

# Dutch home-based pre-reading intervention with children at familial risk of dyslexia

Sandra G. van Otterloo · Aryan van der Leij

Received: 19 June 2008 / Accepted: 15 September 2009 / Published online: 7 November 2009  
© The Author(s) 2009. This article is published with open access at Springerlink.com

**Abstract** Children (5 and 6 years old,  $n=30$ ) at familial risk of dyslexia received a home-based intervention that focused on phoneme awareness and letter knowledge in the year prior to formal reading instruction. The children were compared to a no-training at-risk control group ( $n=27$ ), which was selected a year earlier. After training, we found a small effect on a composite score of phoneme awareness ( $d=0.29$ ) and a large effect on receptive letter knowledge ( $d=0.88$ ). In first grade, however, this did not result in beneficial effects for the experimental group in word reading and spelling. Results are compared to three former intervention studies in The Netherlands and comparable studies from Denmark and Australia.

**Keywords** Early intervention · Familial risk · Home-based intervention · Phoneme awareness · Reading disabilities

## Introduction

Severe reading disability that cannot be explained by sensory deficits, attention deficits, general language problems, intelligence, poor home-literacy environment or poor education is called developmental dyslexia (Snowling, 2000). Because literacy is essential to function well in modern society, it has been suggested that dyslexia should be diagnosed and treated or even prevented early (Elbro & Scarborough, 2004b). All the more, as developmental dyslexia can be very persistent and can easily lead to other problems, such as fear of failure. Because dyslexia runs in families (Finucci, Gottfredson, & Childs, 1985; Gilger, Pennington, & DeFries, 1991; Scarborough, 1990), it is possible to locate and treat children at risk of reading failure before formal reading instruction starts. In the current study, the effects of early home-based intervention for children at higher familial risk are investigated. The intervention was based on phoneme awareness and letter knowledge and conducted in The Netherlands.

---

S. G. van Otterloo (✉) · A. van der Leij  
Department of Education, University of Amsterdam, P.O. Box 94208, 1090GE Amsterdam,  
The Netherlands  
e-mail: S.G.vanOtterloo@uva.nl

Several intervention studies have indicated that young children could benefit from phoneme awareness and letter knowledge training before the beginning of formal reading instruction (e.g. Ball & Blachman, 1991; Bradley & Bryant 1985; Schneider, Roth, & Ennemoser, 2000). Note that these studies are all school-based, in contrast to the current study that is home-based. Meta analyses (Bus & van IJzendoorn, 1999; Ehri, Nunes, Willows, Schuster, Yaghoub Zadeh, & Shanahan, 2001) showed large overall effect sizes on phoneme awareness measures ( $d=0.86$  and  $d=1.04$ , respectively). The overall effects on reading skill were also evident but moderate ( $d=0.44$  and  $d=0.53$ , respectively). At follow-ups, these effect sizes decline to some extent. Although most intervention studies focused on normal children, some selected their sample out of a special ‘at-risk’ population. Children at risk of developing reading difficulties were defined in several ways. Hatcher, Hulme, and Ellis (1994) selected 7-year old children experiencing difficulties in the early stages of learning to read. O’Connor, Jenkins, Leicester, and Slocum (1993) selected preschool children with developmental delays and Schneider and colleagues selected kindergartners with low phonological processing scores.

Six intervention studies selected kindergartners at familial risk of dyslexia (Elbro & Petersen, 2004; Eleveld, 2005; Fielding-Barnsley & Purdie, 2003; Hindson, Byrne, Fielding-Barnsley, Newman, Hine, & Shankweiler, 2005; Regtvoort & van der Leij, 2007; van Otterloo, van der Leij, & Henrichs, 2009). These studies will be described in more detail because they are directly related to the present study. Table 1 shows an overview of relevant aspects of these six studies. The interventions in all six studies took place in the school year before formal reading instruction started. The experimenters selected preschoolers (Fielding-Barnsley and Purdie; Hindson et al.) or kindergartners (Elbro & Petersen; Eleveld; Regtvoort & van der Leij; van Otterloo et al.) with at least one relative (Fielding-Barnsley and Purdie) or at least one parent (other five studies) with an established reading problem.

Elbro and Petersen (2004) investigated the effects of a phoneme awareness and letter knowledge programme that was conducted at the schools of the children. The kindergarten teacher presented the programme to the whole class for about 30 min every school day for 17 weeks. At post-test, the trained at-risk children outperformed the untrained at-risk children in phoneme awareness. Effect sizes were moderate (see Table 1). At follow-ups at the beginning of second and third grade, the trained at-risk children outperformed the untrained at-risk children on all reading measures.

Fielding-Barnsley and Purdie (2003) used an intervention of 8 weeks based on ‘dialogic reading’, a method for structured book reading (see Whitehurst, Arnold, Epstein, Angell, Smith, & Fischel, 1994) in the year prior to formal reading. The programme also included rhyme, vocabulary and concepts about print. During these weeks, parents read eight books about six to seven times with use of the dialogic reading method. Although the intervention is quite different compared to the intervention of the current study, we included the study in Table 1 as the sample was comparable and the intervention was home-based. The 26 experimental children were compared to 23 at-risk children selected at the same schools. At post-test at the beginning of first grade, the experimental children outperformed the control children on pre-reading measures and, at follow-up, a few months later on reading and spelling measures. The researchers concluded that the intervention is advantageous for children at risk of developing a reading disability. However, the comparison between groups did not include pre-test scores.

Hindson et al. (2005) made use of a training programme that focused primarily on phoneme awareness and letter knowledge and also included ‘dialogic reading’. The training sessions were held individually at preschool or at home once a week by a trained assistant

**Table 1** Aspects of intervention studies with children at familial risk of dyslexia

Aspects	Intervention studies					
	Elbro & Petersen 2004	Fielding-Barnsley & Purdie 2003 <sup>a</sup>	Hindson et al. 2005	Eleveld 2005 <sup>b</sup>	Regervoort and van der Leij 2007	van Otterloo et al. 2009
Language	Danish	Australian English	Australian English	Dutch	Dutch	Dutch
<i>n</i> at-risk exp group	35	26	69	E1:21 E2: 25 (Etot:46)	31	23
<i>n</i> at-risk control group	47	23	17	71	26	25
Content of intervention	PA, LK	Rhyme, vocabulary, CAP, Dialogic reading	PA, LK, Dialogic reading	E1: PA, LK, Serial naming speed	PA, LK	PA, LK
Trainer	Kindergarten teacher	Parent	Trained assistant	Trained assistant	Parent	Parent
Individually or whole class	Whole class	Individually	Individually	Individually	Individually; computerised	Individually
Number of sessions	Max. 85	About 40	11 to 17	About 30	Max. 70	Max. 50
Training time per session	About 30 min	Not given (dialogic reading of 1 or 2 books)	About 30 min	20–30 min	About 10 min	About 10 min
Total training time	42 h 30 min	Not given (reading 8 books 6,5 times during 8 weeks)	5 h 30 min–8 h 30 min	E1: 12 h 30 min E2: 25 h	11 h 40 min	8 h 20 min
<i>M</i> (SD) pre-test score on LK <sup>c</sup> at-risk children	5.8 (5.7)	No pre-test	10.0 (5.6)	13.0 (7.1)	12.8 (6.6)	12.6 (6.4)
Interaction time × condition on LK at post-test	<i>d</i> =0.67	Not possible to test interaction. E outperforms C	No effect	Etot-C: <i>d</i> =0.41	<i>d</i> =1.19	<i>d</i> =0.27 (ns)
Interaction time × condition on PA at post-test	<i>d</i> =0.47 to 0.54	Not possible to test interaction. E outperforms C	Significant effect for phoneme identity. Effect size not given	Etot-C: <i>d</i> =0.41	<i>d</i> =0.59	<i>d</i> =0.23 to 0.89
Transfer effects on reading and spelling a year later	<i>d</i> =0.40 to 0.69	Exp outperforms Controls on reading and spelling	Not given	No significant transfer effects on reading or spelling	No significant transfer effects on reading or spelling	<i>d</i> =0.31 to 0.63 (ns)

PA phoneme awareness, LK letter knowledge, RI reading instruction, CAP concepts about print

<sup>a</sup> No real comparison between groups is possible because of the absence of a pre-test

<sup>b</sup> Eleveld had two experimental groups: E1 is best comparable with the other studies, E2 received an extra training a year ahead. At post-test results of combined experimental group (Etot) are presented

<sup>c</sup> Elbro and Peterson measured letter naming while the other studies measured letter sounds

and took about 30 min. The number of sessions was partly governed by the results of criterion tests testing the target letter and ranged from 11 to 17. At post-test, the trained at-risk group outperformed an untrained at-risk group on several pre-reading measures: phoneme identity (a phoneme awareness measure), rhyme, and print familiarity. At follow-up, in the last 2 months of kindergarten (in the year they receive formal reading instruction), trained not-at-risk children outperformed trained at-risk children. Despite these findings, the results of an earlier study within the same research group and using the same instruments (Byrne & Fielding-Barnsley, 1993) gave the opportunity to compare the scores of the trained at-risk children to the scores of untrained unselected control children. The trained at-risk children from the Hindson study scored similar in spelling and word identification to untrained unselected control children from the study of Byrne and Fielding-Barnsley. Although a straight comparison is lacking, this indicates that the intervention might have improved the performance of at-risk children to grade average.

In the Dutch study of Eleveld (2005), a programme of similar intensity as the current study was used. Difference was the focus on naming speed, in addition to practice of phoneme awareness and letter knowledge. The training sessions were held individually at school by a trained assistant. Eleveld compared the results of 21 trained at-risk children (E1) with 25 trained at-risk children (E2) who also received a second training based on phonological skills and serial naming speed a year earlier. Both training groups were compared to a no-training control group of 71 at-risk children. In Table 1, we present the results of the combined intervention group ( $E_{\text{total}}$ ) only because the experimental groups did not differ in their test results. At post-test, interaction of time  $\times$  condition was found for both phoneme awareness and letter knowledge (small to moderate effect sizes). There were no transfer effects on reading or spelling.

