Effects of training phonological awareness on children's reading skills*

Maximilian Pfost[a], Kristine Blatter[b], Cordula Artelt[c], Petra Stanat[d] & Wolfgang Schneider[e]

[a]University of Bamberg, Germany
[b]German Youth Institute, Germany
[c]Leibniz Institute for Educational Trajectories, University of Bamberg, Germany
[d]Institute for Educational Quality Improvement, Humboldt-Universität zu Berlin, Germany
[e]University of Würzburg, Germany

Abstract
Phonological awareness and letter knowledge have been shown to be precursors of children's word reading and reading comprehension. As a consequence, promoting children's phonological awareness should result in better reading skills. In order to evaluate this assumption, we trained phonological awareness and letter knowledge of 370 German preschool children and compared their word reading and reading comprehension skills with those of a group of 99 untrained children. Our findings indicate strong intervention effects on children's phonological awareness and letter knowledge. However, this advantage did not result in better reading skills in general. Within the group of low-performing children, we found small effects on word reading. This effect was mediated by individual differences in phonological awareness. In summary, only for low-performing children do the findings support our theoretical assumptions. This may be attributable to differences in the orthographic transparency between German and English, which restricts a generalization of findings across languages.

Keywords: Phonological awareness; Reading development; Emergent literacy skills;

Accepted Author's Manuscript. NOTICE: This is the author's version of a work that was accepted for publication. A definitive version was published 2019 in *Journal of Applied Developmental Psychology, 65*,101067. https://doi.org/10.1016/j.appdev.2019.101067

© 2019. This manuscript version is made available under the CC-BY-NC-ND 4.0 license
https://creativecommons.org/licenses/by-nc-nd/4.0/

*This study analyzes data from the research project "Short- and long-term effects of training phonological awareness in native and non-native German-speaking kindergarten children", which was supported by the German Federal Ministry of Education and Research (01GJ 0972-74).

Introduction

The idea that learning to read is a long-lasting and multifaceted process which starts in early childhood has found broad scientific acceptance as well as widespread reflection in psychological theories and models of reading development (Chall, 1983; Goswami & Bryant, 1990; Rieben & Perfetti, 1991; Stanovich, 1986; Whitehurst & Lonigan, 1998). From a perspective of individual differences in reading development, this idea also encompasses the underlying concept of a continuous and cumulative development: Early differences in emergent literacy skills (code-related precursors and oral language skills) are inherently linked to later differences in reading (Pfost, 2016; Scarborough, 2002; Stanovich, 1986). As learning to read within an alphabetic script requires an understanding of the alphabetic principle – the idea that printed symbols or graphemes systematically represent phonemes – beginning readers need to acquire sufficient knowledge about the phonological structure of their oral language in order to make an efficient use of this alphabetic principle. Therefore, phonological processing skills such as verbal short-term memory (the ability to store verbal information temporarily) and phonological awareness are assumed to be highly important for reading development (Melby-Lervåg, Lyster, & Hulme, 2012; Wagner & Torgesen, 1987).

Phonological awareness denotes the conscious access to phonemes, syllables as well as intra-syllabic units of oral language. In addition, phonological awareness encompasses the ability to cognitively manipulate these sound representations as reflected in tasks such as blending sounds together or reversing the order of sounds in a word (Anthony & Francis, 2005; Pfost, 2015; Stanovich, 1986; Wagner & Torgesen, 1987). From a developmental perspective, the general trend for phonological awareness goes from relatively large sound units (syllables, onset and rimes) – often acquired before learning to read – to smaller sound units (phonemes; Anthony & Francis, 2005; Goswami, 2002; Goswami & Bryant, 1990; see also Fricke, Szcerbinski, Stackhouse, & Fox-Boyer, 2008, for data on German-speaking children). Regarding the causal role of phonological awareness for students' word reading skills, some proponents emphasize that phonological awareness affects the development of students' reading skills (Bradley & Bryant, 1983; Cardoso-Martins, 1995; Hulme, Bowyer-Crane, Carroll, Duff, & Snowling, 2012; Muter, Hulme, Snowling, & Stevenson, 2004), whereas others underline the primacy of reading processes, which in turn promote children's phonological awareness (Castles & Coltheart, 2004; Morais, Cary, Alegria, & Bertelson, 1979), or assume a bidirectional relationship between phonological awareness and reading (Goswami & Bryant, 1990). Despite and somehow independent of this intense discussion, there is nevertheless strong evidence that early differences in children's phonological awareness are correlated to later differences in students' reading skills (see meta-analyses by Melby-Lervåg et al., 2012; National Early Literacy Panel, 2008; Pfost, 2015; Scarborough, 1998).

From a theoretical point of view, the development of children's reading comprehension has often been described as two strands of code-related /word reading skills and oral language skills that become increasingly more interwoven (Scarborough, 2002; Storch & Whitehurst, 2002). In order to understand text, it is thus necessary that beyond code-related skills, including phonological awareness and letter knowledge, children acquire language competences that contribute to the process of comprehension of words and sentences, for example vocabulary knowledge and grammar skills (Dickinson, Golinkoff, & Hirsh-Pasek, 2010). Furthermore, and as stressed by the comprehensive language approach (e.g. Dickinson, McCabe, Anastasopoulos, Peisner-Feinberg, & Poe, 2003; Rohde, 2015), these competences tend to be interdependent. For example, phonological awareness may itself develop as a consequence of children's vocabulary acquisition due to lexical restructuring processes. That is, due to vocabulary growth, children are increasingly able to differentiate between words with similar sounds, which should lead to a more segmented (phonemic) lexical representation (Goswami, 2002; Metsala, 1999).

Intervention studies on phonological awareness

Based on the knowledge that phonological awareness predicts later reading skills and is necessary for a flexible and accurate phonemic decoding ability as well as the formation of the orthographic lexicon (Phillips & Torgesen, 2006; Share, 1995), several intervention programs that foster children's phonological awareness skills were devised in order to prevent children from developing later reading difficulties. Among the first were Lundberg, Frost, and Petersen (1988) who developed an 8-month program which trains phonological awareness through daily sessions consisting of metalinguistic games, encompassing exercises on the syllabic as well as the phonemic level. Evaluation results were quite promising as participants from the intervention group showed higher levels of phonological awareness after the intervention and outperformed children from the control group in a word reading test in grades 1 and 2. These promising results were replicated and confirmed, as observed in several meta-analyses (Bus & van Ijzendoorn, 1999; Ehri et al., 2001; National Early Literacy Panel, 2008; Suggate, 2016).

In summary, these studies showed large intervention effects for the development of children's phonological awareness and small to medium intervention effects for children's reading skills. Furthermore, effect sizes for reading skills tended to be larger for preschoolers and kindergarten children in comparison to older children. Effect sizes were also larger when the training did not exclusively focus on phonological awareness skills but also involved the use of letters/alphabet knowledge (Bus & van Ijzendoorn, 1999; Ehri et al., 2001). An additional teaching of letters is based on the theoretical argumentation that letter knowledge itself promotes children's phonological awareness and changes children's representation of phonemes which contributes to children's reading development (Barron, 1991; Foulin, 2005).

From the perspective of causal effects, intervention effects on children's reading skills should be attributed to changes in phonological awareness and alphabet knowledge. Testing this mediation hypothesis by relying on data from 152 English-speaking children, Hulme et al. (2012) found strong empirical support for this theoretical assumption: 20 weeks of intervention on phonological awareness and letter knowledge significantly affected children's phoneme awareness and letter knowledge. Children's phoneme awareness and letter knowledge predicted children's literacy skills. Beyond the training effect via these two mediators, no direct training effect was found for children's literacy skills. In sum, the empirical findings fitted well to the theoretical assumptions.

Phonological awareness and learning to read in different alphabetic orthographies

Beyond the question of causality and intervention effects, a third variable, namely, orthographic consistency, which varies across different languages, has attracted considerable attention in recent research on phonological awareness and learning to read.

First, prior research has accumulated clear evidence that the acquisition of basic reading skills in English, due to its comparatively low level of consistency in grapheme-phoneme as well as phoneme-grapheme correspondence, generally tends to be more difficult and therefore takes longer in comparison to reading acquisition in more transparent orthographies (Seymour, Aro, & Erskine, 2003; Ziegler & Goswami, 2006). Second, due to these orthographic differences, beginning readers in different languages face different challenges and adjust to these challenges by relying on different cognitive strategies and mechanisms (Aro, 2006; Ziegler & Goswami, 2005). Within more transparent orthographies, we may assume that phonological awareness skills are less important for learning to read. One explanation is based on the fact that children reach a high level of word reading accuracy earlier on compared to children who are learning to read in a less transparent orthography. Therefore, the impact of individual differences in phonological awareness on reading development in more transparent orthographies might be less enduring (Aro, 2006; Landerl & Thaler, 2006; Pfost, 2015).

This idea is further supported by empirical findings that show lower correlations between phonological awareness and reading in more transparent orthographies (Mann & Wimmer, 2002; Ziegler et al., 2010). In the same vein, training phonological awareness seems more efficient for English students

in comparison to students from other (European) countries (Bus & van Ijzendoorn, 1999; Ehri et al., 2001). Lastly, meta-analyses on phonological awareness intervention studies in Germany – a rather transparent orthography as far as reading is concerned – reported small effect sizes, partially not reaching statistical significance, for children's decoding skills and reading comprehension (Fischer & Pfost, 2015; Wolf, Schroeders, & Kriegbaum, 2016).

