Scientific Studies of Reading
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/hssr20

Gains From Training in Phonological Awareness in Kindergarten Predict Reading Comprehension in Grade 9

Ann-Christina Kjeldsen[a], Antti Kärnä[b], Pekka Niemi[b], Åke Olofsson[c] & Katarina Witting[a]

[a] Åbo Akademi University

[b] University of Turku

[c] University of Umeå

Published online: 20 Aug 2014.

To cite this article: Ann-Christina Kjeldsen, Antti Kärnä, Pekka Niemi, Åke Olofsson & Katarina Witting (2014) Gains From Training in Phonological Awareness in Kindergarten Predict Reading Comprehension in Grade 9, Scientific Studies of Reading, 18:6, 452-467, DOI: 10.1080/10888438.2014.940080

To link to this article: http://dx.doi.org/10.1080/10888438.2014.940080

Scientific Studies of Reading, 18:452–467, 2014

ISSN: 1088-8438 print/1532-799X online
DOI: 10.1080/10888438.2014.940080

Routledge
Taylor & Francis Group

Gains From Training in Phonological Awareness in Kindergarten Predict Reading Comprehension in Grade 9

Ann-Christina Kjeldsen
Åbo Akademi University

Antti Kärnä and Pekka Niemi
University of Turku

Åke Olofsson
University of Umeå

Katarina Witting
Åbo Akademi University

The effects of a kindergarten training program in phonological awareness with 209 Swedish-speaking children were followed up until the end of Grade 9. Initial levels of letter knowledge and phonological awareness were positively associated with the level of decoding skill in Grade 3 but not with its growth afterward. The intervention group performed significantly better in decoding in Grade 3, and the difference was maintained until Grade 6. The trained children also scored higher in Grade 9 reading comprehension. Although the results give empirical support for a connection between early phonological awareness training, later word decoding development, and still later reading comprehension, the theoretical explanation for the link between especially word decoding and reading comprehension is far from clear.

An early diagnosis of prospective reading difficulties is a strong tool in the hands of teaching professionals, provided that it is promptly followed by a well-informed intervention, preferably before the start of schooling. To this end, a sizeable body of knowledge relevant for beginning reading has been produced. The most popular topic is undoubtedly early training of phonological awareness.

The developmental structure of phonological awareness has been debated: Is it a single ability or composed of a set of abilities? If phonological sensitivity consists of distinct abilities, these may be differentially related to reading acquisition and efforts to promote it by phonological training. In the worst scenario the structure of phonological sensitivity may depend on orthographic complexity. These issues have evoked considerable interest but so far no consensus has been

Correspondence should be sent to Ann-Christina Kjeldsen, V.Esplanadg. 12 A, FIN-22100 Mariehamn, Åland, Finland. E-mail: annchristinakjeldsen@gmail.com

reached. Stances have varied from a single-factor construct (Anthony & Lonigan, 2004; Anthony et al., 2002; Lonigan, Burgess, & Anthony, 2000; Stahl & Murray, 1994), through two-factor construct in terms of rime awareness and phonological awareness (Carroll, Snowling, Hulme, & Stevenson, 2003; Muter, Hulme, Snowling, & Taylor, 1998) to a three-factor solution structure consisting of phonological sensitivity to rhyme, syllables, and phonemes (Høien, Lundberg, Stanovich, & Bjaalid, 1995). The circle is closed by a recent longitudinal study of 5- to 7-year-old Greek-speaking children concluding that phonological sensitivity is best understood as a unified construct instead of a set of distinct abilities (Papadopoulos, Kendeou, & Spanoudis, 2012). Moreover, children learning to read in another transparent orthography, Finnish, proceed toward phonological awareness starting from supraphonemic units at the age of 3 to 4 years and reaching the phonemic level about two years later (Puolakanaho, Poikkeus, Ahonen, Tolvanen, & Lyytinen, 2004; Silvén, Poskiparta, & Niemi, 2004).

The debate just presented makes sense if interventions are targeted on but a few subcomponents of phonological awareness. However, the issue can easily be circumvented by deploying a wide range of tasks from rhyming to phonemic exercises as was done by Lundberg, Frost, and Petersen (1988) on the Danish island of Bornholm. Because the ultimate goal of reading instruction is reading comprehension, often described as the product of decoding and language comprehension (Gough & Tunmer, 1986), we argue that studies of phonological intervention do not share the theoretical interest with the structurally guided research on phonological awareness. Recent research suggests that early training of phonological awareness is a causally efficient component in the development of children's early word-level reading skill (e.g., Hulme, Bowyer-Crane, Carroll, Duff, & Snowling, 2012). There is evidence in support of this argument. Bridging the gap between letter strings and spoken words is assumed be based on improved phonological representations (Elbro, Borstrøm, & Petersen, 1998; Elbro & Jensen, 2005), a necessary prerequisite for the quality of word representations in the lexicon (Perfetti, 2007). In their extensive meta-analysis, Melby-Lervåg, Halaas Lyster, and Hulme (2012) concluded that the way to good word reading skills goes via phonemically structured phonological representations. Perfetti and Stafura (2014) extended the argument by stating that word knowledge based on good phonological quality plays a pivotal role in the development of reading comprehension. This view highlights the importance of vocabulary for reading development as also suggested by an extensive longitudinal study (Verhoeven & van Leeuwe, 2011). Thus, if early training of phonological awareness is causally related with the development of children's early word-level reading skill, a question arises about its role when decoding becomes more fluent and accurate (Verhoeven & van Leeuwe, 2009) and variables associated with general language comprehension gradually tend to become more important to reading comprehension (e.g., Cutting & Scarborough, 2006; National Institute of Child Health and Human Development Early Child Care Research Network, 2005; Storch & Whitehurst, 2002). Thus, our study was inspired by the question whether traces of early training of phonological awareness can be found in decoding development in the elementary school and in reading comprehension thereafter. The problem statement is significant in view of the evidence suggesting that, in the absence of proper intervention, those children who occupy the lowest ranks in terms of reading ability by the end of Grade 1 tend to remain in this relative position for the whole school career (de Jong & van der Leij, 2003; Juel, 1988; Landerl & Wimmer, 2008; Lundberg, 1994). There is therefore considerable theoretical and practical interest to answer the question whether initial gains acquired by phonological training prior to Grade 1 can be maintained through the school years.