The intervention in the Dutch study of Regtvoort and van der Leij (2007) combined a phoneme awareness and letter knowledge programme with reading instruction in accordance with the minimal pairing procedure of Beck (1989). In exercises using this procedure, one-syllable words (both VC words and CVC words) slightly change with one letter at a time. The programme was computerised and home-based. The authors compared a group of trained at-risk children to a group of untrained at-risk children and a group of untrained not-at-risk children. At post-test, the trained at-risk children gained more on a combined measure of phoneme awareness than the untrained at-risk children and to the same extent as not-at-risk children. Effect sizes were moderate. On a measure of letter knowledge, the trained at-risk children gained more between pre-test and post-test in comparison to both the untrained at-risk children and the not-at-risk children. Effect sizes were large. However, these effects did not transfer to significant differences between trained and untrained at-risk children in reading and spelling skill in first or second grade. In most cases, the not-at-risk children outperformed the at-risk children.

In a Dutch pilot study of the present study (van Otterloo et al., 2009), a group of kindergartners that received phoneme awareness and letter knowledge training (the experimental group) was compared to a group that received a training focusing on morphology, syntax, and vocabulary (the control group). Both programmes were designed to take about 10 min a day, 5 days a week for 10 weeks. The parents performed the intervention at home. At post-test, the experimental group made more progress than the control group on the phoneme awareness measures but not significantly on letter knowledge. The overall effect on phoneme awareness was large ( $d=1.15$ ). On the follow-up tests at the end of first grade, the experimental children outperformed the control children on reading and spelling skill, but the differences between the groups were not statistically significant (small to moderate effects).

In general, there was a comparable result at post-test across the studies in Table 1, independent of country, language, or kind of tutor. It can be concluded that it is possible to train phoneme awareness to children at familial risk of reading disability in several countries. The effect on letter knowledge, however, was significant in only part of the studies, in the Danish study and in the Dutch studies of Eleveld (2005) and Regtvoort and van der Leij (2007). The relatively good result at letter knowledge in Denmark might be related to the score on letter knowledge at pre-test. Table 1 shows that, in the three Dutch studies, the at-risk children knew more letters at pre-test than in the Danish study (Elbro & Petersen, 2004). Therefore, it seems possible that the children in Denmark had more to gain, which made an intervention there more necessary and more effective. However, it should be noted that Elbro and Petersen measured letter naming instead of letter sounds (receptive letter knowledge), a somewhat different and more difficult skill. Another explanation for the large effect in the Danish study is the much lengthier intervention time (42.5 h versus 5 to 12.5 h in the other studies). In addition, in contrast to the Dutch and Australian studies, the Danish study did not involve individual training but whole class training. The ultimate aim of the early interventions is transfer to reading and spelling as indicated by follow-up measurements. Table 1 shows that the intervention of Elbro and Petersen (2004) had moderate transfer effects on reading and spelling. The study of Fielding-Barnsley and Purdie gives the suggestion of transfer to reading and spelling, but effects cannot be determined in the absence of a pre-test. Hindson et al. (2005) did actually not investigate transfer. The transfer effects in the Dutch intervention study of van Otterloo et al. (2009) were of comparable size as in the Danish study, but not statistically significant, possibly due to smaller sample size. The Dutch studies of Eleveld and Regtvoort and van der Leij did not result in significant transfer effects on reading and spelling.

In the present study, the programme of van Otterloo et al. (2009) that induced moderate transfer was extended and involved longer training time and teaching more letters including more low-frequent letters. It is important to note that the programme was based on the programme that was successfully used in the Danish study of Elbro and Petersen (2004). Because early letter knowledge is an important predictor of reading skill in The Netherlands (Braams & Bosman, 2000; de Jong & van der Leij, 1999), a larger effect on letter knowledge at post-test might also enlarge transfer to reading and spelling in first grade. To increase the power to detect training effects, the sample size in the current study was larger: 57 instead of 42 participants at the end of first grade.

There are two additional points that should be mentioned about the current intervention study. In the present study, the phoneme awareness and letter knowledge programme was home-based with parents functioning as tutors, similar to the studies of Fielding-Barnsley and Purdie (2003), Regtvoort and van der Leij (2007) and van Otterloo et al. (2009) but different to the studies of Eleveld (2005), Elbro and Petersen (2004) and Hindson et al. (2005) who used trained assistants or kindergarten teachers. There is evidence to support our choice of trainers. Several researchers showed that parents can have substantial effect on the pre-reading development of their children by directly teaching letters at home (e.g., Sénéchal & LeFevre, 2002). Torppa and colleagues even concluded that parental teaching of letter names to children at familial risk of dyslexia may have a compensating effect on letter knowledge of at-risk children (Torppa, Poikkeus, Laakso, Eklund, & Lyytinen, 2006). Van Otterloo et al. (2009) concluded that most parents in their sample were able to work with the phoneme awareness and letter knowledge programme properly, resulting in moderate intervention effects.

A second point that should be mentioned concerns treatment integrity. Because the control on treatment integrity in the Dutch intervention study (van Otterloo et al., 2009) was

only based on self-report, monitoring and checking the quality and quantity of administration more precisely might stimulate better treatment integrity, which might lead to larger training effects. When treatment integrity is insufficiently checked and positive training effects are found, the outcome can be caused by something else than the training. In addition, when the training did not have positive results, this could be due to an ineffective treatment as well as an inadequately implemented treatment (Troia, 1999). In the present study, treatment integrity was promoted, checked, and assessed more carefully. The results of the assessments of treatment integrity have been reported in a separate study (van Otterloo, van der Leij, & Veldkamp, 2006), with the same participants as in the experimental group of the current study. It was concluded that overall treatment integrity was sufficient (see ‘Method’ for more details). Therefore, it may be predicted that the outcome of the current study can be ascribed to the intervention.

### Research question, summary of design and expectations

The current study was designed to replicate our earlier intervention study (van Otterloo et al., 2009). However, there were three important differences: (1) The present study used a longer and adjusted version of the intervention programme with more letters and was comparable in training time to the studies of Eleveld (2005) and Regtvoort and van der Leij (2007); (2) there was more attention given to treatment integrity; and (3) the study included a no-training control group instead of a trained control group. The research question was whether this specific training focusing on phoneme awareness and letter knowledge, and administered by the parents, would lead to immediate effects on the trained skills and transfer effects on reading and spelling in first grade for Dutch kindergarten children at familial risk.

Two groups were involved, both at familial risk: an experimental group that received the intervention and a control group that did not receive the intervention. For practical reasons the present study did not include a control group that received a non-specific training. In our earlier intervention study (van Otterloo et al., 2009), non-specific effects of training were controlled for, and the authors were able to conclude that the training resulted in specific effects. The intervention took place during the second half of the final kindergarten year. The selected children did not get formal reading instruction yet. In The Netherlands, reading instruction starts in first grade.

In addition to the comparison between groups, we also looked at individual differences within the experimental group by comparing weaker and less weak participants. We did so to investigate whether weaker children profited equally from the intervention in comparison to less weak children.

Analogous to the results of similar intervention studies (see Table 1), we expected that the experimental group gained more in phoneme awareness. In contrast to the results of our earlier study (van Otterloo et al., 2009), but in line with the studies of Elbro and Petersen (2004), Regtvoort and van der Leij (2007) and Eleveld (2005), we expected the experimental group to gain more in letter knowledge than the control group because more letters were trained. In addition, because an intervention with similar content (Elbro and Peterson) produced moderate transfer to reading and spelling, we expected the experimental group to outperform the control group on reading and spelling measures in first grade. The expectation was also motivated by the finding that letter knowledge is a strong predictor of reading skill (Elbro & Scarborough, 2004a; Vellutino, Fletcher, Snowling, & Scanlon, 2004; Braams & Bosman, 2000). Although naming speed has also been mentioned as a strong predictor of (later) reading skill (Felton & Wood, 1989; van den Bos, 1998; Wolf,

1991), we did not expect the experimental group to make more progress on this skill because it was not trained.

## Method

### Recruitment of participating families

For practical reasons, participants of the experimental group were recruited together with participants of a parallel study. In the greater area of Amsterdam, 130 elementary schools were summoned with the request to spread folders for parents in kindergarten. The folders contained concise information about the study and the request to participate when at least one of the parents had severe reading and spelling problems. Ninety-six schools volunteered to spread in total 4,500 folders. One hundred and nine families with at least one parent that gave a self-report of dyslexia were interested in participation.

This self-report was confirmed by a screening test: a fluency of word reading test, a test on fluency of pseudoword reading and a verbal IQ measure. These measures are described in more detail in ‘Measures’. The parents were tested individually at their homes by the first author or a trained assistant. The parent was selected for participation when his or her scores on both reading tests were equal to or below the 20th percentile of the total population, or the score on one of the reading tests was equal to or below the tenth percentile. In addition, we selected the parent if the percentile score on the verbal IQ measure was extremely higher than the percentile score on one of the reading tests, with a distance of at least 60 (e.g. the score on the verbal competence was above the 90th percentile and the score on one of the reading tests was beneath the 30th percentile). The rationale behind this so-called disparity score was that verbally competent people might show reading scores just above the criterion, probably due to extended experience with reading in higher education and vocation. When their reading scores are much lower than can be expected from their high verbal competency, this means a serious reading handicap for these people.

Of the screened families, 73% was eligible for inclusion in the experimental studies (80 families in total). About half of them, 37 in total, were selected for participation in the experimental group of the current study. One girl was excluded because she already could read at pre-test. Four families withdrew at an early stage because the parents turned out to be too occupied to participate (two), the intervention was too difficult for the parents to administrate (one) or the parents did not agree with the experiment (one). One boy was excluded from the experimental group because he finished less than 3 weeks of training. One girl was excluded from the analyses because she went to special education halfway first grade.

The at-risk control children were recruited a year ahead together with a group of children not at risk for a related study. In the greater area of Amsterdam, 46 elementary schools were summoned with the request to spread folders for parents in kindergarten. Twenty-seven schools volunteered to spread the folders. In sum, 1,440 folders were spread. Because insufficient families were interested in participation, we placed advertisements and articles in regional newspapers. Thirty-four families were screened because they were interested in participation and at least one parent gave a self-report of dyslexia. Of these families, almost 85% were eligible for inclusion in the current study (29 families in total). Two children from this group (one girl or one boy) withdrew because the parents or the child did not want to continue.