However, the assumption that phonological awareness is less important for the reading development of German students may only apply to children with average or above-average emergent literacy and reading skills. Contrary to average and above average performing children, below average performing students might face comparable difficulties in phonological recoding as do students from English-speaking countries due to their lower word reading accuracy. Therefore, for children performing below average who learn to read in German, phonological awareness may be as important as it is for English-speaking children. Finally, we should keep in mind that phonological awareness itself belongs to the family of constrained reading skills (Paris, 2005): Normal readers acquire phonological awareness skills relatively quickly and entirely master phonological awareness tasks relatively early while learning to read (e.g. Fricke et al., 2008, for data on German students). As effects of phonological awareness on reading might only be found as long as individual differences in phonological awareness exist, German-speaking children without deficits in emergent literacy skills might be less affected by a training of such skills due to a high mastery level within this group of children (see also Kjeldsen, Educ, Saarento-Zaprudin, & Niemi, 2019, who report stronger intervention effects for Finish Swedish at-rist students in comparison to not-at-risk students).

Aims of the study

The current paper analyzes data from a longitudinal intervention study of children learning to read within the German orthography. Preschoolers in the intervention group participated in the evaluated and well-documented combined phonological awareness and letter knowledge training "Hören, lauschen, lernen" I and II (HLL; Küspert & Schneider, 2008; Plume & Schneider, 2004). Comparable to the training program by Lundberg et al. (1988), HLL aims at promoting preschoolers' phonological awareness through daily sessions, approximately 15 to 20 min in length, of metalinguistic games and exercises. In the first week HLL starts with listening games and rhymes and continues with words and phrases at the beginning of the third week. Syllables are introduced in week five and metalinguistic exercises on the sub-syllabic level are introduced in week seven. Finally, exercises on the phoneme level start in week eleven, as does the letter knowledge training. Within this part of the training, twelve letters and their corresponding sounds are additionally introduced, with a focus on letter-sound correspondences. As HLL is designed for preschool children, the program explicitly does not intend to teach children to read or write words and letters. However, children should become familiar with letters and learn that these letters are systematically associated with specific phonemes. Games use pictures to identify word onset and to then assign it to a specific letter, and also include motor exercises to represent letter forms (see Plume & Schneider, 2004, for further exercises). The training period is intended to last a minimum of 20 weeks. Intervention effects of the HLL training program were evaluated several times by different authors, which makes HLL the most evaluated phonological awareness intervention program in Germany (Fischer & Pfost, 2015; Wolf et al., 2016). Across all evaluation studies, participants showed significant improvements in phonological awareness with a medium effect size. Small effects were reported for reading, but these effects tend to vanish after a post-test interval of more than one year.

In addition to limiting variables, such as small sample size, prior intervention studies neglected to analyze the possible (causal) mediating mechanism of the intervention effects. Therefore, comparable to the study by Hulme et al. (2012), we intended to analyze the effects of a phonological awareness and letter knowledge intervention in preschool on first graders' reading skills by taking the theoretically implied causal mechanisms of this intervention effect into account. Contrary to the study by Hulme et al. (2012), who exclusively focused on phoneme awareness and letter knowledge as potential mediating

mechanisms, we also included oral language skills, such as vocabulary and grammar skills, as an alternative explanatory model. Therefore, our analyses acknowledge the criticism that oral language skills have been underrepresented in previous studies that explored key determinants of individual differences in reading development (e.g., Dickinson et al., 2010). Accordingly, the conceptual representation of reading development favored in our study is conceived of as a process of interwoven strands of language comprehension and code-related skills (Scarborough, 2002; Storch & Whitehurst, 2002). A consideration of alternative explanatory variables allowed for the testing of the specificity of the training effects on reading via phonological awareness and letter knowledge. Based on prior research, we came up with the following expectations:

First, preschoolers who participated in the combined phonological awareness and letter knowledge intervention program should show a higher level of phonological awareness and letter knowledge in comparison to non-participants. Due to the specificity of the intervention, we expected that the direct effects of the intervention would be limited to the trained skills. Therefore, we did not expect to find direct training effects on oral language skills, such as vocabulary and grammar skills. In addition, we did not expect to find direct training effects on untrained code-related skills, such as verbal short-term memory.

Second, with regard to the study by Hulme et al. (2012) as well as in line with the aforementioned causal perspective of phonological awareness and letter knowledge affecting reading development, we expected to find higher reading skills in grade 1 for those students who had participated in the phonological awareness intervention.

Third, we expected that the intervention effect on reading skills in grade 1 would be mediated by individual differences in phonological awareness and letter knowledge. Oral language skills and verbal short-term memory are expected to predict individual differences in reading as well. However, these skills are not expected to mediate any intervention effect, as these skills should not be directly affected by the intervention. Since reading comprehension is strongly dependent on decoding skills in early school years (Shankweiler et al., 1999; Storch & Whitehurst, 2002), we did not have any clear-cut expectations on differences in the mediation model between different reading measures in grade 1, i.e. word reading and reading comprehension.

Methods

Participants

At the beginning of the research project in 2009, 45 kindergartens located in the regions of Bavaria, Baden-Württemberg and Berlin, Germany, agreed to participate. Within these 45 kindergartens, 572 children took part in the study. In Germany, kindergarten refers to the public or private early childhood education and care settings that children enter before primary school. German primary schools normally start around age six. Kindergarten attendance is not obligatory. The participation of the children was voluntary, and we tested children only if written parental consent was provided. At the last measurement point, in grade 1, 469 children (82.0%) were still participating. Student drop-out mainly occurred between the post-test and follow-up assessments due to the transition into primary school, given that some primary schools refused to participate in the study and that some children did not enter school but stayed in kindergarten for another year. Within the final sample, 370 children were assigned to the intervention group, and 99 children were assigned to the control group. As some institutions already offered the HLL phonological awareness intervention program in their curriculum (and in consequence were automatically assigned to the intervention group), a completely randomized assignment of the children to the intervention or control group was not possible. Therefore, we compared children from the intervention and the control group on pre-test measures and corrected for sample imbalances (see below).

Within the analyzed sample, 50.5% of the children were male and 49.5% were female. At the beginning of the study, the children were on average 5 years and 7 months old. About 54.6% of the

children lived in families without a migration background, whereas 11.1% of the children lived in families with one parent born outside of Germany and further 30.1% of the children lived in families with both parents born outside of Germany. For 4.3% of the children information on their migration background was not available. Due to the sampling procedure, it was intended that children with a migration background were overrepresented (Blatter, 2015).

Design

After the pre-test (t1, Sep.–Nov. 2009), children received 20 weeks of intervention (Dec. 2009–May 2010) – the combined phonological awareness and letter knowledge training HLL I and II (Küspert & Schneider, 2008; Plume & Schneider, 2004). A kindergarten teacher conducted the intervention in everyday sessions of 15 to 20 min within small groups (≤ 10 children). Kindergarten teachers were trained within their institutions in order to implement the intervention within their regular class. Using training protocols, implementation quality was continuously monitored by staff members of the research group, ensuring that the training was implemented as intended. Although not all kindergarten teachers had received intensive training before and supervision during implementation of the intervention, first results showed negligible training differences between these groups (Schneider, Artelt, & Stanat, 2018). The post-test (t2) was assessed when children were still attending kindergarten (May–July 2010). Children's reading skills were tested at the second follow-up assessment in grade 1 (t4, May-July 2011; see Blatter et al., 2013). We did not make use of data from the first follow-up (t3) – a computer-based assessment – due to conceptual concerns about the measures. Only a few children were able to read when entering school, and the phonological awareness measures included in the computer-based assessment turned out to be very easy for most children in the sample.

Measures

In order to assess children's emergent literacy skills, tasks that encompass children's phonological awareness, verbal short-term memory and letter knowledge were used (code-related precursors). Children's oral language skills were assessed using a test of receptive vocabulary as well as tests of productive and receptive grammar skills. Finally, children's reading skills were evaluated with a word reading as well as a reading comprehension test. With the exception of the word reading and reading comprehension test, all estimates of internal consistency and test-retest reliability reported below are based on data of our own sample.

Phonological awareness (t1 and t2). Children's phonological awareness was assessed with six subtests encompassing different sizes of sound units that demand recognition and manipulation. In order to test children's *syllable and rhyme awareness*, two tasks from the "Bielefelder Screening zur Früherkennung von Lese- und Rechtschreibschwierigkeiten" (BISC; Jansen, Mannhaupt, Marx, & Skowronek, 2002) were applied. Within these tasks, children had to (1) identify rhymes, and (2) separate words into syllables. Every task was comprised of 10 items. Therefore, children could reach a maximum score of 20 points. Internal consistency of the scale was 0.77 at t1 and 0.72 at t2.

Children's *phonemic awareness* was assessed with (1) a sound-to-word comparison task that required children to compare a given sound to the onset of a specific word; (2) a phoneme synthesis task. That is, children were slowly presented words, phoneme by phoneme, which they had to recognize; (3) a phoneme deletion task, which required children at first to isolate the first sound of a word, and, secondly, to pronounce the remaining part of the word; and (4) a phoneme analysis task, which required the child to divide simple words into its constituent sounds. Correct answers were added together, resulting in a maximum score of 42 points. Internal consistency of the scale was 0.91 at t1 and 0.94 at t2.