Given that kindergarten training has proven positive effects on learning to decode which, in turn, is a fundamental component of early reading comprehension, then also theories of print exposure (Cunningham & Stanovich, 1997; Sparks, Patton, & Murdoch, 2014; Stanovich & West, 1989) point to the importance of investigating long-term effects of early phonological awareness training. There is still no broad theory of reading that can explain the full width and complexity of reading development from Grade 1 to the threshold of adolescence, although many components are known (Perfetti & Stafura, 2014). Nonetheless, a first step would be to clarify if there are long-term effects of kindergarten training, via word-level reading on later reading comprehension. Somewhat unexpectedly, such long-term prospective training studies are absent even during the 1980s and 1990s when training of phonological awareness was new and fashionable. More specifically, the reported follow-up intervals were short, never exceeding 3 years, as indicated by two meta-analyses and a review: Bus and van IJzendoorn (1999; 32 intervention studies reviewed, 9 longitudinal), Ehri et al. (2001; 52/8), Schneider and Stengård (2000; 44/9). What is more, interest in long-term effects of phonological training appears rather to have declined than increased during the past decade, in spite of the tremendous effort put into training studies. In a recent review based on 97 studies, Slavin, Lake, Davis, and Madden (2011) noted that very few long-term evaluations of phonics-based training have been carried out. The ones they refer to were based on Reading Recovery. Initial gains from first-grade tutoring were robust as compared to nontutored peers. However, the gains largely dissipated in the course of 3 to 4 years resulting in standardized effect sizes of .15 and .20, respectively (p. 19). The small number of relevant studies is surprising given the indisputable initial training gains (e.g., Hulme et al., 2012) but rather equivocal findings concerning their maintenance. There are but three phonological training studies with an unusually long follow-up period. In Australia, Byrne, Fielding-Barnsley and Ashley (2000) observed improved word reading ability up to Grade 5 in an unselected sample. A Danish study found a significant training effect on nonword reading speed for children with a familial risk of dyslexia as far as in Grade 7. For reading comprehension there was a nonsignificant tendency in Grade 7 favoring trained readers at risk over untrained readers at risk (see Table 3 in Elbro & Klint Petersen, 2004). Reading comprehension was not tested in earlier grades. It is worth noting that the Danish study reported no observations between Grades 3 and 7. The third study was conducted in the United Kingdom (Snowling & Hulme, 2011) and involved school beginners who were at risk on the basis of their performance in spelling. The intervention featured integrated phonological and reading-alone elements. Gains in single word reading ability were clear in the posttest, and they sustained after 11 months. Fifty-four of the original 75 participants could be contacted for a follow-up 6 years afterward with no signs of attrition bias. By this time, reading accuracy and comprehension were still within the typical average range, whereas fluency was below it.

The present study is inspired by that by Lundberg et al. (1988) carried out in Danish, an opaque orthography that is challenging to the beginning reader (e.g., Seymour, Aro, & Erskine, 2003). We have previously reported a successful replication with Swedish-speaking children showing maintained training effects on decoding until the end of Grade 2 (Kjeldsen, Niemi, & Olofsson, 2003). The present study reports the reading development of the same students until Grade 9. Specifically, we ask whether the gains in decoding are maintained from Grade 3 onward and whether they extend to reading comprehension in Grade 9.

METHOD

The study was conducted in the autonomous archipelago district of Åland in South-Western Finland. Its population of about 28,000 is up to 92% Swedish speaking, with Swedish also being the only language of teaching throughout the preschool and school. In terms of transparency, the Swedish orthography can be placed in the middle of the continuum extending from shallow to deep orthographies, and it is comparable to Dutch (Seymour et al., 2003). A short description of Swedish orthography can be found in Olofsson (2003). The province of Åland features the same high level of literacy as the rest of Finland. For this study, kindergarten groups were randomized into intervention and control conditions. The geographical distance between experimental and control groups varied, and some diffusion of treatment effects cannot be ruled out. By the time of the intervention in 1997–1998, knowledge of phonological awareness had penetrated the kindergarten and preschool staff, and exercises in it were common. This was likely to further counteract the possible gains due to training because even the control group came into contact with relevant language games.

According to the curriculum, the Swedish language is the most important subject in the comprehensive school of Åland, as evidenced by a generous amount of teaching (five to six lessons per week in Grades 1 through 6 and 3 lessons per week in Grades 7 through 9). Yearly screening tests are performed to identify students in need of individualized support in Swedish. The curriculum emphasizes book reading, the goal being an ability to discuss, analyze, and criticize the text. The role of library is underlined as well as skills related to information technology, nonfiction literature, word lists, dictionaries, newspapers, and magazines.

Participants

The 209 original participants (100 girls, 109 boys) in Kjeldsen et al.'s (2003) study comprised all students entering the kindergarten class in Åland in 1997. The intervention group consisted of 108 students and the control group of 101 students. At the beginning of the study in September 1997, they were on average 6 years 2 months old. At the time of data collection in Grade 9 they were on average 15 years 9 months old. It was ensured at the beginning of the study that Swedish had been mastered by all participants. There were 14 single-parent families in the intervention group and nine in the control group. Mothers' educational background was similar in both groups: compulsory education (50.5% vs. 50.3%), secondary (41.4% vs. 40.4%), and college/university (8.0% vs. 8.9%). Sampling of participants was based on 9 (intervention) and 13 (control) kindergarten classes. It was ensured that intervention and control groups did not differ in terms of average geographical distance between home and school (for details, see Kjeldsen et al., 2003).