According to the selection criteria, 17 participating parents in experimental or control group (30%) scored below or equal to the tenth percentile on both reading measures and 19

parents (33%) scored below or equal to the 20th percentile on both measures and/or below or equal to the tenth percentile on one of these measures. Twenty-one parents (37%) were included because of the ‘disparity-score’. The experimental and control groups did not differ in mean scores on any of these selection measures.

## Participants

Thirty experimental children (19 boys and 11 girls) and 27 control children (14 boys and 13 girls) were available for data analyses. At pre-test, all participating children were in their second kindergarten year (the year before they receive formal reading instruction). A chi-square test indicated that the groups did not differ significantly in the gender distribution. The children were derived from 43 schools, with four children from one school at most.

Table 2 gives a description of the participants in age and control measures. The groups did not differ in age or in performance on the receptive vocabulary measure as well as in performance on the non-verbal IQ measure. In comparison with chronological age-related norms (non-verbal IQ) or didactical age-related norms (receptive vocabulary), the group means was close to the 75th percentile, so substantially above mean and suggesting a middle class population.

Table 3 shows background information related to home literacy environment. The information was collected with the use of the parent questionnaire of Eleveld (2005). It is important to note that home literacy environment and educational level of the parents was relatively high. On average, the social economic status of the families was in the range of middle class as is confirmed by years of schooling. In the Dutch school system, a number of 10 years of schooling from first grade onwards indicates that the individual has had 6 years of schooling at the primary level, followed by 4 years at the secondary level. A number of 11 years upwards indicate extended schooling at the secondary level and/or schooling at the third level (professional college or university). Most parents had at least 10 years of schooling. Chi-square tests indicated that the groups did not differ significantly on these measures. A substantial proportion of the parents reported reading problems of close family members, and many parents reported early language problems of the child. This confirms that the selected participants came from families with reading disabilities.

## Intervention

The intervention, Sounding Sounds and Jolly Letters (Klinkende Klanken en Lollige Letters), was a home-based pre-reading programme with parents as tutors. It was a translation and adaptation of a Danish school-based kindergarten phoneme awareness and

**Table 2** Means and standard deviations on age and control measures

Measures	Experimental group, <i>n</i> =30		Control group, <i>n</i> =27	
	<i>M</i>	SD	<i>M</i>	SD
Age <sup>a</sup> (in month)	69.87	3.22	71.13	4.42
Receptive vocabulary <sup>a</sup>	72.23	9.78	70.74	9.87
Non-verbal IQ <sup>b</sup>	25.57	4.68	26.93	4.58

<sup>a</sup> Measured at pre-test

<sup>b</sup> Measured halfway first grade



**Table 3** Background variables related to home literacy environment

Variables	Experimental group, <i>n</i> =30 (%)	Control group, <i>n</i> =27 (%)
Number of years of schooling <sup>a</sup> mother		
≥15	53	44
10–14	43	44
≤9	3	11
Number of years of schooling <sup>a</sup> father		
≥15	43	44
10–14	37	26
≤9	20	30
Only father dyslexic	77	67
Only mother dyslexic	20	33
Both parents dyslexic	3	0
Reading problems direct family	47	63
Early problems with language (child)	37	37
Number of books in the home		
<100	33	41
100–200	40	37
>200	27	22
Library membership child	73	67
Shared reading		
<2 times a week	3	7
2–4 times a week	30	37
5 times a week or more	67	56

<sup>a</sup>From first grade onwards

letter programme, ‘Towards initial reading: Phonological awareness’ (Borstrøm & Petersen, 1996). The Danish programme was described and applied in a longitudinal study (Borstrøm & Elbro, 1997; Elbro & Petersen, 2004; Petersen, 2000). An addition we made to the training is the use of a mirror along with the articulation exercises and, more importantly, writing exercises of the newly learned letters on a slate. After two earlier studies with the programme, we extended the programme and made adaptations as a consequence of suggestions and remarks we received from the participating parents.

The Dutch programme was designed to take about 10 min a day, 5 days a week for 14 weeks (totally approximately 12 h). The programme focused on the letter–sound correspondences and gave special attention to the articulation of the speech sounds. It progressed slowly with approximately two speech sounds a week. The following 20 letters and speech sound (12 consonants and eight vowels) were used in the programme in the presented order: aa, oo, ee, uu, m, s, p, l, a, o, e, i, t, r, v, b, n, k, d and z, (/a/,/o/,/e/,/y/,/m/,/s/,/p/,/l/,/a/,/ɔ/,/ɛ/,/ɪ/,/t/,/r/,/v/,/b/,/n/,/k/,/d/,/z/). We used both long vowels (in Dutch represented by a homogeneous digraph) and short vowels. Every speech sound and the corresponding letter were introduced in several ways. First, a rhyme or song was read to the child with focus on the speech sound. Then, the child was shown a picture of the letter while the parent produced the corresponding speech sound. If possible, the child also received a semantic cue (/m/is the ‘taste-good’ sound). Afterwards, the parent and child wrote down the letter several times on the slate. Together they thought of people they knew

whose name started with the sound, followed by an articulation exercise with attention to rounding (vowels), place of articulation and manner of articulation. The newly learned speech sounds were repeated in nursery rhymes and language games. Phoneme blending and sound identity of both initial and final sound were trained by several language games. These games included the use of pictures, card games and a hand doll. The other training materials consisted of two programme books, instruction sheets, a slate with chalk and a progress table with funny stickers for the child. The programme book contained detailed descriptions of the daily lessons, including the nursery rhyme of the day and references to which materials to use. The instruction sheets explained the aims of the programme and gave general instructions. The parents were, among others, explicitly instructed to use speech sounds instead of letter names and were explained why.

### Procedure

The first author of this paper organised an instruction meeting for all parents before the programme started. During this meeting, the material was presented, the aims of the programmes were explained and exercises were demonstrated and discussed. We motivated the participants to read the instructions carefully before starting the lesson. The parents also had the opportunity to ask questions. Participants who were not present at the meeting were contacted individually.

After 1 week of training, the parents were phone-called to give them the opportunity to ask questions and to find out if there was any lack of clarity. After about 6 weeks of training, the first author organised a second meeting. During this meeting, the second part of the programme was presented, instructions were repeated, and questions were answered.

The children were tested individually at their schools in a separate room before and after the training, in January and June 2003. There were follow-ups conducted in first grade, in January and June 2004. The test sessions lasted about 45 min each. Several specifically trained graduate students were involved in the testing of the children.

### Treatment integrity

We measured treatment integrity to stimulate integrity and to control if integrity was sufficient. Two aspects of treatment integrity were assessed: the quality and the quantity of the administration of the intervention programme. For the rationale behind these instruments, we refer to the study of van Otterloo et al. (2006). To measure the quality of administration, one lesson was videotaped and judged with an instrument focusing on the interaction between tutor and tutee. The instrument was a sum of five items using 5-point Likert scales (Cronbach's alpha was 0.92). A large number of families (79%) received a high mean observation score, a score between 4 and 5. The mean was 4.2 (SD=0.80). Information about the quantity of administration, the number of weeks the subjects worked on the programme, was gathered with the use of log forms for the parents to be filled in every day. Most families (67%) completed the whole training programme (14 weeks). A few families stopped early in the programme and some later on. One participant completed less than 15 lessons and was therefore excluded from the experiment. The mean number of weeks completed was 12.13 (SD=3.24). Based on these results, van Otterloo et al. concluded that the treatment integrity was sufficient.

One might question whether a parent with dyslexia was sufficiently able to carry out the intervention. We had information about which parent was dyslexic (from the screening) and which parent was as tutor working with the intervention (information gathered with the

log forms). Table 3 shows that, in the experimental group, 22 children (77%) had a dyslectic father, seven children had a mother with dyslexia (20%) and of one child both parents were dyslectic (3%). In nineteen cases, the mother was the tutor (63%); in three cases, the father was the tutor (10%); and in eight cases, the parents varied in their role as tutor (27%). Combining these data, it shows that, in 15 cases, the non-dyslectic parent was the tutor (50%); in seven cases, the dyslectic parent was the tutor (23%); and in eight cases, the dyslectic parent gave part of the intervention (27%). None of the dyslectic tutors that were included in the experiment reported problems with their role of tutor. The quality of administration of the 15 non-dyslectic tutors was somewhat, but not significantly, higher in comparison to the quality of administration of the seven dyslectic tutors,  $t(20)=1.52$ , ns,  $d=0.67$  (mean observation score of 4.2 versus 3.7).

## Measures

### *Screening*

To investigate possible dyslexia of the parents both word recognition and phonological recoding were measured. Because of the important role of reading speed in The Netherlands in comparison to reading accuracy, we used fluency measures. To investigate a possible disparity between reading level and verbal language level, a test of verbal competence was added as a control measure. This measure was also included to exclude persons with very low verbal competence. The same measures were used in several other studies selecting children at familial risk in The Netherlands (e.g. Eleveld, 2005; Koster et al., 2005; Regtvoort & van der Leij, 2007). The measures used for screening are described below.

*Fluency of word reading [EMT (1-min test), Brus & Voeten, 1980]* Word recognition was measured by a card with four columns of words (total, 116) with increasing difficulty. This test measured the number of words the subject could read correctly in 1 min. Mean parallel test reliability (between forms A and B) was  $r_{tt}=0.90$  (range, 0.76–0.96) (van den Bos, Iutje Spelberg, Scheepstra, & de Vries, 1994).

*Fluency of pseudoword reading [de Klepel (the Clapper); van den Bos et al., 1994]* Phonological recoding was measured by a card with four columns of pseudowords with increasing difficulty, analogous to the EMT. The experimenter measured the number of pseudowords that the subject could read and pronounce correctly in 2 min. Reported mean parallel test reliability (between forms A and B) was  $r_{tt}=0.92$  (range, 0.89–0.95).