Verbal short-term memory (t1 and t2). Children listened to a series of ten four-to-six syllable non-words of increasing difficulty from the BISC (Jansen et al., 2002). The non-words were pronounced

like real words and, after presentation, the child was asked to correctly reproduce the non-words. Internal consistency of the scale was 0.61 at t1 and 0.65 at t2.

Letter knowledge (t1 and t2). Children's letter knowledge was assessed using a letter naming task developed by Moser and Berweger (2007). Each child was asked to name 18 uppercase and 10 lowercase letters printed in random order on five pages. Internal consistency of the scale was 0.95 at t1 and t2.

Vocabulary (t1 and t2). In order to assess children's receptive vocabulary, an adapted German version of the "Peabody Picture Vocabulary Test – Revised" (PPVT-R; Roßbach, Tietze, & Weinert, 2005) was used. Internal consistency was not calculated as the test was terminated when children gave six or more false responses in a set of twelve words. Within the control group, test-retest reliability was 0.80 (the analysis was based solely on the control group in order to get an estimate for test-retest reliability not affected by intervention effects).

Productive and receptive grammar skills (t1 and t2). Two tests were used in order to assess children's grammar skills. *Productive grammar skills* were assessed using the plural-singular-formation subtest of the "Heidelberger Sprachentwicklungstest" (HSET; Grimm & Schöler, 1998). Within this test, children had to change words and non-words from singular into plural and from plural into singular by applying regular grammar rules of singular-plural formation. Within the control group, test-retest reliability was 0.81. Children's *receptive grammar skills* were assessed with 36 items from an adapted German version of the "Test for Reception of Grammar" (TROG-D; Fox, 2006). On each test item, the child had to listen to an orally provided sentence and to subsequently choose the illustration that correctly represents the heard content from a list of four pictures. In order to minimize effects of children's lexical knowledge, the test was designed to exclusively make use of easy words. Solving the test items correctly was possible without the use of any expressive language skills. Within the control group, test-retest reliability was 0.83.

Word reading (t4). We used a five-minute word reading test of the "Würzburger Leise Leseprobe – Revision" (WLLP-R; Schneider, Blanke, Faust, & Küspert, 2011). This test requires that children read single words and select the corresponding picture that illustrates the word's meaning from four possible alternatives. The children were instructed to read as many words as possible within five minutes. Children's word reading competence was indicated by the number of correctly identified words/pictures, with a maximum score of 140 points. According to the test manual, test-retest reliability was 0.76.

Reading comprehension (t4). Reading comprehension was measured with two subtests from the "Ein Leseverständnistest für Erst- bis Sechstklässler" (ELFE 1-6; Lenhard & Schneider, 2005). For the word reading comprehension subtest, the children had to choose the word that corresponds best to a given picture from a list of four words. For the sentence reading comprehension subtest, children had to choose the word that complements a given phrase from a list of five words (e.g. "You can [stand/see/walk/steal/lay] with your eyes, and listen with your ears."). Correct answers of both subtests were added together, resulting in a maximum score of 100 points. According to the test manual, test-retest reliability was 0.89 for word reading and 0.87 for sentence reading. With regard to validity, the test authors provided evidence of a strong correlation between these two subtests and teachers' judgements about children's reading skills (r = 0.53/0.58).

Procedure

At first, we analyzed whether children who participated in the phonological awareness intervention differed from children of the control group on our pre-test measures. As a completely randomized assignment of the children to the different research conditions was not feasible, we decided to control for imbalances respective differences on our covariates. In order to get a satisfying estimate of the treatment effect, we applied an inverse probability-of-treatment weighting procedure (Austin & Stuart, 2015; Leite, 2017). The simple idea behind this procedure is the use of sampling weights, which adjust for bias on covariates between children in the intervention and control group. Therefore, the weighted sample represents a pseudo-population in which the distributions of the measured covariates of the

intervention and control group are similar. All weights were stabilized as we multiplied the propensity score weights with the expectancy values of being in the intervention or control group within our sample. Then we evaluated covariate balance. In the next step, we analyzed intervention effects on post-test and follow-up outcomes via regression models using M*plus* (Muthén & Muthén, 1998-2012). We corrected standard errors of the estimates for non-normality using the MLR-estimator as well as the nesting of children within kindergartens by using the type = complex option.

Finally, we estimated a mediation model by using post-test outcomes (code-related precursors and oral language skills) as potential mediators for the follow-up reading measures. Intervention effects were split up into several specific indirect effects via the considered mediators, which add up to a total indirect effect, as well as a possible remaining unexplained or direct intervention effect. All dependent variables were standardized. As a consequence, intervention effects may be directly interpreted as effect size/standardized mean difference. Relations between post-test mediators and follow-up outcomes may be directly interpreted as correlations. Missing data were treated via multiple imputation throughout all analyses (see Graham, 2009). Based on a procedure described by van Buuren and Groothuis-Oudshoorn (2011), five imputed datasets were estimated. For missing data imputation, a broad predictor matrix, including, for example, data on children's sociodemographic background, skills and competences across all points of measurement as well as child and family literacy activities was applied.

Results

Descriptive statistics

Due to characteristics of the study design, children were not randomly assigned to the intervention and control group. Therefore, it was necessary to check for pre-test differences and, if pre-test differences did exist, to control for such differences. As can be seen from columns three to six of Table 1, comparing pre-test differences between children from the intervention and control group, we did not find significant group differences for children's age and sex, socio-economic background (HISEI-score), children's code-related precursors or children's productive grammar skills. However, significant differences for parents' migration and education background were present: children from the intervention group less often had parents with a migration background. Furthermore, these parents more often had completed upper track schooling. With regard to children's oral language skills, children from the intervention group slightly outperformed children from the control group on the vocabulary as well as receptive grammar skills measure. Therefore, in order to control for these differences, we applied a balancing procedure based on stabilized propensity score weights. For the estimation of the propensity score weights, all covariates presented in Table 1 (sociodemographic background, children's pre-test (t1) code-related precursor and oral language skills) were considered. Weighted results are shown in columns seven to ten of Table 1. Weighted results no longer indicated any group differences between children from the intervention and control group on all pre-test measures.

Intervention effects – children's code-related precursors and oral language skills

Intervention effects on children's code-related precursors and oral language skills are presented in Table 2. At the end of kindergarten, the trained children on average had higher phonological awareness skills (phoneme level) and higher letter knowledge than the control group. Furthermore, we observed a significant intervention effect for verbal short-term memory. Effects can be classified as moderate (letter knowledge, verbal short-term memory) to large (phoneme awareness; see Table 2, last column). As expected, we did not find any intervention effects for children's oral language skills – vocabulary and grammar skills.

Table 1

Weighted and unweighted pre-test differences between preschool children of the intervention (IG) and control group

		Unweighted Results					Weighted Results			
	Scale Range[a]	Total *M* (*SD*)	IG *n* = 370 *M*	Control *n* = 99 *M*	*Δ* (IG-Control) *M* (*SE*)	*d*	IG *w* = 369.4 *M*	Control *w* = 96.3 *M*	*Δ* (IG-Control) *M* (*SE*)	*d*
Sociodemographic background										
Sex (1= female)	0-1	0.49 (0.50)	0.49	0.51	-0.01 (0.06)	-0.03	0.49	0.49	0.01 (0.06)	0.01
Age (in months)	57-78	66.63 (4.01)	66.74	66.21	0.54 (0.48)	0.13	66.63	66.63	0.00 (0.50)	0.00
HISEI	16-90	49.26 (18.21)	50.06	46.23	3.83 (2.09)	0.21	49.24	49.32	-0.08 (2.26)	-0.00
Migration – one parent	0-1	0.12 (0.32)	0.11	0.15	-0.05 (0.04)	-0.14	0.12	0.11	0.01 (0.03)	0.02
Migration – both parents	0-1	0.32 (0.47)	0.29	0.42	-0.13 (0.06)*	-0.28	0.32	0.32	0.00 (0.05)	0.01
Parental education – upper track	0-1	0.48 (0.50)	0.52	0.33	0.19 (0.06)**	0.39	0.48	0.45	0.03 (0.07)	0.05
Parental education – middle track	0-1	0.31 (0.46)	0.29	0.37	-0.08 (0.06)	-0.16	0.31	0.30	0.00 (0.06)	0.01
Code-related precursors										
Verbal short-term memory	0-10	6.89 (2.07)	6.92	6.77	0.14 (0.26)	0.07	6.89	6.99	-0.09 (0.25)	-0.05
Phonological awareness – syllables and rhymes	0-20	15.73 (3.46)	15.84	15.36	0.48 (0.42)	0.14	15.74	15.68	0.07 (0.42)	0.02
Phonological awareness – phonemes	0-42	19.76 (8.01)	20.09	18.52	1.57 (0.91)	0.20	19.75	19.61	0.14 (1.01)	0.02
Letter knowledge	0-28	9.03 (7.75)	9.25	8.19	1.07 (0.86)	0.14	9.06	8.93	0.13 (0.99)	0.02
Oral language skills										
Vocabulary	0-175	42.68 (23.56)	44.10	37.39	6.71 (2.65)*	0.29	42.63	42.31	0.32 (3.06)	0.01
Productive grammar skills	0-36	15.59 (8.37)	15.67	15.31	0.36 (1.04)	0.04	15.59	16.18	-0.59 (1.05)	-0.07
Receptive grammar skills	0-36	27.10 (6.46)	27.58	25.32	2.26 (0.79)**	0.35	27.10	27.00	0.10 (0.76)	0.02

Note. All results are based on multiple imputed datasets ($m = 5$); Estimator = MLR. IG = Intervention group. Weighted results are based on stabilized propensity score weights (see methods section; Austin & Stuart, 2015; Leite, 2017); Propensity scores weights consider all variables presented in Table 1. [a]Except for age, the theoretical scale range is given.