There were no missing values for the independent variables (i.e., gender, intervention, phonological awareness, letter knowledge, and vocabulary). For dependent variables (i.e., decoding in Grades 3, 4, and 6 in addition to reading comprehension in Grade 9), the percentages of missing values in the intervention and control groups ranged from 6.9 to 21.3, depending on the study group and outcome (see Table A1 in the appendix). In general, there was somewhat more attrition among the students with low scores in prereading skills and for them especially in the intervention group. The rate of missing data was, however, usually at most 10%, and the missingness was clearly associated with the available kindergarten measures: Students with

missing data in both groups scored lower in letter knowledge and phonological awareness (see the appendix, Table A1). We therefore performed multiple imputation of missing data using Bayesian analysis (Rubin, 1987; Schafer, 1997). Specifically, the missing data were imputed from an unrestricted H1 model by storing the generated missing data values at every 300th iteration after the MCMC sequence had converged. In this manner, altogether 50 imputed data sets were created. The imputation variance–covariance model used included the four variables with missing data and a number of complete variables (i.e., phonological awareness, letter knowledge, vocabulary, intervention as well as the two interaction components Phonological Awareness × Intervention and Letter Knowledge × Intervention). After generating the filled-in data sets, we fitted the growth curve models into each data set and used Rubin's (1987) rules to combine the parameter estimates and standard errors into a single set of results.

Measures

Kindergarten, Fall Term

Phonological awareness and letter knowledge. We used a letter knowledge test (α = .98) and a sum score of eight tests of phonological awareness (α = .96) at the beginning of the kindergarten year. Swedish versions of the same tests as in Lundberg et al. (1988) were used (for descriptions, see Kjeldsen et al., 2003).

Vocabulary. A 22-item picture naming test was used (Ege, 1985; for details, see Kjeldsen et al., 2003). Based on an item-analysis, five items were excluded and the vocabulary score is based on the remaining 17 items. Cronbach's alpha was .64

Grades 3, 4 and 6, Spring Term

Decoding. In an attempt to measure word attack skills economically and in the absence of contextual support, Jacobson's (2001) word chain test was used. In this group test, participants were told to draw a slash, as quickly and accurately as possible, between words written together without space. For Grade 3, the word chains consisted of three words varying from two to six graphemes in length. For Grades 4 and 6, the word chains consisted of four words varying from two to six letters in length. The words were nouns, verbs, adjectives, or numerals. The time limit was 2 min. The test manual gives the test–retest reliability .89 for Grade 3 and .84 for Grade 5 (no estimates are available for Grades 4 and 6).

For the purpose of model testing, the Grade 3 data had to be equated with those of Grades 4 and 6. This was done by means of linear equating procedure and a new sample of Grade 3 students randomized into two groups (with $n = 30$ and $n = 32$), for which the tests were administered in different orders (Crocker & Algina, 2008, pp. 459–460). Based on these data, the Grade 3 measurements were then transformed to Grade 4 equivalent measurements by $Y^* = 0.54 \times (X - 32.98) + 14.74$, where Y^* is the Grade 4 equivalent score and X is the original Grade 3 measurement.

Grade 9, Spring Term

Reading comprehension. A group test was used (Johansson, 1992). It consists of 10 short texts with gradually increasing length and difficulty. For each text there are two tasks: choosing out of five alternatives the best fitting heading and determining which ones (min = 2, max = 4) out of six statements are correct. Zero point is given if the participant chooses more than one heading for a text. Each correctly chosen statement gives 1 point and each incorrectly chosen one gives –1 point. Because the student can choose anything between zero and six possibilities as correct, this procedure is recommended by the author to prevent guessing from inflating the scores. The maximum score is 36. The few cases with a negative total score were converted to 0 because this score would be expected on the basis of pure guessing. According to the test manual, the Cronbach's alpha is .72.

Procedure

The training program was adopted from Lundberg et al. (1988; see also Adams, Foorman, Lundberg, & Beeler, 1998). A detailed description of the program can also be found in Kjeldsen et al. (2003). The training started in September 1997 and continued for 8 months. A 15- to 20-min session was arranged 5 days a week. The program was structured in a sequence consisting of six groups of tasks with increasing difficulty. In each group there were seven to 12 tasks. A typical training session consisted of five to nine short games or tasks, many of which were repetitions from previous sessions to support learning. In addition to the whole program, a small number of the children received a training dose corresponding to 60% of the program. As it turned out later, the smaller training dose exerted an effect on reading and spelling in Grades 1 and 2 similar to that of the whole program (Kjeldsen et al., 2003). To ensure treatment fidelity, kindergarten teachers in the training group were carefully instructed both in groups and individually in how to present the program. During the year preceding the study, the theoretical background was presented by in-service courses open to all teachers in Åland. Moreover, a pilot program with different students was implemented for the teachers of the training group in order to familiarize them with the program and the tests. During the intervention fidelity was assessed by questionnaires to kindergarten teachers who also wrote diaries about their work.