*Verbal competence* As a control measure, we used the subtest ‘Similarities’ from the Dutch translation of the Wechsler Adult Intelligence Scale—Revised (Wechsler, 1981). In this test, the experimenter asked the subject to give the similarity between two concepts. Every answer gave 0, 1 or 2 points with a maximum score of 26.

### *Control measures children*

Both verbal IQ (receptive vocabulary) and non-verbal IQ were measured to describe the sample and exclude children with very low verbal competence or cognitive ability.

*Receptive vocabulary* This test is part of the Taaltoets Allochtone Kinderen (language test foreign children; Verhoeven & Vermeer, 1996). Each item consists of four pictures and a

spoken word. The child had to choose the alternative that best matches the given word. The test had 98 items that increases in difficulty. The administration of the test was stopped when the child failed six or more of the last eight items. For this test, the Cronbach's alpha was 0.89. The test was administered during pre-test (time 1).

*Non-verbal IQ* The Coloured Progressive Matrices (Raven, Court, & Raven, 1984) has three subtests with 12 test items each. Every item consists of a rectangular pattern in which a part is missing. The child has to look for the missing part and must choose between six alternatives. The maximum score is 36. Split half reliability (corrected for test-reduction) was 0.68, for 6-year-olds 0.82 and for 7-year-olds 0.84 (van Bon, 1986). The test was administered during first follow-up (time 3).

### *Effect measurement*

*Design* The intervention took place in the second half of the final kindergarten year, in the months of February to May 2003. All children were pre-tested in January or at the beginning of February (time 1) and post-tested in June or at the beginning of July (time 2). Table 4 presents an overview of constructs, measures used for effect measurement and their occasion of administration.

At pre- and post-test (time 1 and 2), we measured several skills that were correlates of reading or spelling. We selected phoneme awareness measures and a letter knowledge-measure to investigate whether the experimental programme gave immediate results. Rapid naming measures were included to investigate whether training effects were specific. Because the children did not receive formal reading instruction yet, we did not include reading or spelling measures.

In the first grade, after several months of reading instruction, we measured reading and spelling to investigate whether specific training effects lead to transfer on these skills. These follow-ups took place in January (time 3) and June (time 4). Measures of phoneme awareness and letter knowledge were included, but several measures could not be used as extreme ceiling effects were to be expected. Measures used for effect measurement are described below.

**Table 4** Overview of constructs, measures and occasion of administration

Construct	Measures	Kindergarten		First grade	
		Time 1	Time 2	Time 3	Time 4
		Intervention ↓			
Phoneme awareness	Alliteration	X	X		
	Phoneme blending	X	X	X	
	Phoneme segmentation	X	X	X	
Letter knowledge	Receptive letter knowledge	X	X	X	
	Productive letter knowledge			X	X
Rapid naming speed	Rapid naming-pictures	X	X		
	Rapid naming-colours	X	X		
Word recognition	Fluency of word reading I			X	
	Fluency of word reading II				X
Spelling	Spelling				X

*Alliteration (Irausquin, unpublished manuscript)* The experimenter sounded out three words. Two of these words started with the same speech sound. One word started with a different speech sound. The child had to name the word that started with the different sound ('odd one out'). Examples are kat–kip–mes (cat–chicken–knife). The test consisted of three practice items and ten test items and had a maximum score of 10. Eleveld (2005) reported Cronbach's alphas of 0.74–0.83.

*Phoneme blending [Toets voor Auditieve Synthese (Test for Phoneme Blending; Verhoeven, 1993a)]* The experimenter gave the child a word in separate phonemes and asked the child to blend the phonemes and pronounce the correct word (e.g., /p/-/l/-/a/-/n/-/t/ makes *plant*). The maximum score was 20. Cronbach's alpha was above 0.85 (Verhoeven, 2000).

*Phoneme segmentation [Toets voor Auditieve Analyse (Test for Phoneme segmentation; Verhoeven, 1993b)]* The experimenter gave a whole word and the child was required to segment the word in separate phonemes [e.g. *raam* (window) makes /r/-/a/-/m/]. The maximum score was 20. Cronbach's alpha was above 0.85 (Verhoeven, 2000).

*Receptive letter knowledge* Receptive knowledge of 32 letters or digraphs was tested. Letters with a very low frequency in Dutch (e.g. c, q, x, y) were not presented. The experimenter showed the child a sheet with rows with six printed lowercase letters each. Then, she gave a speech sound and asked the child to indicate the printed letter in the row that matched the sound. Every correct answer gave one point (Verhoeven, 2002). The reliability (Cronbach's alpha) in second kindergarten year was 0.87–0.89 (Eleveld, 2005).

*Productive letter knowledge* In the 'Grafemetoets' (Grapheme test) (Verhoeven, 1993a), the child has to read out loud 34 separate graphemes. We measured both accuracy and speed. The maximum accuracy score was 34. Cronbach's alpha was above 0.85 (Verhoeven, 2000).

*Rapid naming pictures* The experimenter showed the child a sheet with five rows of ten pictures. There were five different pictures (a bike, a tree, a fish, a bed, and a chair) represented in varying order. The experimenter required the child to name the pictures as quickly as possible and measured the time. The time per item (in seconds) was calculated. Split-half reliability for kindergarten was 0.73. For first grade the mean of split-half and test–retest reliability was 0.81 (van den Bos, 2003).

*Rapid naming colours* This measure was similar to the rapid naming pictures measure but with coloured blocks instead of pictures. There were five different colours used: black, yellow, red, green and blue. Split-half reliability for kindergarten was 0.80. For first grade, the mean of split-half and test–retest reliability was 0.88 (van den Bos, 2003).

*Fluency of word reading I* Word recognition at time 3 was measured by the first test card (1C) of the Drie-Minuten-Toets (3-min test; Verhoeven, 1995), which contains five columns of CV (consonant–vowel), VC, and CVC words, increasing in difficulty. The children were instructed to read the words correctly and as quick as possible. Cronbach's alpha was 0.88 (Verhoeven & van Leeuwe, 2003).

*Fluency of word reading II* Word recognition at time 4 was measured with the use of cards 1A, 2A and 3A of the Drie-Minuten-Toets (3-min test; Verhoeven, 1995). Card 1A

contained five columns of CV (consonant–vowel), VC and CVC words, increasing in difficulty. Card 2A also contained five columns with one-syllable words but, this time, with consonant clusters (CCVC, CVCC, CCVCC and CCCVC words). Card 3A contained four columns with words with more than one syllable. For every card, the instruction was the same. The children were asked to read the words correctly and as quick as possible. The score per card is the number of words read correctly in 1 min. The composite measure is the mean of the raw scores on the separate cards. Cronbach's alpha of the first, second and third card was, respectively, 0.88, 0.94 and 0.92 (Verhoeven & van Leeuwe, 2003).

*Spelling* We used the dictation E3A from the 'Schaal Vorderingen in Spellingvaardigheid 1' (Progress in Spelling Skill Scale; van den Bosch, Gillijns, Krom, & Moelands 1993). The experimenter dictated a target word, which the child had to write down. The dictation consisted of two parts. The first part contained 20 one-syllable words with consonant clusters (CVCC, CCVCC, VCCC, CVCCC and CCCVC words). The second part was a bit easier and contained 17 one-syllable words (CVC, CVCC and CCVC words). Every correctly spelled word gave one point, with a maximum of 37. Reliability of the test was 0.87 (Moelands & Kamphuis, 2001).

## Results

### Missing data

In addition to dropouts mentioned in 'Method', overall, less than 1% of data were missing. There were no variables with more than 5% (three subjects) of data missing. Although missing data were not missing (completely) at random because one subject from the control group was not present at pre-test, they were estimated and imputed by a missing value analysis using the regression method with addition of random residues (SPSS 10.0 Syntax Reference Guide, 1999).

### Effect measurement

#### *Pre-test (time 1)*

Table 5 shows the means and standard deviations on all separate pre- and post-test measures. It also shows the results on composite measures of phoneme awareness and rapid naming. The separate phoneme awareness measures were transformed into the proportion of correct answers, so we could calculate the mean of the three variables and make a composite score. We choose to use this transformation instead of to standardise the scores because means and standard deviations are easier to interpret this way. A composite score based on standardise scores does not give other results. Composite measures were calculated for the two separate rapid naming measures by calculating the mean of the separate tests. Independent *t* tests indicated that the groups did not differ significantly on any pre-test measure.

#### *Post-test (time 2)*

After post-test, we found an interaction of time  $\times$  condition (two repeated measurements) on receptive letter knowledge. The effect is considerable. See Table 5 for *F* and *d* values. As expected, the experimental group makes more progress on letter knowledge.

**Table 5** Means and standard deviations on pre- and post-test measures, MANOVA results and effect sizes

Measures	Experimental group, <i>n</i> =30				Control group, <i>n</i> =27				Interaction Time × condition	Effect size <i>d</i> <sup>a</sup> Interaction Time × condition
	Pre-test (time 1)		Post-test (time 2)		Pre-test (time 1)		Post-test (time 2)			
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>F</i> (1, 55)	
Receptive letter knowledge	12.87	6.08	20.47	5.34	12.44	6.91	16.71	7.79	10.53*** <sup>b</sup>	0.88
Phoneme awareness	33.94	25.97	50.72	24.40	31.54	23.51	43.61	28.85	1.19 <sup>b</sup>	0.29
Phoneme blending	7.00	6.61	10.50	5.80	6.30	5.98	9.34	7.02	0.28 <sup>b</sup>	0.14
Phoneme segmentation	4.50	5.07	8.93	6.17	4.41	5.29	7.24	6.71	2.59* <sup>b</sup>	0.43
Alliteration	4.43	2.81	5.50	3.24	4.11	2.90	4.79	3.34	0.16 <sup>b</sup>	0.11
Rapid naming	1.82	0.62	1.73	0.77	1.87	0.55	1.62	0.48	2.38	-0.41
Rapid naming-pictures	1.73	0.47	1.59	0.57	1.87	0.65	1.57	0.49	1.71	-0.35
Rapid naming-colours	1.91	0.84	1.88	0.99	1.88	0.56	1.66	0.58	1.69	-0.35

<sup>a</sup> Transformation of  $\eta^2$  into Cohen's *d* using the formula:  $d = [(4\eta^2)/(1-\eta^2)]^{1/2}$  (see Rosenthal, Rosnow, & Rubin, 2000, p.15)

<sup>b</sup> The original *p* value was divided in two. This is justified as the interaction of time × condition with two times of measurement gives exactly the same *p* value as an independent *t* test on the difference of the scores at time 2 and 1. Dividing the *p* value is common when testing one-tailed with a *t* test

\**p*<0.10; \*\**p*<0.05; \*\*\**p*<0.001

On the composite measure of phoneme awareness the interaction of time × condition is not significant, the effect size is small ( $d=0.29$ ) but potentially meaningful given the relatively small sample size. This effect is mainly caused by a tendency to a moderate effect on phoneme segmentation ( $d=0.43$ ). On the other two measures, phoneme blending and alliteration, the experimental group did not significantly make more progress.