* $p < .05$. ** $p < .01$.

Table 2

Intervention effects on preschool code-related precursors and oral language as well as first grade reading skills (Weighted results)

Dependent variable	ICC	IG *M (SD)*	Control *M (SD)*	*Δ* (IG-Control) *M* (*SE*)	Semi-standardized effect *B* (*SE*)
Code-related precursors (t2)					
Verbal short-term memory	0.110	7.47 (1.99)	6.70 (2.13)	0.77 (0.37)*	0.377 (0.183)*
Phonological awareness – syllables and rhymes	0.196	17.64 (2.39)	16.56 (3.11)	1.08 (0.75)	0.416 (0.288)
Phonological awareness – phonemes	0.336	29.52 (7.94)	22.43 (9.73)	7.08 (2.19)**	0.803 (0.248)**
Letter knowledge	0.113	15.97 (7.60)	12.30 (8.84)	3.67 (1.39)**	0.459 (0.173)**
Oral language skills (t2)					
Vocabulary	0.387	55.82 (23.46)	54.36 (25.33)	1.46 (7.54)	0.061 (0.316)
Productive grammar skills	0.197	18.88 (8.56)	18.90 (9.34)	-0.03 (2.29)	-0.003 (0.263)
Receptive grammar skills	0.310	29.70 (5.14)	29.35 (5.68)	0.35 (1.44)	0.067 (0.275)
Reading skills (t4)					
Word reading	0.255	38.32 (18.02)	35.19 (21.85)	3.13 (5.50)	0.165 (0.291)
Reading comprehension	0.240	21.34 (12.58)	19.74 (14.21)	1.60 (3.53)	0.124 (0.273)

Note. All results are based on multiple imputed datasets ($m = 5$); Estimator = MLR; Type = Complex. Results are weighted by using stabilized propensity score weights. ICC = Intraclass correlation, children nested within kindergartens; IG = Intervention group.

* $p < .05$. ** $p < .01$.

Intervention effects – children's reading skills

In grade 1, one year after the intervention program, we did not find any significant differences between children from the intervention and control group. Also with regard to descriptive results, the effect sizes were quite small. In sum, our results show that the participation in a phonological awareness intervention in kindergarten did not lead to better word-reading or reading comprehension skills one year after the intervention.

Models with direct and indirect intervention effects

In order to test the validity of the assumed theoretical model of the relation between children's code-related precursors, oral language and reading skills, we combined the aforementioned variables into a single (mediation) model. Although estimated within one model, the results are presented separately for word reading (Figure 1) and reading comprehension (Figure 2) in order to keep the figures clearly arranged. As all endogenous variables (code-related precursors, oral language and reading skills) were standardized, we may interpret effects between children from the intervention and control group as standardized mean differences. Relations between code-related precursors, oral language and reading skills may be directly interpreted as correlation coefficients with a scale range from -1 to 1. As expected, individual differences in phoneme awareness and letter knowledge at the end of kindergarten significantly predicted individual differences in grade 1 word reading and reading comprehension. The size of the correlations was moderate. The awareness of syllables and rhymes correlated with word reading as well, but the effect size was quite small. With regard to the oral language measures, children's vocabulary significantly predicted word reading and reading comprehension in grade 1. These correlations reached small to moderate effect sizes. Interestingly, although about half of the variance of the word reading and reading comprehension measures was explained by all variables within our model, we neither found a significant direct ($d_{word\ reading}$ = -0.215, $d_{reading\ comprehension}$ = -0.245), indirect ($d_{word\ reading}$ = 0.380; $d_{reading\ comprehension}$ = 0.369), or total ($d_{word\ reading}$ = 0.165; $d_{reading\ comprehension}$ = 0.124) intervention effect for the reading outcomes.

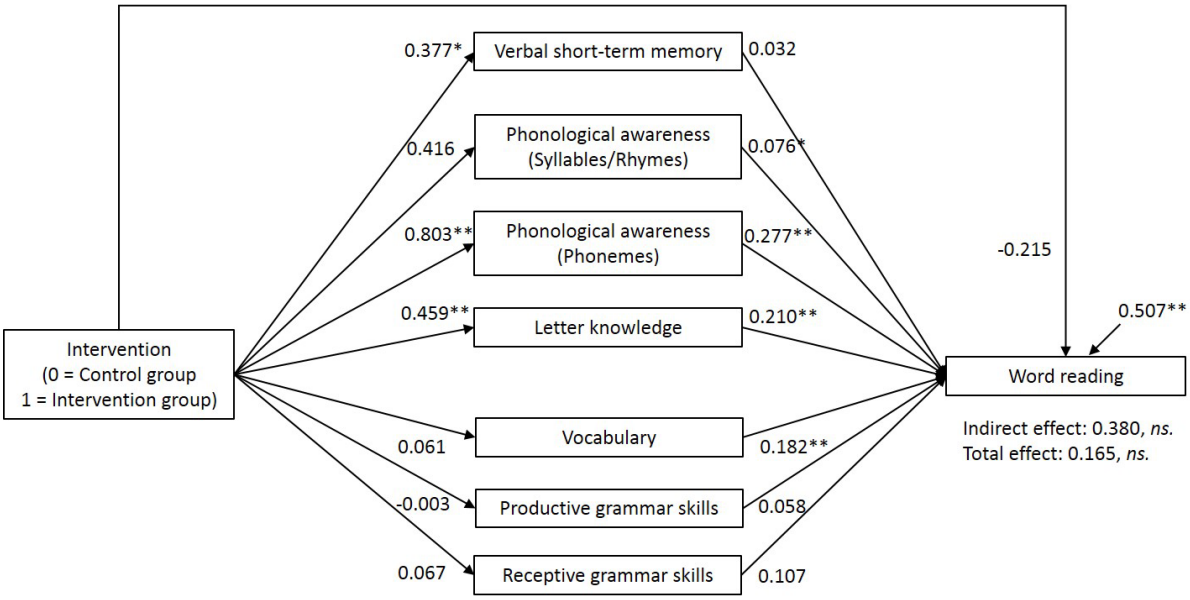

Figure 1. Effect of a training of phonological awareness on preschoolers' code-related precursors, oral language and word reading skills. Results show a semi-standardized solution (all variables were standardized to a mean of 0 and a standard deviation of 1, except for the variable indicating whether students belonged to the intervention group, coded as 1, or the control group, coded as 0). Results are weighted by using stabilized propensity score weights. Estimator = MLR. Type = complex. ns = not significant. * $p < .05$. ** $p < .01$.

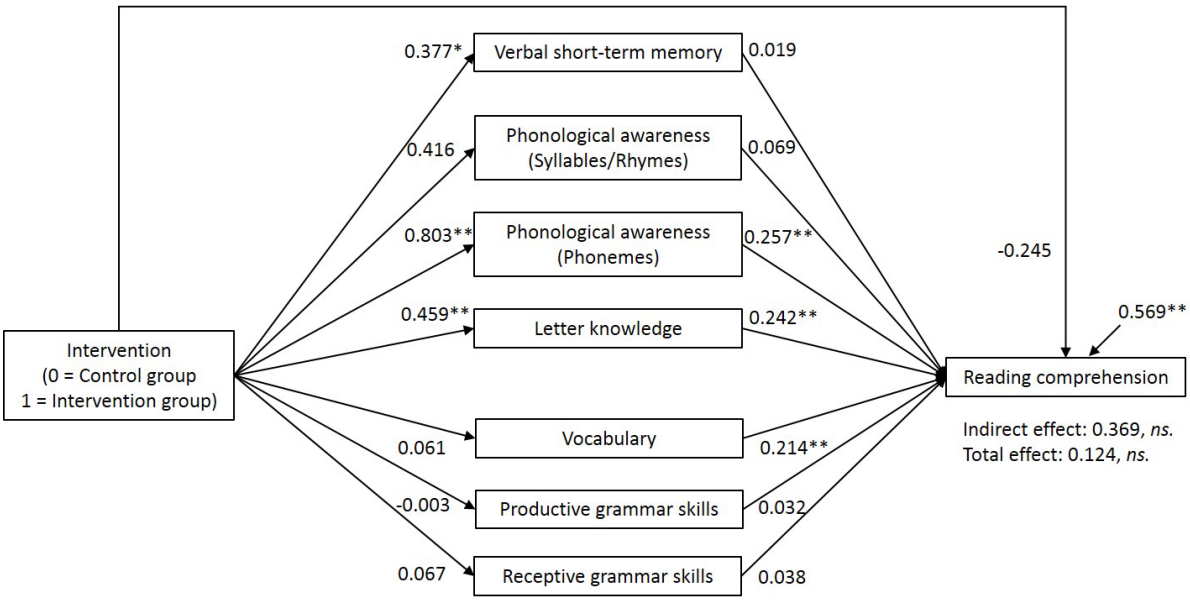

Figure 2. Effect of a training of phonological awareness on preschoolers' code-related precursors, oral language and reading comprehension skills. Results show a semi-standardized solution (all variables were standardized to a mean of 0 and a standard deviation of 1, except for the variable indicating whether students belonged to the intervention group, coded as 1, or the control group, coded as 0). Results are weighted by using stabilized propensity score weights. Estimator = MLR. Type = complex. ns = not significant. * $p < .05$. ** $p < .01$

Preliminary discussion

As expected, children who participated in the combined phonological awareness and letter knowledge training substantially improved their phoneme awareness and letter knowledge. Interestingly, children from the intervention group also showed higher verbal short-term memory skills in comparison to children from the control group after the intervention. Contrary to our expectations, we could not find any effects for children's word reading and reading comprehension one year after the intervention, and results of the mediation model underline this unexpected finding. Although, on the one hand, positive intervention effects were found for phoneme awareness and letter knowledge, and, on the other hand, individual differences in phoneme awareness and letter knowledge positively correlated with children's word reading and reading comprehension skills, this positive link did not lead to improvements in children's reading skills in the end. Therefore, the assumed causal theoretical model (cf. Hulme et al., 2012) did not hold.