RESULTS

Descriptive Statistics

The descriptive statistics for the study variables are presented in Table 1. There were approximately equal proportions of boys and girls in the intervention and control groups. The control group scored higher in initial phonological awareness and letter knowledge, whereas the intervention group scored higher on the decoding and reading comprehension variables. According to the point-biserial correlations (with p values equaling to those from independent samples t tests; see Table 1, row 9 Intervention), there were no statistically significant differences between intervention and control groups on preschool measurements. However, compared to the control group, the intervention group performed marginally significantly better on decoding in Grades 3 and 6 and

TABLE 1
Descriptive Statistics and Correlations Between Study Variables

	Control		*Intervention*		*Correlations*							
Variables	M	SD	M	SD	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*
1. Boy	0.51	0.50	0.55	0.50	—							
2. PA PS	32.86	13.60	30.77	14.92	−0.29***	—						
3. LK PS	27.12	17.71	25.73	18.95	−0.32***	0.65***	—					
4. Vocabulary PS	33.96	3.80	33.94	3.77	0.01	0.42***	0.36***	—				
5. Decoding G3	14.69	6.01	16.22	5.68	−0.32***	0.41***	0.48***	0.16*	—			
6. Decoding G4	18.87	6.29	21.54	6.08	−0.29***	0.33***	0.36***	0.16*	0.73***	—		
7. Decoding G6	26.73	7.42	28.59	7.15	−0.25***	0.20**	0.29***	0.09	0.51***	0.54***	—	
8. RC G9	16.74	9.34	20.14	5.86	−0.16*	0.33***	0.27***	0.28***	0.44***	0.56***	0.37***	—
9. Intervention	—	—	—	—	0.04	−0.07	−0.04	0.00	0.13†	0.21**	0.13†	0.22**

Note. $N = 209$ (imputed data). PA = phonological awareness; LK = letter knowledge; RC = reading comprehension; PS = preschool; G = grade.
†$p < .10$. *$p < .05$. **$p < .01$. ***$p < .001$.

highly significantly better on decoding in Grade 4 and on reading comprehension in Grade 9. The skewness and kurtosis values indicated that there was some departure from normal distribution in the dependent variables. However, the values were clearly below levels that have been shown to be suspect in some methodological studies (univariate skewness of 2.0 and kurtosis of 7.0; e.g., Curran, West, & Finch, 1996).

Latent Curve Models

In latent curve models (LCMs; Bollen & Curran, 2006), it is assumed that all members of a population have developmental curves of the same functional form (e.g., linear or quadratic), but the parameters that describe those curves are allowed to differ. In other words, individuals may differ in their initial level (intercept) as well as in their growth rate (slope) over time. LCMs can be used to estimate (a) the parameters for an average initial level and an average rate of change and (b) between-subject variability in these growth parameters. The main task is often, however, to explain this between-subject variability in individual development (Willett & Sayer, 1994).

We fitted LCMs to data in order to study the intervention effects on the decoding trajectories extending across Grades 3, 4, and 6. In doing this, it was tested (a) whether the intervention has an indirect effect on reading comprehension in Grade 9 via influencing the level (intercept) or growth rate (slope) of decoding and (b) whether the intervention effects on decoding or reading comprehension depend on the initial levels of phonological awareness or letter knowledge measured in kindergarten (interaction). The intervention effects were examined controlling the effects of student gender and the level of vocabulary measured prior to the intervention.

Because a linear LCM did not fit the data, $\chi^2(9) = 20.175$, $p = .017$, a nonlinear trajectory was specified for the development of decoding by fixing the first loading on slope factor to 0, freely estimating the second loading, and fixing the third loading to 1 (Bollen & Curran, 2006, pp. 98–103; Figure 1). The reading comprehension and vocabulary variables were included in the models as single-indicator latent variables to take their measurement errors into account (e.g.,

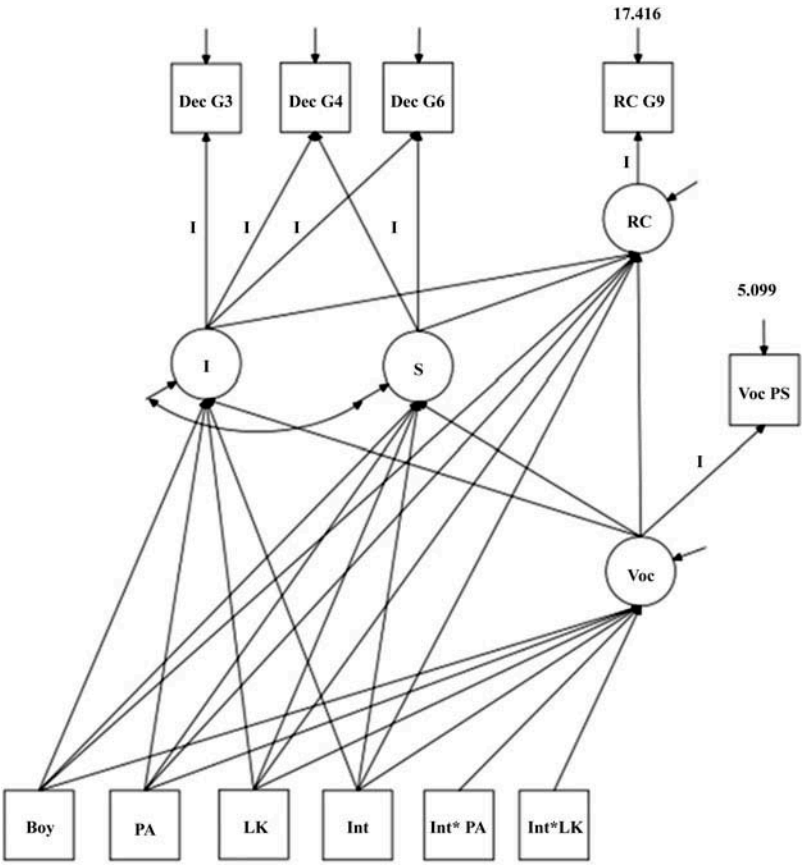

FIGURE 1 A latent curve model for testing intervention effects on decoding (Dec) in Grades 3, 4, and 6 and reading comprehension (RC) in Grade 9. *Note.* I = intercept; S = slope; Voc = vocabulary; PA = phonological awareness; LK = letter knowledge; Int = intervention. Numbers indicate the values to which parameters were fixed.