As expected, the experimental group did not make more progress in rapid naming. Unexpectedly, there was a tendency towards more progress in the control group on the rapid-naming composite. The effect size was small. The effects on the separate rapid naming measures were a bit smaller.

Independent *t* tests showed that these differences in progress between groups resulted in the experimental group scoring better at post-test on receptive letter knowledge,  $t(45.38)=2.10$ ,  $p<0.05$  (one-tailed),  $d=0.48$ . The control group had a larger standard deviation on receptive letter knowledge,  $F(1, 55)=6.16$ ,  $p<0.05$  (Levene test), but outliers did not cause this difference. Group differences on phoneme awareness were smaller than expected, and not statistically significant, but in the expected direction: composite measure phoneme awareness,  $t(55)=1.01$ , ns,  $d=0.25$ ; phoneme blending,  $t(55)=0.68$ , ns,  $d=0.17$ ; phoneme segmentation,  $t(55)=0.99$ , ns,  $d=0.25$ ; and alliteration,  $t(55)=0.81$ , ns,  $d=0.21$ . The groups did not significantly differ on rapid naming measures, which conformed to the expectation.

### First follow-up in first grade (time 3)

The first part of Table 6 indicates the means and standard deviations on all January measures of first grade (time 3) and effect sizes of the group differences. It is clear that the experimental group did not outperform the control group, as the means do not differ much between the groups. The largest difference in favour of the experimental group is on phoneme segmentation. However, an independent *t* test showed that the experimental group also did not do significantly better on phoneme segmentation,  $t(55)=0.71$ , ns,  $d=0.18$ . It should be noted that all accuracy measures of phoneme awareness and letter knowledge were skewed to the left and several measures approached ceiling. The skewness (SE) of receptive letter knowledge was  $-0.84$  (0.32), the skewness of productive letter knowledge–accuracy was  $-1.19$  (0.32), of phoneme blending  $-2.09$  (0.32) and of phoneme segmentation  $-1.36$  (0.32).

On the fluency of word reading test, the standard deviation in the control group seemed larger than in the experimental group. A Levene test confirmed a tendency,  $F(1, 55)=3.92$ ,  $p=0.053$ . When we look closer at the data, the control group had two participants with relatively high scores, 55 and 48, which were possible outliers. Removing these participants from the analysis reduced the mean and standard deviation,  $M$  (SD)=19.04 (9.55), but still, the experimental group did not do better.

The repeated measures analysis did not show an interaction of time  $\times$  condition on phoneme segmentation anymore, which conforms to the skewness of the variable. Receptive letter knowledge still resulted in a significant interaction of time  $\times$  condition,  $F(2, 54)=6.70$ ,  $p<0.01$ ,  $d=1.00$ . Figure 1 shows that the experimental group made more progress in letter knowledge between pre- and post-test but that the control group made more progress between post-test and follow-up, resulting in almost the same progress in total. A contrast analysis (simple contrast) indicated that both interactions of time  $\times$

**Table 6** Means and standard deviations on first-grade measures and effect sizes of group differences

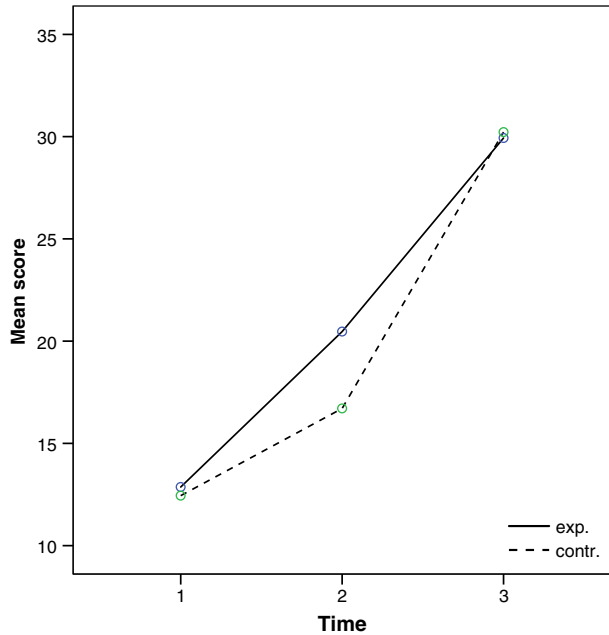
Measures	Experimental group, $n=30$		Control group, $n=27$		Effect size of group difference $d^b$
	$M$	SD	$M$	SD	
Follow-up I (time 3)					
Fluency of word reading I	17.80	8.55	21.44	12.65	-0.29
Receptive letter knowledge	29.93	1.82	30.22	1.72	-0.17
Productive letter knowledge—accuracy	31.63	2.28	32.37	2.45	-0.30
Productive letter knowledge—speed <sup>a</sup>	41.71	14.61	40.80	17.14	0.05
Phoneme blending	18.00	2.56	18.11	2.81	-0.04
Phoneme segmentation	17.27	3.38	16.59	3.80	0.18
Follow-up II (time 4)					
Fluency of word reading II	22.73	11.98	28.89	16.55	-0.37
Spelling	27.87	7.75	28.70	7.47	-0.11
Productive letter knowledge—accuracy	32.70	1.26	32.52	2.19	0.08
Productive letter knowledge—speed <sup>a</sup>	33.03	13.41	32.33	8.12	-0.09

<sup>a</sup> The lower the score, the faster the letters were produced

<sup>b</sup> Calculation based on control group SD



**Fig. 1** Progress in receptive letter knowledge



condition were significant. Apparently, the early head start in receptive letter knowledge of the experimental group did not last long because most children learned the letters in the first half of first grade. The figure clearly illustrates the ceiling effect, as the maximum score of the test was 32.

#### *Second follow-up in first grade (time 4)*

The second part of Table 6 shows the means and standard deviations on the final follow-up of June first grade (time 4). It is clear from these data that the experimental group did not outperform the control group on fluency of word reading, spelling and productive letter knowledge speed. An independent  $t$  test indicated that the same is true for productive letter knowledge accuracy,  $t(55)=0.39$ , ns,  $d=0.08$ . This was to be expected as the variable was already skewed at time 3 and reached ceiling at time 4, skewness (SE) =  $-1.84$  (0.32).

On the speed measure of productive letter knowledge, the standard deviation in the experimental group was larger than in the control group. A Levene test confirms this,  $F(1, 55)=5.08$ ,  $p=0.028$ . Leaving out an outlier (with score 73) that partly causes this, reduced the mean and standard deviation of the experimental group,  $M$  (SD) =  $31.66$  (11.28) but did not give significantly different results.

#### Differences between subgroups

To be able to investigate whether weaker children profit equally from the intervention in comparison to less weak children, the data were split based on the median of the pre-test score on a phoneme awareness and letter knowledge composite. This resulted in two subgroups. The relatively good-performing group consisted of 15 children from the experimental group and 14 children from the control group. The relatively weak-performing

group consisted of 15 children from the experimental group and 13 children from the control group. When 'level' (good versus weak) was put in the repeated measures analyses as an extra between factor, there turned out to be a significant three-way interaction of time (pre- to post-test)  $\times$  condition  $\times$  level on receptive letter knowledge,  $F(1, 53)=5.12$ ,  $p<0.05$ ,  $d=0.62$ . Separate analyses on the relatively weak and the relatively good group showed that in the weaker half, the experimental group made considerably more progress than the control group,  $F(1, 26)=17.86$ ,  $p<0.001$ ,  $d=1.66$ , while in the relatively good-performing half, the progress did not differ between the conditions.

Analogous to these results, we also found a three-way interaction of time (pre- to post-test)  $\times$  condition  $\times$  level on the phoneme awareness composite,  $F(1, 53)=3.95$ ,  $p=0.052$ ,  $d=0.54$ . Again, separate analyses on the relatively weak and the relatively good group showed that, in the weaker half, the experimental group made more progress than the control group,  $F(1, 26)=5.11$ ,  $p<0.05$ ,  $d=0.89$ , while in the relatively good-performing half, there was no significant difference between the conditions.

We did not find significant interactions of condition  $\times$  level on the reading and spelling measures in first grade, meaning that the (lack of) difference between the experimental group and the control group was the same for the good-performing or the weaker-performing half.

## Correlations

Because we did not find significant transfer effects of the training, we calculated Pearson correlations to verify the assumption that there is a relation between the trained skills phoneme awareness and letter knowledge in kindergarten and reading and spelling in first grade. Table 7 presents Pearson correlations, zero-order and partial, between the trained skills phoneme awareness (composite) and letter knowledge and the untrained skill rapid naming (composite) at time 2 and spelling and fluency of word reading at time 4. The correlations were tested one-tailed, as positive correlations were expected.