Given that German in comparison to English has a more transparent orthography, findings from English students may not necessarily generalize to German students' reading acquisition (Ziegler & Goswami, 2006). As first grade students from English-speaking countries normally show lower word reading skills in comparison to first grade students from Germany (Seymour et al., 2003), we reanalyzed our data separately for children with above-average and below-average code-related precursors and oral language skills. Due to a lower orthographic transparency, children from English-speaking countries face more challenges in efficiently mapping oral to written language/phonological recoding, emphasizing phonological awareness skills for reading development (Landerl et al., 2013; Ziegler & Goswami, 2005). We assumed that German children who perform below average also more often have difficulties in phonological recoding. Therefore, these children might benefit more from participating in phonological awareness training in comparison to children who are performing above average. Consequently, the result pattern for low-performing German children might be more comparable to the previously discussed findings from studies that were run in English-speaking countries. As we did not

have clear-cut expectations on such group differences before conducting this study, this part of the study may be seen as exploratory.

Intervention effects for children with above- and below-average emergent literacy skills

In order to categorize children according to their emergent literacy skills, we first standardized children's t1 code-related precursors (verbal short-term memory, syllable and rhyme awareness, phonemic awareness, letter knowledge) and oral language (vocabulary, productive and receptive grammar) skills and calculated the mean test score on these measures for every child. Next, we used this aggregated test score to separate two groups at the 50th percentile (median split). Stabilized propensity score weights were estimated separately for each group and used again for all further analyses (see statistics on covariate balance in the Appendix, Table A1). Findings on intervention effects for code-related precursors and oral language skills at t2 (Model 1), word reading and reading comprehension at t4 (Model 2) as well as the combined mediation model (Model 3) are depicted in Table 3 for the low performing (below the 50th percentile) and in Table 4 for the high performing (above the 50th percentile) children.

For children performing below average (N_{IG} = 179, $N_{Control}$ = 56), our results indicated significant intervention effects for all code-related precursors. The lowest effect was found for verbal short-term memory (d = 0.388), whereas the highest effect was found for phoneme awareness (d = 1.255). Furthermore, in grade 1, we found a small to moderate intervention effect for children's word reading, just missing the 5% significance level (d = 0.444, p = .064). A closer look at the mediation model shows a moderate significant relation between children's word reading skills and phoneme awareness. Furthermore, the specific indirect intervention effect via phoneme awareness (B = 0.387, t = 3.442, p < .01) reached statistical significance as did the sum of all indirect effects. Children's vocabulary and productive grammar skills also correlated with children's word reading, but were not affected by the intervention. Concerning reading comprehension, the intervention effect was not significant (d = 0.439, p = .075), although we also found a specific indirect intervention effect via phoneme awareness (B = 0.325, t = 3.042, p < .01).

For children performing above average (N_{IG} = 191, $N_{Control}$ = 43), our results indicated a moderate to strong intervention effect for children's phoneme awareness (d = 0.725). Contrary to the group of children performing below average, intervention effects for verbal short-term memory, syllable and rhyme awareness as well as letter knowledge were low and did not reach statistical significance. However, it should be noted that ceiling effects were present for these variables. Furthermore, we did not find any intervention effects for children's grade 1 word reading and reading comprehension skills. An inspection of the mediation model reveals significant positive correlations between children's syllable and rhyme awareness and word reading as well as reading comprehension. Furthermore, children's letter knowledge was related to children's word reading and reading comprehension. Finally, we did not find any specific indirect intervention effects.

Discussion

At first, with regard to the immediate intervention effects, our findings clearly supported our expectations. Twenty weeks of intensive training of phonological awareness and letter knowledge substantially increased children's syllable and rhyme awareness, phoneme awareness, and letter knowledge. Effect sizes were moderate to large and well in line with the meta-analytic findings presented by Fischer and Pfost (2015) or Wolf et al. (2016), both focusing on German-speaking children. Effect sizes also corresponded well to the meta-analytic findings presented by Ehri et al. (2001) or the National Early Literacy Panel (2008), both integrating findings across several languages. Therefore, from an evaluation perspective, the immediate goal of this training on phonological awareness and letter knowledge was reached.

Table 3

Regression effects for children with below-average code-related and oral language skills

	Model 1 *B* (*SE*)	Model 2 *B* (*SE*)	Model 3 *B* (*SE*)
Intervention (IV) on DV:			
Code-related precursors			
Verbal short-term memory	0.388 (0.152)*		0.388 (0.152)*
Phonological awareness – syllables and rhymes	0.689 (0.250)**		0.689 (0.250)**
Phonological awareness – phonemes	1.255 (0.147)**		1.255 (0.147)**
Letter knowledge	0.880 (0.140)**		0.880 (0.140)**
Oral language skills			
Vocabulary	0.023 (0.372)		0.023 (0.372)
Productive grammar skills	0.100 (0.230)		0.100 (0.230)
Receptive grammar skills	-0.081 (0.280)		-0.081 (0.280)
Reading skills			
Word reading		0.444 (0.239)	-0.063 (0.175)
Reading comprehension		0.439 (0.247)	-0.008 (0.196)
Code-related and oral language skills (IV) on Word reading (DV)			
Verbal short-term memory			0.035 (0.063)
Phonological awareness – syllables and rhymes			0.015 (0.073)
Phonological awareness – phonemes			0.308 (0.093)**
Letter knowledge			0.083 (0.079)
Vocabulary			0.286 (0.069)**
Productive grammar skills			0.156 (0.071)*
Receptive grammar skills			-0.012 (0.073)
Code-related and oral language skills (IV) on Reading comprehension (DV)			
Verbal short-term memory			0.016 (0.062)
Phonological awareness – syllables and rhymes			0.021 (0.085)
Phonological awareness – phonemes			0.259 (0.085)**
Letter knowledge			0.091 (0.086)
Vocabulary			0.306 (0.058)**
Productive grammar skills			0.076 (0.067)
Receptive grammar skills			-0.083 (0.083)
Total indirect effect for word reading			0.507 (0.179)**
Total indirect effect for reading comprehension			0.448 (0.143)**

Note. All results are based on multiple imputed datasets (m = 5); Estimator = MLR; Type = Complex. Results are weighted by using stabilized propensity score weights. IV = independent variable; DV = dependent variable. Except for intervention (Intervention group = 1, Control group = 0), all variables are standardized.

* $p < .05$. ** $p < .01$.

Table 4

Regression effects for children with above-average code-related and oral language skills

	Model 1 *B* (*SE*)	Model 2 *B* (*SE*)	Model 3 *B* (*SE*)
Intervention (IV) on DV:			
Code-related precursors			
Verbal short-term memory	0.427 (0.298)		0.427 (0.298)
Phonological awareness – syllables and rhymes	0.348 (0.208)		0.348 (0.208)
Phonological awareness – phonemes	0.725 (0.209)**		0.725 (0209)**
Letter knowledge	0.281 (0.239)		0.281 (0.239)
Oral language skills			
Vocabulary	0.086 (0.199)		0.086 (0.199)
Productive grammar skills	-0.267 (0.225)		-0.267 (0.225)
Receptive grammar skills	0.160 (0.168)		0.160 (0.168)
Reading skills			
Word reading		-0.028 (0.183)	-0.307 (0.214)
Reading comprehension		0.011 (0.182)	-0.273 (0.221)
Code-related and oral language skills (IV) on Word reading (DV)			
Verbal short-term memory			0.015 (0.069)
Phonological awareness – syllables and rhymes			0.122 (0.062)*
Phonological awareness – phonemes			0.157 (0.103)
Letter knowledge			0.299 (0.086)**
Vocabulary			0.052 (0.046)
Productive grammar skills			-0.021 (0.077)
Receptive grammar skills			0.143 (0.077)
Code-related and oral language skills (IV) on Reading comprehension (DV)			
Verbal short-term memory			0.011 (0.053)
Phonological awareness – syllables and rhymes			0.155 (0.072)*
Phonological awareness – phonemes			0.150 (0.095)
Letter knowledge			0.331 (0.080)**
Vocabulary			0.076 (0.050)
Productive grammar skills			-0.009 (0.068)
Receptive grammar skills			0.091 (0.061)
Total indirect effect for word reading			0.280 (0.136)*
Total indirect effect for reading comprehension			0.284 (0.120)*

Note. All results are based on multiple imputed datasets (m = 5); Estimator = MLR; Type = Complex. Results are weighted by using stabilized propensity score weights. IV = independent variable; DV = dependent variable. Except for intervention (Intervention group = 1, Control group = 0), all variables are standardized.

* $p < .05$. ** $p < .01$.