Kline, 2011). To stay in the regression-based framework of Mplus 7.11, we regressed our latent exogenous variable (Vocabulary) on the observed exogenous variables, which allowed the relationships between these variables to differ from zero. The growth factors and the Grade 9 reading comprehension were predicted by the variables boy, intervention, phonological awareness, letter knowledge, vocabulary, and interaction terms Intervention × Phonological Awareness as well as Intervention × Letter Knowledge. Furthermore, the Grade 9 reading comprehension was predicted by the growth factors as well. This model had a good fit to the data, $\chi^2(8) = 13.397$, $p = .099$. We performed a Wald test to examine whether the intervention effects on growth factors and reading comprehension were moderated by the levels of phonological awareness or letter knowledge. A simultaneous test of the six coefficients for the interaction terms Intervention × Phonological Awareness and Intervention × Letter Knowledge indicated that they did not differ

statistically significantly from zero, Wald test, $\chi^2(6) = 6.956$, $p = .325$. Fixing these coefficients of the interaction terms to zero resulted in a more parsimonious model, $\chi^2(14) = 22.333$, $p = .072$, and the results for this model are presented in Table 2.

According to the model results, the mean level of decoding for Grade 3 control group girls with average levels of phonological awareness, letter knowledge and vocabulary was 15.464 ($p < .001$). Boys as a group scored on the average 2 points lower than girls ($b = –1.992$, $p = .010$), and the intervention group students 2 points higher than control group students ($b = 2.129$, $p = .003$). Furthermore, both phonological awareness ($b = 0.078$, $p = .039$) and letter knowledge ($b = 0.100$, $p < .001$) predicted the level of decoding positively. The average increase in decoding skills between Grades 3 and 6 was 12.163 points ($p < .001$) of which 39.8% occurred between Grades 3 and 4 (cf. the loading of Grade 4 decoding on slope growth factor). Finally, the intercept (i.e., the level of decoding) had a direct effect on the Grade 9 reading comprehension ($b = 0.908$, $p = .003$).

We investigated the effect of intervention on Grade 9 reading comprehension further by examining the indirect effects. The total effect of intervention on Grade 9 reading comprehension was 3.692, 95% confidence interval (CI) [1.638, 5.746], $p < .001$. About 22% of this effect consisted of the (statistically nonsignificant) direct effect ($b = 0.811$), 95% CI [–5.071, 6.693], $p = .787$, whereas 78% of this total effect was indirect ($b = 2.881$), 95% CI [–2.938, 8.700], $p = .332$. The indirect intervention effect passed mainly through the intercept of decoding ($b = 1.932$), 95% CI [0.115, 3.749], $p = .037$; the effect through the slope was smaller ($b = 0.949$), 95% CI [–4.425, 6.323], $p = .729$, and statistically nonsignificant. Finally, we calculated standardized effect sizes for these indirect effects by dividing them by the control group standard deviation. This resulted in effect sizes of 0.31 for the total indirect effect and 0.21 and 0.10 for the effects via the intercept and the slope, respectively.

DISCUSSION

This longitudinal study followed up the reading development of 209 Swedish-speaking students from the beginning of kindergarten until the end of Grade 9. For a half of the participants, training in phonological awareness was given during the kindergarten year according to the guidelines set by Lundberg et al. (1988). In the subsequent follow-up, the potential associations between the intervention and reading progress through Grades 3 to 9 were delineated. The focus was on decoding through Grades 3 to 6 and reading comprehension in Grade 9.

There were no interactions between the intervention and the initial level of reading readiness, that is, letter knowledge and phonological awareness. Instead, the intervention was associated with a general difference in the level of decoding in Grade 3, which sustained until the end of Grade 6. This result corroborates the positive effect of intervention on decoding in Grades 1 and 2 with the same students (Kjeldsen et al., 2003). A notable detail is that the difference between trained and untrained students was in place already in Grade 1 and remained conspicuously stable through the elementary school. In other words, the continuous growth in decoding skills, albeit impressive, could not be predicted on the basis of the intervention. Although girls outperformed boys significantly, the growth in decoding ability was not related to gender. These findings are in line with ample evidence suggesting a stability of the students' relative position as a reader in the lower grades (Juel, 1988; Lundberg, 1994), until Grade 6 (de Jong & van der Leij, 2003),

TABLE 2
Latent Curve Model Results for the Intervention Effects on Word Decoding in Grades 3, 4, and 6 as Well as on Reading Comprehension in Grade 9

	Estimate	SE	p
Loadings			
Intercept[a]			
Decoding G3	1.000	—	—
Decoding G4	1.000	—	—
Decoding G6	1.000	—	—
Slope[a]			
Decoding G3	0.000	—	—
Decoding G4	0.398	0.027	<.001
Decoding G6	1.000	—	—
Vocabulary[a]			
Vocabulary PS	1.000	—	—
Reading comprehension[a]			
Reading comprehension G9	1.000	—	—
Structural coefficients			
Intercept			
Boy	−1.992	0.775	.010
Phonological awareness	0.078	0.038	.039
Letter knowledge	0.100	0.027	<.001
Vocabulary	−0.068	0.198	.732
Intervention	2.129	0.727	.003
Intervention × Phonological Awareness	0.000	—	—
Intervention × Letter Knowledge	0.000	—	—
Slope			
Boy	−0.885	1.038	.394
Phonological awareness	−0.067	0.052	.200
Letter knowledge	−0.033	0.036	.364
Vocabulary	0.203	0.278	.466
Intervention	0.620	0.959	.518
Intervention × Phonological Awareness	0.000	—	—
Intervention × Letter Knowledge	0.000	—	—
Vocabulary			
Boy	1.290	0.493	.009
Phonological awareness	0.065	0.033	.052
Letter knowledge	0.067	0.026	.009
Intervention	0.161	0.461	.727
Intervention × Phonological Awareness	0.041	0.043	.342
Intervention × Letter Knowledge	−0.046	0.033	.165
Reading comprehension			
Boy	1.160	3.590	.747
Phonological awareness	0.119	0.220	.590
Letter knowledge	−0.034	0.121	.778
Vocabulary	0.458	0.811	.572
Intervention	0.811	3.001	.787
Intervention × Phonological Awareness	0.000	0.000	—