**Table 7** Pearson  $r$  correlations between phoneme awareness, letter knowledge and rapid naming at time 2 and spelling and reading at time 4 ( $N=57$ )

Variables	1	2	3	4	5
Zero-order correlations					
1. PA (time 2)	–	0.73***	0.25*	0.28*	0.23*
2. LK (time 2)		–	0.35*	0.36*	0.30*
3. RAN (time 2)			–	0.34*	0.36*
4. Spelling (time 4)				–	0.63***
5. Fluency of word reading (time 4)					–
Partial correlations					
PA controlled for LK (time 2)				0.03	0.02
PA controlled for RAN (time 2)				0.22	0.16
LK controlled for PA (time 2)				0.23*	0.19
LK controlled for RAN (time 2)				0.27*	0.20
RAN controlled for PA and LK (time 2)				0.24*	0.28*

PA phoneme awareness, LK letter knowledge, RAN rapid naming

\* $p<0.05$ ; \*\* $p<0.01$ ; \*\*\* $p<0.001$  (one-tailed)

The first part of Table 7 shows that phoneme awareness and letter knowledge, as well as rapid naming at time 2, were correlated significantly with spelling and reading skills at time 4. To exclude the influence of the overlap between the skills at time 2 in the correlations with the skills at time 4, partial correlations were calculated. Phoneme awareness and letter knowledge were strongly intercorrelated ( $r=0.73$ ). Partial correlations show that the correlation of phoneme awareness with both spelling and reading vanished when it was controlled for letter knowledge ( $r=0.02$  and  $r=0.03$ ). The influence of letter knowledge remained significant in the correlation with spelling skill, when it was controlled for phoneme awareness ( $r=0.23$ ) but not in the correlation with reading.

The correlations of both phoneme awareness and letter knowledge with reading skill were not significant any more when they were controlled for rapid naming ( $r=0.16$  and  $r=0.20$ ). The correlation of rapid naming with reading skill was still significant when it was controlled for both phoneme awareness and letter knowledge ( $r=0.28$ ). This strongly indicates the strength of rapid naming as a predictor for reading skill.

The correlation between spelling skill with letter knowledge remained significant when it was controlled for rapid naming ( $r=0.27$ ). The correlation between spelling with phoneme awareness, when controlled for rapid naming, was not significant any more ( $r=0.22$ ). The correlation of rapid naming with spelling skill, when it was controlled for both phoneme awareness and letter knowledge, was still significant ( $r=0.24$ ). Therefore, both letter knowledge and rapid naming uniquely contributed to the prediction of later spelling skill.

### Stability of reading skill

To investigate the stability of first grade reading skill over time, most children in our sample ( $N=53$ ) were tested again on fluency of word reading II at the end of second grade (time 5). The large Pearson correlation between fluency of word reading II at time 4 and 5 ( $r=0.90$ ) suggests that the individual differences in reading skill within this group of children are quite stable.

To provide information about the progress in reading skill of our at-risk sample in comparison to unselected children, the distribution of the participants over five percentile classes were calculated at time 4 and 5 (see Table 8). These percentile classes were based on a normative sample of unselected children with the same educational age, as presented by the test constructors (Verhoeven, 1995). According to the norms, class A includes percentile scores above 75, class B includes percentile 50–75, class C 25–50, class D 10–25 and class E includes 10 and below. Table 8 shows that, at time 5, less children of the at-risk sample (both in the experimental group and in the control group) performed in the highest class and more children of this group performed in the lowest class in comparison to time 4 (16% versus 31% at class A and 28% versus 15% at class E for the total sample). In comparison to the normative sample, the relative reading level of the entire at-risk group declined from the end of grade 1 to end of grade 2. Because the results in the experimental and control groups were comparable, they have been combined.

## Discussion

After training, the experimental group showed more progress than the control group on receptive letter knowledge. The training effect was considerable ( $d=0.88$ ). The combined effect on phoneme awareness was small ( $d=0.25$ ) and mainly caused by a moderate effect

**Table 8** Distribution in proportions over different percentile classes on fluency of word reading II at the end of first grade (time 4) and the end of second grade (time 5)

Classes	Percentile	Total sample, $N=53$	
		First grade (%)	Second grade (%)
A	>75	31	16
B	>50–75	13	20
C	>25–50	26	21
D	>10–25	15	15
E	≤10	15	28

on phoneme segmentation ( $d=0.43$ ). Although the effects on PA were not statistically significant with an alpha of 0.05, they are potentially meaningful given the relatively small sample size of the study. Interpreting insignificant results when effects are in the expected direction and of comparable size as in similar intervention is in line with the call that research should focus on effect size and practical significance and not only statistical significance (e.g. Kirk, 2001; Wilkinson & the Task Force on Statistical inference, 1999).

These effects resulted in moderately and significantly higher scores on the receptive letter knowledge measure at post-test ( $d=0.48$ ) and in higher, but not significantly higher, scores in phoneme awareness ( $d=0.25$ ). On an untrained skill, rapid naming, the experimental group did not make more progress. These results supported our expectations, with the exception of the relatively small effect on phoneme awareness. It can be concluded that the effects were specific because there was no effect on rapid naming, an untrained skill. Within the subgroup of children at 'double risk' who combined familial risk with low pre-reading skills at pre-test (the weakest half of the sample), the training effects for letter knowledge and phoneme awareness at post-test were both large and significant. This is encouraging, since these children needed the intervention most. Contrary to our expectations, there was no effect on phoneme awareness and letter knowledge in first grade. Because of the skewness of all accuracy measures on phoneme awareness and letter knowledge at that time, the tests did not sufficiently discriminate between the better scoring children, and therefore, we were not sufficiently able to investigate effects on these skills in first grade. More importantly, there was no transfer to reading and spelling. The experimental group did not seem to benefit from their head start in letter knowledge and phoneme awareness.

The specific moderate to small effects on letter knowledge and phoneme awareness in the total sample and large effects in the subgroup of weakest children suggest that the immediate goals of the programme were achieved. In addition, we concluded that parents were able to train their children at familial risk effectively in these skills. In the study of van Otterloo et al. (2006), it was demonstrated that the treatment integrity was sufficient.

In comparison to other intervention studies with children at familial risk (Table 1), the immediate effect on letter knowledge in the current study was larger than in the studies of Elbro and Petersen (2004), Eleveld (2005), Hindson et al. (2005), and van Otterloo et al. (2009) but not larger than in the study of Regtvoort and van der Leij (2007), which gives an effect of comparable size. The effects on phoneme awareness were, except for a moderate effect on phoneme segmentation, smaller than the studies in Table 1 that reported effect size.

However, the transfer effects at follow-up on reading and spelling measures that were moderate in the Danish study of Elbro and Petersen (2004) and the Dutch study of van

Otterloo et al. (2009) were not replicated in the present study. This finding is in line with the Dutch studies of Eleveld (2005) and Regtvoort and van der Leij (2007) that also did not demonstrate transfer to reading and spelling. The moderate transfer effect of our earlier study (van Otterloo et al.) was not statistically significant and can therefore be attributed to chance. The Australian study of Hindson et al. (2005) did not report effect sizes of follow-up results. In the other Australian study of Fielding-Barnsley and Purdie (2003), there is a suggestion of transfer to reading and spelling, but no real comparison between groups is possible because of the absence of the pre-test. It can be concluded that, at least in comparison to the Danish study, the Dutch studies do not support the assumption that early training of phoneme awareness, and letter knowledge of 8–12.5 h results in better reading and spelling skills of children at familial risk of dyslexia.

To answer the question why there were no transfer effects in this Dutch study, it is important to include the results of the intervention studies mentioned in Table 1 that were implemented in The Netherlands. The disappointing findings of four Dutch intervention studies that were comparable in methods and samples, and only differed in content of the programme and mode of instruction (computer-assisted or not; at school or at home), invite us to reconsider the possibility of prevention of dyslexia in The Netherlands in carefully selected at-risk children.

The first point to consider is that the at-risk samples of the four studies were very specific. Not only were the children derived from families with a history of reading disabilities, but many of them also had relatively highly educated parents, possibly because participation was voluntary. As a consequence, the scores on the receptive vocabulary and non-verbal IQ measure were above mean. Effects of the study can therefore not be compared to studies using other kinds of samples. The study of Reitsma and Wesseling (1998), for example, did show transfer effects on first grade reading measures of computer-assisted training in phoneme blending in kindergarten. Their study was also implemented in The Netherlands, but their sample was considerably different. The participants were of comparable age but were not tested on familial risk, and 32% of the sample had a different ethnic background. Moreover, at pre-test, the children recognised about six or seven letters, which was fewer than the 12 or 13 letters in our sample. It may be assumed that there was more to gain in the study of Reitsma and Wesseling because of the different social background and mastery level of the participants.

However, it is possible to compare the results of the current study and the studies of Eleveld (2005), Regtvoort and van der Leij (2007), and van Otterloo et al. (2009) to studies carried out in other countries that also used children at familial risk in the pre-reading phase (Elbro & Petersen, 2004; Fielding-Barnsley & Purdie, 2003; Hindson et al., 2005). To be able to draw cross-linguistic comparisons, differences in language and teaching methods have to be taken into account. To investigate this, it was decided to gather additional information about Denmark from Elbro (February 2005) and about Australia from Byrne (June 2005). Table 9 gives an overview of the most important differences between the countries.

It is evident that, in comparison to both English and Danish, Dutch has a rather shallow orthography, although less shallow than for instance Spanish or Finnish (Seymour, Aro, & Erskine, 2003). As a consequence, learning to read and spell takes less effort in countries with a shallow language such as The Netherlands because it can rely more on phonological decoding and children learn to read faster. Correspondingly, these children master phoneme awareness early, which makes it a less strong predictor of later reading ability (de Jong & van der Leij, 1999; Wimmer, Mayringer, & Landerl, 2000). The finding of the present study that phoneme awareness at the end of kindergarten did not correlate with reading and spelling at the end of second grade when controlled for letter knowledge supports this view. Consequently, training phoneme awareness may be less effective.