Interestingly, the children from the intervention group also improved their verbal short-term memory, which was beyond the immediate scope of the applied training program. According to the comprehensive language approach (Dickinson et al., 2003; Rohde, 2015), emergent literacy skills develop interdependently. Although prior research on the development of children's reading skills often highlighted the critical role of children's verbal short-term memory (Swanson, Zheng, & Jerman, 2009; Wagner & Torgesen, 1987), past empirical research has not provided conclusive evidence for the effectiveness of short-term memory training on reading development (Redick, Shipstead, Wiemers, Melby-Lervåg, & Hulme, 2015). However, the improved test scores in verbal short-term memory might also reflect familiarity effects with such forms of test material due to the training practices and may not necessarily reflect improvements in the underlying cognitive memory system. A positive or negative replication of this finding may shed further light on this open question.

Word reading and reading comprehension

Regarding overall intervention effects for children's grade 1 reading skills, we neither found a significant positive effect for children's word reading skills nor for children's reading comprehension. Furthermore, effect sizes were quite small in descriptive terms. At first glance, and in comparison to the meta-analytic findings reported by Bus and van Ijzendoorn (1999), Ehri et al. (2001), Suggate (2016), or the National Early Literacy Panel (2008), this finding does not seem to replicate the often reported positive effects of a training of phonological awareness. However, this near-to-zero effect fits well into meta-analytic findings of studies with German-speaking children (Fischer & Pfost, 2015; Wolf et al., 2016) and may reflect the higher transparency of the German orthography in comparison to the English orthography. Prior research has shown that German students, in comparison to English-speaking students, reach a high level of reading accuracy quite early, due to a high mastery of phonological recoding (Frith, Wimmer, & Landerl, 1998; Seymour et al., 2003). Furthermore, reading instruction in the first year of primary school in Germany typically relies on systematic phonics instruction, directly teaching grapheme-phoneme correspondences and their usage for word reading (Landerl & Thaler, 2006). Additionally, teachers often make use of alphabet sound charts. Therefore, an intensive and effective phonics teaching that is characteristic of German first grade instruction might have contributed to the observation of no significant long-term effects for reading. To summarize, in addition to differences in transparency between the German and English orthographies, differences in reading instruction might also have contributed to our observation that the early advantage in phonological awareness resulting from training these skills in kindergarten had disappeared at the end of grade 1.

Nevertheless, and this seems interesting to us from a theoretical perspective, individual differences in phonological awareness and letter knowledge were strongly related to individual differences in first graders' word reading and reading comprehension skills. The inconsistency, therefore, is that although individual differences in children's phonological awareness and letter knowledge predict individual differences in reading, improving these skills does not necessarily improve children's reading skills. Although just preliminary, one possible explanation of this inconsistency may lie in the assumed causal model (see Scarborough, 2002). Instead of the dominance of a causal chain model (phonological awareness affects reading development), as reported by Hulme et al. (2012), for example, this finding may represent the dominance of an underlying condition model (a third variable, e.g. verbal or general cognitive abilities, affecting both individual differences in phonological awareness as well as individual differences in reading development). Furthermore, as suggested by Plomin, Shakeshaft, McMillan, and Trzaskowski (2014), relations that are observed in natural settings (the "what is" topic) do not necessarily relate to questions on the "what could be" topic of intervention effects. Due to variables such as the stability of the family environment (e.g., Sameroff, Seifer, Baldwin, & Baldwin, 1993), intervention effects do not automatically persist over longer time periods beyond the time of intervention or reproduce in the formation of further (reading) skills.

Intervention effects for children with above- versus below-average emergent literacy skills

As we could not confirm our expected finding of a positive intervention effect for children's word reading and reading comprehension skills, we further explored whether we could confirm our expectations for children with above-average and below-average emergent literacy skills. Within the group of children with below-average code-related precursors and oral language skills, we found strong intervention effects for children's phonological awareness and letter knowledge. Furthermore, at least in descriptive terms, we found small to medium intervention effects for children's word reading and reading comprehension skills. These effects just missed the 5% significance level, a fact that is also attributable to the reduced sample size. Furthermore, the mediation model reveals a positive indirect intervention effect via children's phoneme awareness. In sum, the overall picture for those children performing below average better corresponds to the expected (causal) model of improvements in phonological awareness leading to better reading skills (see Hulme et al., 2012). Children with below-average emergent literacy skills later reach a high level of word reading accuracy, in comparison to children with above-average verbal and code-related skills (National Early Literacy Panel, 2008). Comparable to children learning to read in a less transparent orthography, these children tend to depend more on phonological recoding while reading, due to a slower development of detailed orthographic representations (Share, 1995), which emphasizes individual differences in phonological awareness and facilitates positive training effects.

Within the group of children with above-average verbal and code-related skills, we only observed a minor intervention effect for children's phoneme awareness. In addition, we did not observe any direct intervention effects for children's reading skills. In sum, better performing children who were learning to read within the German orthography did not profit much from a phonological awareness and letter knowledge intervention program.

Limitations

Despite its strengths – a comparable large sample size and a broad use of different code-related precursors and oral language measures – some limitations of this study are equally worth mentioning. First, children were not randomly assigned to the intervention or control group. As some kindergartens already had implemented the phonological awareness training HLL (Küspert & Schneider, 2008; Plume & Schneider, 2004) into their routines, a complete randomization was not feasible. Nevertheless, observed pre-test differences were low and we corrected for such differences by using stabilized propensity score weights (Austin & Stuart, 2015; Leite, 2017). In addition, as the intervention program had already found its way into practice, the ecological validity of our findings is high. Second, concerning the mediation effects, some difficulties with regard to the causal interpretation of our findings remain. In addition to the linearity and no-interaction assumption, a causal interpretation of mediation effects requires the assumptions of sequential ignorability that may hardly be tested directly and may not hold necessarily (Imai, Keele, & Tingley, 2010). Third, we did not have any information on children's word reading skills, neither before the beginning nor directly after the phonological awareness intervention. Therefore, we could neither rule out the possibility that children already had substantial word reading skills before the beginning of the intervention which might have restricted effects of the phonological awareness training, nor could we rule out the possibility of immediate direct intervention effects on word reading skills when children still attended kindergarten. Finally, we did not analyze treatment effects separately for the different phonological awareness subtests due to small numbers of items within each subtest.

Conclusion

We analyzed the effects of training phonological awareness on children's emergent literacy and reading skills. Although the training was quite effective at promoting children's phonological awareness and letter knowledge, we found near-to-zero transfer effects for children's reading skills. This pattern was

inconsistent with our assumed causal model of phonological awareness affecting reading development. Effects were slightly stronger for children with below-average oral language and code-related skills. In our opinion, these small effects may be attributable to the more consistent German orthography in comparison to the English orthography and differences in first grade reading instruction.

From a theoretical perspective, our findings are at odds with the often implicitly stated assumption that causal models of reading acquisition are closely comparable across different orthographies (see Share, 2008, for further criticism). At least for the German orthography, future research might consider reformulating and further evaluating the importance of phonological awareness when learning to read. Specific phonological awareness interventions might be worth considering for children with deficits in emergent literacy skills. However, our findings cast doubt on the assumption that phonological awareness equally contributes to individual differences in reading skills for all children independent of their skill level, and across all orthographies. Apparently, reading researchers need to be careful when deriving generalizations of findings from the English orthography into other, more transparent orthographies and vice versa. Furthermore, and in addition to the question of differences among orthographies, future research might take characteristics of first grade reading instruction into account. This seems especially interesting as typical modes of reading instruction might not just relate but also interact with orthographic characteristics, which opens the field for more intervention studies across orthographies and methods of reading instruction.

Finally, from a practical point of view, we think that it might be worth reflecting more carefully on how training programs for phonological awareness could be better adapted to the individual needs of children (but see Kjeldsen et al., 2019, for a different view). Children with deficits in emergent literacy skills might profit considerably from a structured phonological awareness training in preschool. Children with no such deficits, however, tend to profit less from this type of structured add-on intervention. As children already in kindergarten show considerable individual differences in their competence profiles, it seems worth discussing how limited time resources can be used best and adapted to these differences. Screening procedures could help to identify children with deficits in language and emergent literacy skills and to offer adapted training possibilities. These trainings might also go beyond traditional phonological awareness intervention procedures, including other language skill components and ensuring that children's self-concept and motivation will be positively affected by the training experience.

References

Anthony, J. L., & Francis, D. J. (2005). Development of phonological awareness. *Current Directions in Psychological Science, 14*, 255-259. doi:10.1111/j.0963-7214.2005.00376.x

Aro, M. (2006). Learning to read: The effect of orthography. In R. M. Joshi & P. G. Aaron (Eds.), *Handbook of orthography and literacy* (pp. 531-550). Mahwah, NJ: Lawrence Erlbaum Associates.

Austin, P. C., & Stuart, E. A. (2015). Moving toward best practice when using inverse probability of treatment weighting (IPTW) using the propensity score to estimate causal treatment effects in observational studies. *Statistics in Medicine, 34*, 3661-3679. doi:10.1002/sim.6607

Barron, R. W. (1991). Proto-literacy, literacy and the acquisition of phonological awareness. *Learning and Individual Differences, 3*, 243-255. doi:10.1016/1041-6080(91)90010-X

Blatter, K. (2015). *Familiale sprachbezogene Förderung und frühe (schrift-)sprachliche Kompetenzen. Zusammenhänge bei Kindern mit und ohne Migrationshintergrund [Home literacy environment activities and early language and literacy skills. Interrelations in children with and without a migration background]*. Otto-Friedrich-Universität Bamberg: Bamberg, Germany.