(*Continued*)

TABLE 2
(Continued)

	Estimate	SE	p
Intervention × Letter Knowledge	0.000	0.000	—
Intercept	0.908	0.304	.003
Slope	1.499	2.963	.613
Intercepts			
Decoding G3	0.000	—	—
Decoding G4	0.000	—	—
Decoding G6	0.000	—	—
Intercept	15.464	0.629	<.001
Slope	12.163	0.847	<.001
Vocabulary	−0.756	0.419	.071
Reading comprehension	−15.142	38.380	.693
Residual variances			
Decoding G3	6.577	2.607	.012
Decoding G4	11.770	1.816	<.001
Decoding G6	30.890	5.943	<.001
Intercept	18.121	3.000	<.001
Slope	5.984	7.792	.443
Vocabulary PS	5.099	—	—
Vocabulary	5.929	1.085	<.001
Reading comprehension G9	17.416	—	—
Reading comprehension	11.421	21.003	.587
Residual covariances			
Intercept-Slope	−2.951	4.006	.461

Note. An em dash indicates that the parameter was fixed.
[a]Latent variables of the measurement model.

and even until upper grades (Landerl & Wimmer, 2008; Snowling, Muter, & Carroll, 2007). The significant gains in decoding produced by the intervention were rather modest in absolute terms. However, they do speak for an early intervention of phonological awareness instead of one given in school, which must also incorporate elements of reading. A comparison can be made with the study by Hurry and Sylva (2007) in which the training took place when children had already attended school for 1 year. Initial gains in reading had disappeared by the follow-up 3.5 years later. This finding is in line with that by Poskiparta, Niemi, and Vauras (1999; Niemi, Poskiparta, & Vauras, 2001), another study with gains disappearing 2.5 years after the termination of training given in Grade 1.

The present results also suggest a modest advantage in reading comprehension for the trained students as far as in Grade 9. The gain was approximately equally divided in direct and indirect contributions of the intervention. However, only the indirect effects through decoding development reached significance. To our knowledge, only one study has previously explored long-term gains in comprehension. Elbro and Klint Petersen (2004) found a nonsignificant tendency toward improved reading comprehension among seventh graders with a familial risk of dyslexia. The present results lend support to their finding by showing a somewhat stronger sustained gain in reading comprehension in an unselected sample consisting of middle school students 2 years older.

It is well established that orthographies differ in how easily they lend themselves to reading acquisition. A consistent pattern is that in opaque orthographies children need considerably more time to reach good accuracy and fluency. On the other hand, cross-linguistic comparisons also show that phonological awareness plays an important role in all orthographies (e.g., Caravolas, Lervåg, Defior, Seidlová Málková, & Hulme, 2013; Moll et al., 2014; Ziegler et al., 2010). Cross-language comparisons have focused on regular classroom reading progress, and to our knowledge there are no studies directly comparing effects of phonological training as a function of orthographic depth. Such a study would be needed to test whether phonological awareness is of equal importance regardless of orthographic depth.

The strengths of the present study are the unusually long follow-up period and low attrition rate. The results lend support to notion that training in phonological awareness can have long-lasting beneficial effects on reading that generalize to comprehension. The study also has obvious limitations. Only the reading milestones were measured across the years, whereas data are lacking concerning the formative processes taking place in classroom and outside during the 10-year-long follow-up. Although the development after training was beneficial in terms of subsequent decoding (Kjeldsen et al., 2003), and turning to a comprehension gain in Grade 9, the role of several confounding factors cannot be ruled out. For example, receiving special training prior to school in a small community may launch processes favoring learning in school rather than just promoting reading acquisition. The declining interest in many-year-long follow-up studies of phonological training is likely to reflect the strenuous nature of this type of a research as well as the rather modest outcomes and notable lack of information on what takes place between measurement points. There is need for more research on the details of the causal processes operating between the development of initial word-level decoding skill and effective reading comprehension. The present results suggest that an early phonological training is associated with a systematic reading gain through Grades 3 to 9. However, the gain is rather small in absolute terms. The implication is that it would be desirable to delineate activities building on the initial gain. One promising line of research that deserves continued attention is the work on print exposure and amount of reading (Pfost, Doerfler, & Artelt, 2013; Sparks et al., 2014). Other areas concern the connections between word identification processes, vocabulary learning, and reading comprehension (Perfetti & Stafura, 2014).

ACKNOWLEDGMENTS

We thank the students, teachers, and parents for their devoted participation in the study.

FUNDING

The first author received financial support from Government of Åland Islands.

REFERENCES

Adams, M., Foorman, B., Lundberg, I., & Beeler, T. (1998). *Phonemic awareness in young children*. Baltimore, MD: Brookes.

Anthony, J., & Lonigan, C. (2004). The nature of phonological awareness: Converging evidence from four studies of preschool and early grade school children. *Journal of Educational Psychology*, *96*, 43–55.

Anthony, J., Lonigan, C., Burgess, S., Driscoll, K., Phillips, B., & Cantor, B. (2002). Structure of preschool phonological sensitivity: Overlapping sensitivity to rhyme, words, syllables, and phonemes. *Journal of Experimental Child Psychology*, *82*, 65–92.

Bollen, K. A., & Curran, P. J. (2006). *Latent curve models: A structural equation perspective*. New York, NY: Wiley.

Bus, A., & van IJzendoorn, M. (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., & Ashley, L. (2000). Effects of preschool phoneme identity training after six years: Outcome level distinguished from rate of response. *Journal of Educational Psychology*, *92*, 659–667.

Caravolas, M., Lervåg, A., Defior, S., Seidlová Málková, G., & Hulme, C. (2013). Different patterns, but equivalent predictors, of growth in reading in consistent and inconsistent orthographies. *Psychological Science*, *24*, 1398–1407.

Carroll, J., Snowling, M., Hulme, C., & Stevenson, J. (2003). The development of phonological awareness in preschool children. *Developmental Psychology*, *39*, 913–923.