**Table 9** Aspects of the countries involved in intervention studies with children at familial risk of dyslexia

Aspects	Countries		
	Denmark	Australia	The Netherlands
Orthography	Deep	Deep	Relatively shallow
Age first year formal RI	6–7 years	4–5 years	5–6 years
Phonics in regular RI	About 20% of teaching time	About 30% of teaching time	Dominant teaching method
<i>N</i> hour RI in 1st year formal RI	3–4 h per week	About 5 h per week	6.5–7 h per week
PA and LK in year prior to formal reading instruction	<1 h per week	Little	About 1 h per week

*RI* reading instruction, *PA* phoneme awareness, *LK* letter knowledge

Possibly in relation with orthography, there also seem to be some differences in teaching across countries. In comparison to The Netherlands, the children in Australia are a year younger and the children in Denmark are a year older when formal reading instruction begins. More importantly, practice of grapheme–phoneme correspondences (phonics) is the dominant teaching method in learning to read Dutch in first grade. The same is true for learning to read in most other shallow orthographies. In both Denmark and Australia, phonics plays a much less prominent role in learning to read, according to Elbro and Byrne. Instead, more time is spent on whole language and sight words. In addition, the total time that is spent on reading instruction in first grade ranges from about 3–4 h per week in Denmark to 6.5–7 h per week in The Netherlands. In Australia, formal reading instruction starts at the beginning of the first school year, which they call kindergarten. Right from the start, the children get about 5 h of reading instruction per week. In the year prior to formal reading instruction, there is only some attention paid to pre-reading skills in all three countries. In Australia, the year prior to kindergarten is called preschool. Not all children attend preschool and many children only for a few days a week. It may be assumed that, in the year prior to reading instruction, there are no important differences between the three countries. In first grade, however, there may be important differences. In The Netherlands, more time is spent on first grade reading instruction in general and on phonics in particular. This may explain the absence of (delayed) transfer effects on reading and spelling in The Netherlands. It has been suggested that most early reading difficulties are primarily caused by instructional deficits (see for instance Vellutino, Fletcher, et al. 2004). In combination with a relatively inconsistent orthography to be learned, insufficient or inadequately early reading instruction in Australia and Denmark may result in more children with reading disability or more children with persistent reading problems. As a consequence, early intervention may be more effective. However, these suggestions are not derived from a comparative study and therefore speculative and should be tested in a cross-linguistic study.

Because the present and our earlier study and the studies of Eleveld (2005) and Regtvoort and van der Leij (2007) indicate that early intervention of about 8–12.5 h, with a content mainly based on phoneme awareness and letter knowledge and with several instruction modes (computerised, at home with parents, or at school with professionals), does not result in effects on the longer term in samples of children who combine a familial risk of dyslexia with a relatively good verbal and non-verbal intelligence, the question may be raised on how to design intervention programmes that may be more effective. One option is a more intensive intervention. However, it seems very difficult to find parents willing to

cooperate to spend more than 10 min a day working with such a programme, especially when we keep in mind the age of the children and the busy family life. Adding an extra intervention (another 12.5 h) a year ahead as Eleveld (2005) did also did not give larger effects.

The results at the end of second grade (Table 8) may give a hint to another option. In comparison to the end of first grade, the group of at-risk children showed a declination of performance. The proportion of good readers decreased from 31% in first grade to only 16% in second grade, and the proportion of children with very low reading skills in the sample went up from 15% to 28% between first and second grade, indicating that familial risk may be accompanied by relatively late emerging reading problems in a substantial subgroup of at-risk children in a fairly consistent orthography such as Dutch. The increase in low scoring children is in line with the results of the study by Leach, Scarborough, and Rescola (2003) who reported late-emerging reading disabilities. One group of children in their sample with reading disabilities was identified only after the third grade, in contrast to children with early identified reading disability. There was, however, a difference between the two groups in reading scores: The children with late-emerging reading disabilities were relatively better. The authors suggest that this indicates that the reading disabilities were not just late identified but actually late emerging. The findings of the Dutch studies suggest that there is a greater chance of late manifestation of dyslexia in samples of children with familial risk but intellectual proficiency who encounter relatively favourable instructional circumstances in the pre-reading phase and in the first grade. As a consequence of the manifestation at a later stage, when the process of automatization should lead to higher fluency, it may be recommended that remedial instruction and practice should be extended from kindergarten to grades 1 and 2 to gain positive effects. Our findings that rapid naming is the strongest predictor of reading supports the view that training of fluency should be taken into account. To support this suggestion, Struiksma, van der Leij, and Stoel (2009) were able to show that about two thirds of the weakest Dutch readers in second grade (lowest 10%) can be remediated by an intensive and individual fluency-oriented intervention during a 6-month period (total training time, 30 h). The remaining third of these weakest readers (about 3–4% of all children) had a persistent reading disability and needed intensive treatment in a clinical setting.

Before we draw conclusions, some limitations of the present study should be acknowledged. The first limitation to mention is the skewness of the accuracy measures of phoneme awareness and letter knowledge at the first follow-up in first grade (time 3) and the accuracy measure of productive letter knowledge at the second follow-up (time 4). Because measurements approached ceiling, the tests did not sufficiently discriminate between children with high test scores and therefore limits the likelihood of finding any significant differences. As a consequence, we were not able to assess whether the lack of effect on reading and spelling in first grade was due to extinction of the direct training effect over time or to a lack of transfer from phoneme awareness and letter knowledge to reading and spelling.

Other limitations are related to the quasi-experimental design of the study. In this experimental field trial, we were not able to randomly assign the children to the experimental conditions. Instead, we worked with a historical control group. Random assignment is the best procedure to be reasonably confident that group differences at the end of the experiment are not the result of preexisting differences between the groups. These unintended differences between groups may have reduced internal validity (Cook & Campbell, 1979).

Among these is the threat of *selection*. Different kind of people might be interested in different experimental conditions. Because of the requested amount of effort, parents volunteering to participate in the experimental condition might be other kind of parents in comparison to parents of control children resulting in other home-literacy environments.

However, it must be noted that parents could not choose between conditions they were asked to participate in the training study (experimental condition) or they were asked to participate in a research programme that monitors the results of familial at-risk children (control condition). Moreover, several aspects of home literacy environment were measured (see Table 3), and no significant differences were found. The threat of *history*, events that coincides with the treatment and affect the outcome, is also possible in a quasi-experimental design. Because of the historical control group that is used, also an interaction of *selection* and *history* is possible. The different groups might experience other history events between pre- and post-test, which influences the groups differently. We have no knowledge of special events during this quasi-experiment that could have influenced the outcome. There is a threat of *instrumentation* because there was a change of part of the assistants over time. We minimized the effects of this threat by giving all testing assistants the same training.

However, to control for these threats to internal validity, groups were compared in important background variables (age, IQ, home literacy environment, see Tables 2 and 3) and pretest scores. There were no significant differences. Children were derived from many different schools, so also no systematic differences of school environment between groups were to be expected.

In sum, to return to the results of the present study, together with the results of Regtvoort and van der Leij (2007), Eleveld (2005) and van Otterloo et al. (2009), it may be concluded that early training of phoneme awareness and letter knowledge in The Netherlands gives specific and immediate results. However, these results do not lead to significant transfer effects on first grade measures of reading and spelling. As the direct training results on letter knowledge and phoneme awareness did not transfer to reading and spelling, the assumption that early training of 8–12.5 h of phoneme awareness and letter knowledge prevents reading problems of children at familial risk of dyslexia may not apply to The Netherlands, at least not to a sample with relatively good intelligence and educational background of the parents. Because most Dutch children seem to master phoneme awareness relatively easy and early, and the reading instruction in first grade is intensive with an emphasis on phonics, the training seems also less effective than in studies in countries with different circumstances. It may be suggested that the focus of early intervention in The Netherlands (and countries with comparable orthographies and teaching methods) should not be solely on pre-reading skills as phoneme awareness and letter knowledge but also more directly on the initial reading skills (recoding of words that meet the grapheme–phoneme correspondence rules) and on fluency of reading at the lexical and sublexical level (see for an example Struiksmá et al., 2009). To support this suggestion, the findings that the relative reading level of the entire at-risk group declined and that reading disabilities of a substantial proportion of the children emerged in the course of second grade indicate that continuation of intervention or extra remediation is necessary for many at-risk children.

**Acknowledgements** The present study is associated with the research programme “Early diagnosis and treatment of developmental dyslexia: The role of learning mechanisms, speech manipulation, phonological skills and rapid naming”, partly granted by the Netherlands Organization for Scientific Research (NWO). The research programme is a cooperation between the University of Groningen, the Radboud University (Nijmegen), and the University of Amsterdam.

Part of the results of the current study are previously presented at the Sixth BDA International Conference (van Otterloo & van der Leij, 2004) and at the Eleventh Annual Meeting of TripleSR (van Otterloo, Regtvoort, & van der Leij, 2004).

We thank the participating children, parents and schools for their time and effort. We thank prof. Carsten Elbro and prof. Brian Byrne for their information on the early reading education in Denmark and Australia respectively. We thank Judith Bekebrede for collecting and supplying data of the final follow-up. We thank the students and research assistants including Nienke Bolderheij, Sharon Geurtsen, Inge Klaij, Anna Pires, Esther Veldkamp, and Marieke van Wees, for their help in collecting the data.