Blatter, K., Faust, V., Jäger, D., Schöppe, D., Artelt, C., Schneider, W., & Stanat, P. (2013). Vorschulische Förderung der phonologischen Bewusstheit und der Buchstaben-Laut-Zuordnung: Profitieren auch Kinder nichtdeutscher Herkunftssprache? [How effective is a preschool training of phonological awareness and letter sound correspondence for student's with migration background?]. In A. Redder & S. Weinert (Eds.), *Sprachförderung und Sprachdiagnostik: Interdisziplinäre Perspektiven* (pp. 218-238). Münster, Germany: Waxmann.

Bradley, L., & Bryant, P. E. (1983). Categorizing sounds and learning to read: A causal connection. *Nature, 301*, 419-421. doi:10.1038/301419a0

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. doi:10.1037/0022-0663.91.3.403

Cardoso-Martins, C. (1995). Sensitivity to rhymes, syllables and phonemes in literacy acquisition in Portuguese. *Reading Research Quarterly, 30*, 808-828. doi:10.2307/748199

Castles, A., & Coltheart, M. (2004). Is there a causal link from phonological awareness to success in learning to read? *Cognition, 91*, 77-111. doi:10.1016/S0010-0277(03)00164-1

Chall, J. S. (1983). *Stages of reading development*. New York, NY: McGraw-Hill.

Dickinson, D. K., Golinkoff, R. M., & Hirsh-Pasek, K. (2010). Speaking out for language: Why language is central to reading development. *Educational Researcher, 39*, 305-310. doi:10.3102/0013189X10370204

Dickinson, D. K., McCabe, A., Anastasopoulos, L., Peisner-Feinberg, E. S., & Poe, M. D. (2003). The comprehensive language approach to early literacy: The interrelationships among vocabulary, phonological sensitivity, and print knowledge among preschool-aged children. *Journal of Educational Psychology, 95*, 465-481. doi:10.1037/0022-0663.95.3.465

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. doi:10.1598/RRQ.36.3.2

Fischer, M. Y., & Pfost, M. (2015). Wie effektiv sind Maßnahmen zur Förderung der phonologischen Bewusstheit? Eine meta-analytische Untersuchung der Auswirkungen deutschsprachiger Trainingsprogramme auf den Schriftspracherwerb [How effective are trainings of phonological awareness? A meta-analysis of German language training programs and their effects on the acquisition of reading and spelling skills]. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie, 47*, 35-51. doi:10.1026/0049-8637/a000121

Foulin, J. N. (2005). Why is letter-name knowledge such a good predictor of learning to read? *Reading and Writing, 18*, 129-155. doi:10.1007/s11145-004-5892-2

Fox, A. V. (2006). *Test zur Überprüfung des Grammatikverständnisses (TROG-D) [Test for reception of grammar]*. Idstein, Germany: Schulz-Kirchner Verlag.

Fricke, S., Szcerbinski, M., Stackhouse, J., & Fox-Boyer, A. V. (2008). Predicting individual differences in early literacy acquisition in German. *Written Language & Literacy, 11*, 103-146. doi:10.1075/wll.11.2.02fri

Frith, U., Wimmer, H., & Landerl, K. (1998). Differences in phonological recoding in German- and English-speaking children. *Scientific Studies of Reading, 2*, 31-54. doi:10.1207/s1532799xssr0201_2

Goswami, U. (2002). Early phonological development and the acquisition of literacy. In S. B. Neuman & D. K. Dickinson (Eds.), *Handbook of early literacy research. Volume 1* (pp. 111-125). New York, NY: Guilford Press.

Goswami, U., & Bryant, P. E. (1990). *Phonological skills and learning to read*. Hove, UK: Psychology Press.

Graham, J. W. (2009). Missing data analysis: Making it work in the real world. *Annual Review of Psychology, 60*, 549-576. doi:10.1146/annurev.psych.58.110405.085530

Grimm, H., & Schöler, H. (1998). *Heidelberger Sprachentwicklungstest (HSET) [Heidelberg language development test]*. Göttingen, Germany: Hogrefe.

Hulme, C., Bowyer-Crane, C., Carroll, J. M., Duff, F. J., & Snowling, M. J. (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. doi:10.1177/0956797611435921

Imai, K., Keele, L., & Tingley, D. (2010). A general approach to causal mediation analysis. *Psychological Methods, 15*, 309-334. doi:10.1037/a0020761

Jansen, H., Mannhaupt, G., Marx, H., & Skowronek, H. (2002). *BISC. Bielefelder Screening zur Früherkennung von Lese-Rechtschreibschwierigkeiten (2nd ed.) [BISC. The Bielefeld screening for an early detection of reading and writing difficulties]*. Göttingen, Germany: Hogrefe.

Kjeldsen, A.-C., Educ, L., Saarento-Zaprudin, S. K., & Niemi, P. O. (2019). Kindergarten training in phonological awareness: Fluency and comprehension gains are greatest for readers at risk in grades 1 through 9. *Journal of Learning Disabilities, 52*, 366-382. doi:10.1177/0022219419847154

Küspert, P., & Schneider, W. (2008). *Hören, lauschen, lernen. Sprachspiele für Kinder im Vorschulalter. Würzburger Trainingsprogramm zur Vorbereitung auf den Erwerb der Schriftsprache [The Würzburg preschool training programme to promote children's reading acquisition]* (6 ed.). Göttingen: Vandenhoeck & Ruprecht.

Landerl, K., Ramus, F., Moll, K., Lyytinen, H., Leppänen, P. H. T., Lohvansuu, K., . . . Schulte-Körne, G. (2013). Predictors of developmental dyslexia in European orthographies with varying complexity. *Journal of Child Psychology and Psychiatry, 54*, 686-694. doi:10.1111/jcpp.12029

Landerl, K., & Thaler, V. (2006). Reading and spelling acquisition and dyslexia in German. In R. M. Joshi & P. G. Aaron (Eds.), *Handbook of orthography and literacy* (pp. 121-134). Mahwah, NJ: Erlbaum.

Leite, W. (2017). *Practical propensity score methods using R*. Thousand Oaks, CA: Sage.

Lenhard, W., & Schneider, W. (2005). *ELFE 1-6: Ein Leseverständnistest für Erst- bis Sechstklässler. [ELFE 1-6. A reading comprehension test for students from the 1st through the 6th grade]* (1 ed.). Göttingen, Germany: Hogrefe.

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.

Mann, V., & Wimmer, H. (2002). Phoneme awareness and pathways into literacy: A comparison of German and American children. *Reading and Writing, 15*, 653-682. doi:10.1023/A:1020984704781

Melby-Lervåg, M., Lyster, S.-A. H., & Hulme, C. (2012). Phonological skills and their role in learning to read: A meta-analytic review. *Psychological Bulletin, 138*, 322-352. doi:10.1037/a0026744

Metsala, J. L. (1999). Young children's phonological awareness and nonword repetition as a function of vocabulary development. *Journal of Educational Psychology, 91*, 3-19. doi:10.1037/0022-0663.91.1.3

Morais, J., Cary, L., Alegria, J., & Bertelson, P. (1979). Does awareness of speech as a sequence of phones arise spontaneously? *Cognition, 7*, 323-331. doi:10.1016/0010-0277(79)90020-9

Moser, U., & Berweger, S. (2007). *wortgewandt & zahlenstark. Lern- und Entwicklungsstand bei 4- bis 6-Jährigen [wortgewandt & zahlenstark. A test of early academic skills for 4 to 6 year old children]*. Zürich, Suisse: Lehrmittelverlag des Kantons Zürich.

Muter, V., Hulme, C., Snowling, M. J., & Stevenson, J. (2004). Phonemes, rimes, vocabulary, and grammatical skills as foundations of early reading development: Evidence from a longitudinal study. *Developmental Psychology, 40*, 665-681. doi:10.1037/0012-1649.40.5.665

Muthén, L. K., & Muthén, B. O. (1998-2012). *Mplus user's guide. Seventh edition.* Los Angeles, CA: Muthén & Muthén.

National Early Literacy Panel. (2008). *Developing early literacy: Report of the National Early Literacy Panel*. Washington, DC: National Institute for Literacy.

Paris, S. G. (2005). Reinterpreting the development of reading skills. *Reading Research Quarterly, 40*, 184-202. doi:10.1598/RRQ.40.2.3

Pfost, M. (2015). Children's phonological awareness as a predictor of reading and spelling: A systematic review of longitudinal research in German-speaking countries. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie, 47*, 123-138. doi:10.1026/0049-8637/a000141

Pfost, M. (2016). *Stabilität, Kontinuität und interindividuelle Unterschiede in der Entwicklung schriftsprachlicher Kompetenzen [Stability, continuity and interindividual differences in reading competence development]*. Bamberg: University of Bamberg Press.

Phillips, B. M., & Torgesen, J. K. (2006). Phonemic awareness and reading: Beyond the growth of initial reading accuracy. In D. K. Dickinson & S. B. Neuman (Eds.), *Handbook of early literacy research* (Vol. 2, pp. 101-112). New York, NY: Guilford Press.

Plomin, R., Shakeshaft, N. G., McMillan, A., & Trzaskowski, M. (2014). Nature, nurture, and expertise. *Intelligence, 45*, 46-59. doi:10.1016/j.intell.2013.06.008

Plume, E., & Schneider, W. (2004). *Hören, lauschen, lernen 2. Spiele mit Buchstaben und Lauten für Kinder im Vorschulalter. Würzburger Buchstaben-Laut-Training [The Würzburg preschool training programme for promoting children's reading acquisition including a training of letter-sound correspondence]*. Göttingen: Vandenhoeck & Ruprecht.