Crocker, L., & Algina, J. (2008). *Introduction to classical and modern test theory*. Mason, OH: Cengage.

Cunningham, A. E., & Stanovich, K. E. (1997). Early reading acquisition and its relation to reading experience and ability ten years later. *Developmental Psychology*, *33*, 934–945.

Curran, P. J., West, S. G., & Finch, J. F. (1996). The robustness of test statistics to nonnormality and specification error in confirmatory factor analysis. *Psychological Methods*, *1*, 16–29.

Cutting, L. E., & Scarborough, H. S. (2006). Prediction of reading comprehension: Relative contributions of word recognition, language proficiency, and other cognitive skills can depend on how comprehension is measured. *Scientific Studies of Reading*, *10*, 277–299.

de Jong, P., & van der Leij, A. (2003). Developmental changes in the manifestation of a phonological deficit in dyslexic children learning to read a regular orthography. *Journal of Educational Psychology*, *95*, 22–40.

Ege, B. (1985). *Sproglig Test 1* [Language Test 1]. Herning, Denmark: Specialpedagogisk Forlag.

Ehri, L., Nunes, D., Willows, D., Schuster, B., 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. Borstrøm, I., & Petersen, D. (1998). Predicting dyslexia from kindergarten: The importance of distinctness of phonological representations of lexical items. *Reading Research Quarterly, 33*, 36–60.

Elbro, C., & Jensen, M. (2005). Quality of phonological representations, verbal learning, and phoneme awareness in dyslexic and normal readers. *Scandinavian Journal of Psychology*, *46*, 375–384.

Elbro, C., & Klint Petersen, D. (2004). Long-term effects of phoneme awareness and letter sound training: An intervention study with children at risk for dyslexia. *Journal of Educational Psychology*, *96*, 660–670.

Gough, P., & Tunmer, W. (1986). Decoding, reading, and reading disability. *RASE: Remedial & Special Education*, *7*, 6–10.

Høien, T., Lundberg, I., Stanovich, K., & Bjaalid, I. (1995). Components of phonological awareness. *Reading and Writing: An Interdisciplinary Journal*, *7*, 171–188.

Hulme, C., Bowyer-Crane, C., Carroll, J., Duff, F., & Snowling, M. (2012). The causal role of phoneme awareness and letter-sound knowledge in learning to read: Combining intervention studies with mediation analyses. *Psychological Science*, *23*, 572–577.

Hurry, J., & Sylva, K. (2007). Long-term outcomes of early intervention. *Journal of Research in Reading*, *30*, 227–248.

Jacobson, C. (2001). *Läskedjor* [Reading chains]. Stockholm, Sweden: Psykologiförlaget AB.

Johansson, M-G. (1992). *Klassdiagnoser i läsning och skrivning för högstadiet och gymnasiet. Handledning* [Classroom diagnoses of reading and writing for junior and senior high school]. Stockholm, Sweden: Psykologiförlaget AB.

Juel, C. (1988). Learning to read and write: A longitudinal study of 54 children from first through fourth grades. *Journal of Educational Psychology*, *80*, 437–447.

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

Kline, R. B. (2011). *Principles and practice of structural equation modeling* (3rd ed.). New York, NY: Guilford.

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

Lonigan, C., Burgess, S., & Anthony, J. (2000). Development of emergent literacy and early reading skills in preschool children: Evidence from a latent-variable longitudinal study. *Developmental Psychology*, *36*, 596–613.

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

Lundberg, I., Frost, J., & Petersen, O-P. (1988). Effects of an extensive program for stimulating phonological awareness in preschool children. *Reading Research Quarterly*, *23*, 263–284.

Melby-Lervåg, M., Halaas Lyster, S-A., & Hulme, C. (2012). Phonological skills and their role in learning to read: A meta-analytic review. *Psychological Bulletin*, *138*, 322–352.

Moll, K., Ramus, F., Bartling, J., Bruder, J., Kunze, S., Neuhoff, N., . . . Schulz, E. (2014). Cognitive mechanisms underlying reading and spelling development in five European orthographies. *Learning and Instruction*, *29*, 65–77.

Muter, V., Hulme, C., Snowling, M., & Taylor, S. (1998). Segmentation, not rhyming, predicts early progress in learning to read. *Journal of Experimental Child Psychology*, *71*, 3–27.

National Institute of Child Health and Human Development Early Child Care Research Network. (2005). Pathways to reading: The role of language in the transition to reading. *Developmental Psychology*, *41*, 428–442.

Niemi, P., Poskiparta, E., & Vauras, M. (2001). Benefits of training in linguistic awareness dissipate by Grade 3? *Psychology. The Journal of Hellenic Psychological Society*, *8*, 330–337.

Olofsson, Å. (2003). The dyslexic reader and the Swedish language. In N. Goulandris (Ed.), *Dyslexia: Cross-linguistic comparisons* (pp. 137–156). London, UK: Whurr.

Papadopoulos, T., Kendeou, P., & Spanoudis, G. (2012). Investigating the factor structure and measurement invariance of phonological abilities in a sufficiently transparent language. *Journal of Educational Psychology*, *104*, 321–336.

Perfetti, C. (2007). Reading ability: Lexical quality to comprehension. *Scientific Studies of Reading*, *11*, 357–383.

Perfetti, C., & Stafura, J. (2014). Word knowledge in a theory of reading comprehension. *Scientific Studies of Reading*, *18*, 22–37.

Pfost, M., Doerfler, T., & Artelt, C. (2013). Students' extracurricular reading behavior and the development of vocabulary and reading comprehension. *Learning and Individual Differences*, *26*, 89–102.

Poskiparta, E., Niemi, P., & Vauras, M. (1999). Who benefits from training in linguistic awareness in the first grade, and what components show training effects? *Journal of Learning Disabilities*, *32*, 437–446, 456.

