijdschrift voor Orthopedagogiek, 43,* 312–324.

van den Bos, K. P. (2004). *Test serieel benoemen en woorden lezen* [Test of serial naming and word reading]. Unpublished test materials. Groningen: RUG.

van Doorn, F. (1996). *Werkhoudingslijst* [Task-assignment scale]. Amsterdam: Advies- en Begeleidingscentrum.

Verhoeven, L., & van Leeuwe, J. (2003). Ontwikkeling van decodeervaardigheid in het basisonderwijs [Development of decoding skills in primary school]. *Pedagogische Studiën, 80,* 257–271.

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

Walsh, D. J., Price, G. G., & Gillingham, M. G. (1988). The critical but transitory importance of letter naming. *Reading Research Quarterly, 23,* 108–122.

Weiner, P. (1985). The value of follow-up studies. *Topics in Language Disorders, 5,* 78–92.

Wesseling, R., & Reitsma, P. (2000). The transient role of explicit phonological recoding for reading acquisition. *Reading and Writing, 13,* 313–336.

Wentink, H., & Verhoeven, L. (2001). *Protocol voor leesproblemen en dyslexie* [Protocol for reading difficulties and dyslexia]. Nijmegen, the Netherlands: Expertisecentrum Nederlands.