Redick, T. S., Shipstead, Z., Wiemers, E. A., Melby-Lervåg, M., & Hulme, C. (2015). What's working in working memory training? An educational perspective. *Educational Psychology Review, 27*, 617-633. doi:10.1007/s10648-015-9314-6

Rieben, L., & Perfetti, C. A. (Eds.). (1991). *Learning to read: Basic research and its implications*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Rohde, L. (2015). The comprehensive emergent literacy model: Early literacy in context. *SAGE Open, 5*, 1-11. doi:10.1177/2158244015577664

Roßbach, H.-G., Tietze, W., & Weinert, S. (2005). *Peabody Picture Vocabulary Test - Revised*. Unpublished German research version of the PPVT-R (C.M. Dunn & L.M. Dunn, 1981). Otto-Friedrich-Universität Bamberg, Germany; Freie Universität Berlin, Germany.

Sameroff, A. J., Seifer, R., Baldwin, A., & Baldwin, C. (1993). Stability of intelligence from preschool to adolescence: The influence of social and family risk factors. *Child Development, 64*, 80-97. doi:10.1111/j.1467-8624.1993.tb02896.x

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

Scarborough, H. S. (2002). Connecting early language and literacy to later reading (dis)abilities: Evidence, theory, and practice. In S. B. Neuman & D. K. Dickinson (Eds.), *Handbook of early literacy research* (pp. 97-110). New York, NY: Guilford Press.

Schneider, W., Artelt, C., & Stanat, P. (2018). Kurz- und langfristige Effekte eines Trainings zur phonologischen Bewusstheit bei Kindergartenkindern deutscher und nichtdeutscher Herkunftssprache (Verbundvorhaben) [Short- and long-term effects of training phonological awareness in native and non-native German-speaking kindergarten children]. Retrieved 16. October, 2018, from https://www.empirische-bildungsforschung-bmbf.de/de/654.php

Schneider, W., Blanke, I., Faust, V., & Küspert, P. (2011). *Würzburger Lese Leseprobe - Revision (WLLP-R) [Würzburg silent reading test - revised version]*. Göttingen, Germany: Hogrefe.

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

Shankweiler, D., Lundquist, E., Katz, L., Stuebing, K. K., Fletcher, J. M., Brady, S., . . . Shaywitz, B. A. (1999). Comprehension and decoding: Patterns of association in children with reading difficulties. *Scientific Studies of Reading, 3*, 69-94. doi:10.1207/s1532799xssr0301_4

Share, D. L. (1995). Phonological recoding and self-teaching: Sine qua non of reading acquisition. *Cognition, 55*, 151-218. doi:10.1016/0010-0277(94)00645-2

Share, D. L. (2008). On the anglocentrics of current reading research and practice: The perils of overreliance on an "outlier"orthography. *Psychological Bulletin, 134*, 584-615. doi:10.1037/0033-2909.134.4.584

Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly, 21*, 360-407. doi:10.1598/RRQ.21.4.1

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. doi:10.1037/0012-1649.38.6.934

Suggate, S. (2016). A meta-analysis of the long-term effects of phonemic awareness, phonics, fluency, and reading comprehension interventions. *Journal of Learning Disabilities, 49*, 77-96. doi:10.1177/0022219414528540

Swanson, H. L., Zheng, X., & Jerman, O. (2009). Working memory, short-term memory, and reading disabilities: A selective meta-analysis of the literature. *Journal of Learning Disabilities, 42*, 260-287. doi:10.1177/0022219409331958

van Buuren, S., & Groothuis-Oudshoorn, K. (2011). mice: Multivariate imputation by chained equations in R. *Journal of Statistical Software, 45 (3)*, 1-67. doi:10.18637/jss.v045.i03

Wagner, R. K., & Torgesen, J. K. (1987). The nature of phonological processing and its causal role in the acquisition of reading skills. *Psychological Bulletin, 101*, 192-212. doi:10.1037/0033-2909.101.2.192

Whitehurst, G. J., & Lonigan, C. J. (1998). Child development and emergent literacy. *Child Development, 69*, 848-872. doi:10.2307/1132208

Wolf, K. M., Schroeders, U., & Kriegbaum, K. (2016). Metaanalyse zur Wirksamkeit einer Förderung der phonologischen Bewusstheit in der deutschen Sprache [Promoting phonological awareness in German: A meta-analysis]. *Zeitschrift für Pädagogische Psychologie, 30*, 9-33. doi:10.1024/1010-0652/a000165

Ziegler, J. C., 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. doi:10.1177/0956797610363406

Ziegler, J. C., & Goswami, U. (2005). Reading acquisition, developmental dyslexia, and skilled reading across languages: A psycholinguistic grain size theory. *Psychological Bulletin, 131*, 3-29. doi:10.1037/0033-2909.131.1.3

Ziegler, J. C., & Goswami, U. (2006). Becoming literate in different languages: Similar problems, different solutions. *Developmental Science, 9*, 429-453. doi:10.1111/j.1467-7687.2006.00509.x

Appendix A

The first three columns of Table A1 show unweighted mean scores and the standard deviation of the whole sample (column 1) as well as for the subgroup of children categorized as performing above (column 2) and below average (column 3) on pre-test code-related precursors and oral language skills. The standardized difference between these two subgroups is presented in column 4. Children categorized as performing above average on code-related precursors and oral language skills were significantly older and had better educated parents with a higher socio-economic status (HISEI). Furthermore, parents less often had a migration background. Differences in code-related precursors and oral language skills were about one standard deviation when compared with the children categorized as performing below average on code-related precursors and oral language skills.

Columns 5 and 6 show unweighted standardized mean differences between children from the intervention group and control group within the above average and below average performing groups. Especially within the group of children categorized as performing below average on code-related precursors and oral language skills (column 6), children in the intervention group had better educated parents with a higher socio-economic status (HISEI) in comparison to children from the control group. Furthermore, children in the intervention group less often had parents with a migration background. Finally, children in the intervention group had a higher vocabulary and better receptive grammar skills. The last two columns (columns 7 and 8) show weighted standardized mean differences between children from the intervention group and control group within the above average and below average performing groups. Weighted results did not show any differences (covariate imbalance) between children from the intervention and control group for all measures of the sociodemographic background and pre-test skills.

Table A1

Weighted and unweighted pre-test differences between a) preschool children categorized as performing above average (high) or below average (low) on pre-test code-related precursors and oral language skills; b) preschool children of the intervention (IG) and control group within different emergent literacy skills categories

Weigths:	Unweighted						Weighted	
Pre-Test Skills:	Total	High	Low		High	Low	High	Low
	M (*SD*)	*M* (*SD*)	*M* (*SD*)	Δ (High-Low) *d*	Δ (IG-Control) *d*	Δ (IG-Control) *d*	Δ (IG-Control) *d*	Δ (IG-Control) *d*
Sociodemographic background								
Sex (1=female)	0.49 (0.50)	0.53 (0.50)	0.46 (0.50)	0.124	0.034	-0.095	0.176	-0.089
Age (in months)	66.63 (4.01)	67.21 (3.92)	66.05 (4.03)	0.288**	0.109	0.114	0.017	-0.067
HISEI	49.26 (18.21)	55.57 (17.40)	42.97 (16.78)	0.693**	-0.004	0.310*	-0.025	0.030
Migration – one parent	0.12 (0.32)	0.09 (0.28)	0.15 (0.36)	-0.191*	-0.127	-0.122	0.012	0.038
Migration – both parents	0.32 (0.47)	0.16 (0.36)	0.48 (0.50)	-0.701**	-0.025	-0.365*	0.063	0.056
Parental education – upper track	0.48 (0.50)	0.61 (0.49)	0.34 (0.48)	0.537**	0.229	0.467**	0.089	0.117
Parental education – middle track	0.31 (0.46)	0.30 (0.46)	0.31 (0.46)	-0.017	-0.119	-0.198	-0.083	0.018
Code-related precursors								
Verbal short-term memory	6.89 (2.07)	7.78 (1.56)	6.00 (2.15)	0.857**	0.027	-0.017	0.048	-0.054
Phonological awareness – syllables and rhymes	15.73 (3.46)	17.55 (2.14)	13.92 (3.57)	1.050**	-0.191	0.187	0.059	0.058
Phonological awareness – phonemes	19.76 (8.01)	25.01 (6.96)	14.53 (5.01)	1.310**	0.223	0.005	0.031	0.074
Letter knowledge	9.03 (7.75)	12.95 (7.95)	5.13 (5.15)	1.011**	0.089	0.039	-0.045	0.143
Oral language skills								
Vocabulary	42.68 (23.56)	57.54 (18.42)	27.88 (18.22)	1.260**	0.079	0.365*	-0.040	0.045
Productive grammar skills	15.59 (8.37)	20.32 (7.04)	10.89 (6.78)	1.129**	-0.465**	0.287	-0.099	0.030
Receptive grammar skills	27.10 (6.46)	31.21 (3.00)	23.00 (6.38)	1.272**	0.192	0.381*	0.052	0.007

Note. All results are based on multiple imputed datasets ($m = 5$); Estimator = MLR; IG = Intervention group. Weighted results make use of stabilized propensity score weights.
* $p < .05$. ** $p < .01$.