**Open Access** This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

## References

- Ball, E. W., & Blachman, B. A. (1991). Does phoneme awareness training in kindergarten make a difference in early word recognition and developmental spelling? *Reading Research Quarterly*, 26, 49–66.
- Beck, I. (1989). *Reading today and tomorrow: Teachers edition for grades 1 and 2*. Austin: Holt.
- Borström, I., & Petersen, D. K. (1996). *På vej til den første læsning. Fonologisk opmærksomhed [On the way to initial reading. Phonological awareness]*. Copenhagen: Alinea.
- Borström, I., & Elbro, C. (1997). Prevention of dyslexia in kindergarten: Effects of phoneme awareness training with children of dyslexic parents. In C. Hulme & M. Snowling (Eds.), *Dyslexia: Biology, cognition and intervention* (pp. 235–253). London: Whurr.
- 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.
- Bradley, L., & Bryant, P. (1985). *Rhyme and reason in reading and spelling*. Ann Arbor: University of Michigan Press.
- Brus, B. T., & Voeten, M. J. M. (1980). *Eén-minuut-test, vorm a en b. Verantwoording en handleiding [One-minute-test, form a and b. Justification and manual]*. Berkhout: Nijmegen.
- Bus, A. G., & van IJzendoorn, M. H. (1999). Phonological awareness and early reading: a meta-analysis of experimental training studies. *Journal of Educational Psychology*, 91, 403–414.
- Byrne, B., & Fielding-Barnsley, R. (1993). Evaluation of a programme to teach phonemic awareness to young children: A 1-year follow-up. *Journal of Educational Psychology*, 85, 104–111.
- Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation. Design and analysis issues for field settings*. Chicago: Rand McNally College.
- 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.
- Ehri, L. C., Nunes, S. R., Willows, D. M., Schuster, B. V., Yaghoub Zadeh, Z., & Shanahan, T. (2001). Phonemic awareness instruction helps children learn to read: evidence from the national reading panel's meta-analysis. *Reading Research Quarterly*, 36, 250–287.
- Elbro, C., & Petersen, D. K. (2004). Long-term effects of phoneme awareness and letter sound training. An intervention study with children at risk of dyslexia. *Journal of Educational Psychology*, 96, 660–670.
- Elbro, C., & Scarborough, H. S. (2004a). Early identification. In P. Bryant & T. Nunes (Eds.), *International handbook of children's reading* (pp. 339–359). Dordrecht: Kluwer.
- Elbro, C., & Scarborough, H. S. (2004b). Early intervention. In P. Bryant & T. Nunes (Eds.), *International handbook of children's reading* (pp. 361–381). Dordrecht: Kluwer.
- Eleveld, M. A. (2005). *At risk for dyslexia. The role of phonological abilities, letter knowledge, and speed of serial naming in early intervention and diagnosis*. Groningen: Garant.
- Felton, R. H., & Wood, F. B. (1989). Cognitive deficits in reading disability and attention deficit disorder. *Journal of Learning Disabilities*, 22, 3–13.
- Fielding-Barnsley, R., & Purdie, N. (2003). Early intervention in the home for children at risk of reading failure. *Support for Learning*, 18, 77–82.
- Finucci, J. M., Gottfredson, L. S., & Childs, B. (1985). A follow-up study of dyslexic boys. *Annals of Dyslexia*, 35, 117–136.
- Gilger, J. W., Pennington, B. F., & DeFries, J. C. (1991). Risk for reading disability as a function of parental history in three family studies. *Reading and Writing*, 3, 205–217.
- Hatcher, P. J., Hulme, C., & Ellis, A. W. (1994). Ameliorating early reading failure by integrating the teaching of reading and phonological skills: the phonological linkage hypothesis. *Child Development*, 65, 41–57.
- Hindson, B. A., Byrne, B., Fielding-Barnsley, R., Newman, C., Hine, D. W., & Shankweiler, D. P. (2005). Assessment and early instruction of preschool children at risk for reading disability. *Journal of Educational Psychology*, 97(4), 687–704.
- Kirk, R. E. (2001). Promoting good statistical practices: Some suggestions. *Educational and Psychological Measurement*, 61, 213–218.

- Koster, C., Been, P. H., Krikhaar, E. M., Zwarts, F., Diepstra, H. D., & van Leeuwen, T. H. (2005). Differences at 17 month: productive language patterns in infants at familial risk for dyslexia and typically developing infants. *Journal of Speech, Language, and Hearing Research, 48*, 426–438.
- Leach, M. J., Scarborough, H. S., & Rescorla, L. (2003). Late-emerging reading disabilities. *Journal of Educational Psychology, 95*, 211–224.
- Moelands, F., & Kamphuis, F. (2001). *Verantwoording van de toetsen uit de pakketten Schaal vorderingen in spellingvaardigheid (SVS) 1, 2 en 3 [Justification of the spellingtests SVS 1, 2, and 3]*. Arnhem: CITO.
- O'Connor, R. E., Jenkins, J. R., Leicester, N., & Slocum, T. A. (1993). Teaching phonological awareness to young children with learning disabilities. *Exceptional Children, 59*, 532–546.
- Petersen, D. K. (2000). What do you get if you add/mmm/to ice? Training phoneme awareness in kindergarten: An intervention study with children of dyslexic parents. In N. A. Badian (Ed.), *Prediction and prevention of reading failure* (pp. 247–273). Baltimore: York Press.
- Raven, J. C., Court, J. H., & Raven, J. (1984). *Manual for Raven's progressive matrices and vocabulary scales*. London: Lewis.
- Regtvoort, A. G. F. M., & van der Leij, A. (2007). Early intervention with children of dyslexic parents: Effects of computer-based reading instruction at home on literacy acquisition. *Learning and Individual Differences, 17*(1), 35–53.
- Reitsma, P., & Wesseling, R. (1998). Effects of computer-assisted training of blending skills in kindergartners. *Scientific Studies of Reading, 2*(4), 301–320.
- Rosenthal, R., Rosnow, R. L., & Rubin, D. B. (2000). *Contrasts and effect sizes in behavioral research. A correlational approach*. Cambridge: Cambridge University Press.
- Scarborough, H. S. (1990). Very early language deficits in dyslexic children. *Child Development, 61*, 1728–1743.
- Schneider, W., Roth, E., & Ennemoser, M. (2000). Training phonological skills and letter knowledge in children at risk for dyslexia: A comparison of three kindergarten intervention programmes. *Journal of Educational Psychology, 92*, 284–295.
- Sénéchal, M., & LeFevre, J. (2002). Parental involvement in the development of children's reading skill: A five-year longitudinal study. *Child Development, 73*(2), 445–460.
- Seymour, P. H., Aro, M., & Erskine, J. M. (2003). Foundation literacy acquisition in European orthographies. *British Journal of Psychology, 94*, 143–174.
- Snowling, M. J. (2000). *Dyslexia* (2nd ed.). Malden: Blackwell.
- Struiksma, A. J. C., van der Leij, A., & Stoel, R. (2009). Response to fluency-oriented intervention of Dutch poor readers. *Learning and Individual Differences*, doi:10.1016/j.lindif.2009.07.005.
- Torppa, M., Poikkeus, A.-M., Laakso, M.-L., Eklund, E., & Lyytinen, H. (2006). Predicting delayed letter knowledge and its relation to grade 1 reading achievement among children with and without familial risk for dyslexia. *Developmental Psychology, 42*, 1128–1142.
- Troia, G. A. (1999). Phonological awareness intervention research: A critical review of the experimental methodology. *Reading Research Quarterly, 34*(1), 28–52.
- van Bon, W. H. J. (1986). *Raven's Coloured Progressive Matrices; Nederlandse normen en enige andere uitkomsten van onderzoek [Dutch standards and some other outcomes of research]*. Lisse: Swets & Zeitlinger.
- van den Bos, K. P. (1998). IQ, phonological awareness and continuous-naming speed related to Dutch poor decoding children's performance on two word identification tests. *Dyslexia, 10*, 265–288.
- van den Bos, K. P. (2003). *Snel Serieel Benoemen; Experimentele versie. [Rapid naming; Experimental version]*. Groningen: University of Groningen.
- van den Bos, K. P., Iutje Spelberg, H. C., Scheepstra, A. J. M., & de Vries, J. R. (1994). *De Klepel. Vorm a en b. Verantwoording, handleiding, diagnostiek en behandeling. [The Clapper. Form a and b. Justification, manual, diagnostics, and treatment.]*. Nijmegen: Berkhout.
- van den Bosch, L., Gillijns, P., Krom, R., & Moelands, F. (1993). *Schaal vorderingen in spellingvaardigheid 1. Handleiding. [scale progress in spelling skill. Manual.]*. Arnhem: CITO.
- van Otterloo, S. G., & van der Leij, A. (2004, March). Prevention of Developmental Dyslexia. Paper presented at the Sixth BDA International Conference, Coventry, UK.
- van Otterloo, S. G., Regtvoort, A. G. F. M., & van der Leij, A. (2004, June). Does early intervention make a difference for Dutch children at risk: A comparison of two intervention programs. Poster session presented at the Eleventh Annual Meeting of TripleSR, Amsterdam, The Netherlands.
- van Otterloo, S. G., van der Leij, A., & Henrichs, L. F. (2009). Early home-based intervention in The Netherlands for children at familial risk of dyslexia. *Dyslexia, 15*, 187–217.
- van Otterloo, S. G., van der Leij, A., & Veldkamp, E. (2006). Treatment integrity in a home-based pre-reading intervention program. *Dyslexia, 12*, 155–176.

- Vellutino, F. R., Fletcher, J. M., Snowling, M. J., & Scanlon, D. M. (2004). Specific reading disability (dyslexia): What have we learned in the past four decades? *Journal of Child Psychology and Psychiatry*, *45*, 2–40.
- Verhoeven, L. (1993a). *Grafemtoets. Toets voor auditieve synthese. Handleiding. [Grapheme test. Test for phoneme blending. Manual]*. Arnhem: CITO.
- Verhoeven, L. (1993b). *Toets voor auditieve analyse. Fonemendictee. Handleiding [Test for phoneme segmentation. Dictation of phonemes. Manual]*. Arnhem: CITO.
- Verhoeven, L. (1995). *Drie-minuten-toets. Handleiding [three-minutes-test. Manual]*. Arnhem: CITO.
- Verhoeven, L., & van Leeuwe, J. (2003). Ontwikkeling van decodeervaardigheid in het basisonderwijs [Development of recoding skill in primary education]. *Pedagogische Studiën [Pedagogical Studies]*, *80*, 257–271.
- Verhoeven, L., & Vermeer, A. (1996). *Taaltoets Allochtone Kinderen. Handleiding [Language test for children of ethnic minorities. Manual]*. Tilburg: Zwijssen.
- Wechsler, D. (1981). *WAIS-R: Wechsler Adult Intelligence Scale-Revised*. San Antonio, TX: Psychological Corporation.
- Whitehurst, G. J., Arnold, D. S., Epstein, J. N., Angell, A. L., Smith, M., & Fischel, J. E. (1994). A picture book reading intervention in day care and home for children from low-income families. *Developmental Psychology*, *30*, 679–689.
- Wilkinson, L., & the Task Force on Statistical Inference. (1999). Statistical methods in psychology journals: Guidelines and explanations. *American Psychologist*, *54*, 594–604.
- Wimmer, H., Mayringer, H., & Landerl, K. (2000). The double-deficit hypothesis and difficulties in learning to read a regular orthography. *Journal of Educational Psychology*, *92*, 668–680.
- Wolf, M. (1991). Naming speed and reading: The contribution of the cognitive neurosciences. *Reading Research Quarterly*, *26*, 123–141.