Puolakanaho, A., Poikkeus, A-M., Ahonen, T., Tolvanen, A., & Lyytinen, H. (2004). Emerging phonological awareness as a precursor of risk in children with and without familial risk for dyslexia. *Annals of Dyslexia*, *54*, 221–243.

Rubin, D. B. (1987). *Multiple imputation for nonresponse in surveys*. New York, NY: Wiley.

Schafer, J. L. (1997). *Analysis of incomplete multivariate data*. London, UK: Chapman & Hall.

Schneider, W., & Stengård, C. (2000). *Inventory of European longitudinal studies of reading and spelling*. Luxembourg: Office for Official Publications of the European Communities.

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

Silvén, M., Poskiparta, E., & Niemi, P. (2004). The odds of becoming a precocious reader of Finnish. *Journal of Educational Psychology*, *96*, 152–164.

Slavin, R., Lake, C., Davis, S., & Madden, N. (2011). Effective programs for struggling readers: A best-evidence synthesis. *Educational Research Review*, *6*, 1–26.

Snowling, M., & Hulme, C. (2011). Evidence-based interventions for reading and language difficulties: Creating a virtuous circle. *British Journal of Educational Psychology*, *81*, 1–23.

Snowling, M., Muter, V., & Carroll J. (2007). Children at family risk of dyslexia: A follow-up in early adolescence. *Journal of Child Psychology and Psychiatry*, *48*, 609–618.

Sparks, R., Patton, J., & Murdoch, A. (2014). Early reading success and its relationship to reading achievement and reading volume: replication of "10 years later." *Reading and Writing: An Interdisciplinary Journal*, *27*, 189–211.

Stahl, S., & Murray, B. (1994). Defining phonological awareness and its relationship to early reading. *Journal of Educational Psychology*, *86*, 221–234.

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

Storch, S. A., & Whitehurst, G. J. (2002). Oral language and code-related precursors to reading. Evidence from a longitudinal structural model. *Developmental Psychology*, *38*, 934–947.

Verhoeven, L., & van Leeuwe, J. (2009). Modeling the growth of word-decoding skills: Evidence from Dutch. *Scientific Studies of Reading*, *13*, 205–223.

Verhoeven, L., & van Leeuwe, J. (2011). Vocabulary growth and reading development across the elementary school years. *Scientific Studies of Reading*, *15*, 8–25.

Willett, J. B., & Sayer, A. G. (1994). Using covariance structure analysis to detect correlates and predictors of individual change over time. *Psychological Bulletin*, *116*, 363– 381.

Ziegler, J., Bertrand, D., Tóth, D., Csépe, V., Reis, A., Faísca, L., . . . Blomert, L. (2010). Orthographic depth and its impact on universal predictors of reading: A cross-language investigation. *Psychological Science*, *21*, 551–559.

APPENDIX

Missing Value Analyses

TABLE A1

Mean Comparisons for Students with Missing and Nonmissing Values: Intervention and Control Groups

	Phonological Awareness		*Letter Knowledge*		*Vocabulary*		*Decoding Grade 3*		*Decoding Grade 4*		*Decoding Grade 6*		*Reading Comprehension Grade 9*	
	Control	*Interv.*	*Control*	*Interv.*	*Control*	*Interv.*	*Control*	*Interv.*	*Control*	*Interv.*	*Control*	*Interv.*	*Control*	*Interv.*
Decoding Grade 3														
No. present	93	99	93	99	93	99	93	99	92	99	92	85	92	97
No. missing	8	9	8	9	8	9	0	0	1	0	2	0	2	0
% missing	7.9	8.3	7.9	8.3	7.9	8.3	0.0	0.0	1.1	0.0	2.1	0.0	2.1	0.0
M present	33.71	32.96	28.13	27.43	33.95	34.26	15.00	16.53	19.22	21.77	26.88	28.49	17.02	20.42
M missing	23.00	10.44	15.37	7.00	34.12	30.33	NA	NA	13.00	NA	25.00	NA	19.00	NA
Decoding Grade 4														
No. present	93	99	93	99	93	99	92	99	93	99	92	85	92	97
No. missing	8	9	8	9	8	9	1	0	0	0	2	0	2	0
% missing	7.9	8.3	7.9	8.3	7.9	8.3	1.1	0.0	0.0	0.0	2.1	0.0	2.1	0.0
M present	33.96	32.96	27.87	27.43	33.91	34.26	15.02	16.53	19.15	21.77	26.74	28.49	17.14	20.42
M missing	20.12	10.44	18.37	7.00	34.50	30.33	13.13	NA	NA	NA	31.50	NA	13.50	NA
Decoding Grade 6														
No. present	94	85	94	85	94	85	92	85	92	85	94	85	93	84
No. missing	7	23	7	23	7	23	1	14	1	14	0	0	1	13
% missing	6.9	21.3	6.9	21.3	6.9	21.3	1.1	14.1	1.1	14.1	0.0	0.0	1.1	13.4
M present	33.65	31.88	27.56	28.11	33.98	34.00	14.97	16.27	19.13	21.39	26.84	28.49	17.09	19.90
M missing	22.29	28.13	21.14	16.96	33.71	33.70	17.45	18.15	21.00	24.07	NA	NA	15.00	23.77
Reading comprehension Grade 9														
No. present	94	97	94	97	94	97	92	97	92	97	93	84	94	97
No. missing	7	11	7	11	7	11	1	2	1	2	1	1	0	0
% missing	6.9	10.2	6.9	10.2	6.9	10.2	1.1	2.0	1.1	2.0	1.1	1.2	0.0	0.0
M present	33.98	33.33	28.05	27.95	33.89	34.28	15.14	16.66	19.26	21.81	26.97	28.63	17.06	20.42
M missing	17.86	11.27	14.57	6.18	34.86	30.91	2.32	10.43	9.00	19.50	15.00	17.00	NA	NA

Note. Interv. = intervention; NA = not applicable.