

PDF hosted at the Radboud Repository of the Radboud University Nijmegen

The following full text is a publisher's version.

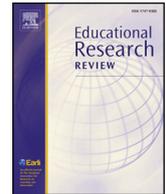
For additional information about this publication click this link.

<https://hdl.handle.net/2066/217239>

Please be advised that this information was generated on 2020-11-24 and may be subject to change.

Contents lists available at [ScienceDirect](https://www.sciencedirect.com)

Educational Research Review

journal homepage: www.elsevier.com/locate/edurev

Computer-supported early literacy intervention effects in preschool and kindergarten: A meta-analysis

Ludo Verhoeven^{a,c,*}, Marinus Voeten^a, Ellie van Setten^{a,b}, Eliane Segers^a^a Behavioural Science Institute, Radboud University, the Netherlands^b University Medical Center Groningen, the Netherlands^c University of Curaçao, Curaçao

ARTICLE INFO

Keywords:

Phonological awareness
 Early literacy
 Orthographic depth
 Computer-supported intervention
 Preschool
 Kindergarten
 Study design

ABSTRACT

A meta-analysis was conducted on the effects of computer-supported early literacy interventions (strict phonological awareness training, combined phonological awareness and letter training, and use of e-books) on phonological-awareness (syllabic awareness, word blending, rhyme, phoneme awareness) and reading-related skills (concept about print, letter knowledge, decoding, spelling) across different languages in preschool and kindergarten since 1995. A total of 59 studies were identified with a total amount of 339 effect sizes, involving 6786 preschool and kindergarten students. A multilevel approach was followed to estimate the average effect size and to examine the moderation of effects. On average, a small positive effect size of 0.28 with 95% CI (0.21, 0.35), was evidenced across treatments and across outcome measures. Large variation in effect sizes was observed between studies and especially between comparisons within studies. If the intervention was part of an integrated learning system in the classroom, the effects were better. The present analysis also shows the importance of methodological rigor of study designs being used in that effect sizes were higher in cases of absence of randomization and comparisons with classroom teaching as usual instead of an active control group.

From an early age on, children learn about literacy in communicative settings around a broad range of language games and storybook reading (Neuman & Dickinson, 2011). Different models have been proposed in the literature to define the knowledge and abilities that are involved in early literacy. Whitehurst and Lonigan (2001) proposed that early literacy combines outside-in and inside-out processes, the former being related to language comprehension and background knowledge, the latter to procedural knowledge related to phonological awareness and letter knowledge. Along the same lines, Sénéchal, LeFevre, Smith-Chant, and Colton (2001) make a distinction between language-related skills, metalinguistic skills, and basic literacy skills (i.e., letter knowledge, decoding and spelling). In her comprehensive emergent literacy model, Rohde (2015) defines early literacy as a construct of partly overlapping components related to language, phonological awareness and print awareness, and claims that each of these have their own developmental trajectory.

Communication and language can be seen as the common ground on which the emergence of literacy is built on (Moats, 2000). Phonological awareness - or the ability to attend to the sounds of language independent of meaning - is generally regarded conditional for learning to read (Anthony & Lonigan, 2004; Caravolas, Lervåg, Defior, Seidlová Málková, & Hulme, 2013; Goswami, 2001; Lonigan, Burgess, & Anthony, 2000). In order to grasp the alphabetic principle, children must become aware of the segmental nature of speech and learn that words consist of phonemes to which letters can be assigned (Liberman & Shankweiler, 1985; Torgesen,

* Corresponding author. Behavioural Science Institute, Radboud University, P.O. box 9104, 6500 HE, Nijmegen, the Netherlands.
 E-mail address: L.Verhoeven@pwo.ru.nl (L. Verhoeven).

<https://doi.org/10.1016/j.edurev.2020.100325>

Received 16 August 2019; Received in revised form 9 January 2020; Accepted 23 February 2020

Available online 04 March 2020

1747-938X/ © 2020 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

2002). Although implicit sensitivity to language starts already early in life, there is a spurt in the development of phonological awareness in preschool and kindergarten (Torgesen & Mathes, 2000). This is often promoted by playing games that help children to focus on the phonological constituents of language. Given the intrinsic relation between phonological awareness and literacy, it is assumed that a combined training of letters and sounds can be even more effective (Bryant & Bradley, 1985). It has also been argued that early literacy training should be offered to children in a meaningful context. Research has indeed shown that interactive activities during storybook reading may help children to get insight into the functions and structure of written language and to uncover the written code (see Mol, Bus, & de Jong, 2009).

In the past two decades, there has been an enormous flux of computer-based programs to facilitate children's early literacy. Review studies have shown that such programs are promising in that a tailored approach can be offered in response to the needs of the individual child (Bus & van IJzendoorn, 1999; Mol et al., 2009). Although technology-enhanced learning has become mainstream in many schools, it is by no means clear under what conditions software programs are most effective. It is also not clear whether effects of early-literacy treatments are sensitive to the language involved. In older studies, the focus was mainly on literacy learning in English, which can be considered an opaque orthography, whereas in more recent studies technology-enhanced early literacy learning in more transparent languages was also focused on. Therefore, in this meta-analysis an attempt will be made to evaluate the differential outcomes of a broad variety of computer-based training programs fostering children's early literacy across different languages in preschool and kindergarten.

1. Development of early literacy

When children are engaged in an environment with literates, they have the opportunity to learn that print carries meaning, that written texts may have various forms and functions, and that ideas can be expressed with conventional writing. In the case of alphabetic languages, children learn that words consist of phonemes which can be represented by letters. There is general agreement that in the case of alphabetic writing systems the acquisition of literacy involves the rediscovering of the principles of phonological recoding with phonological awareness as key component (Castles, Rastle, & Nation, 2018; Ehri, 2005, 2014). Phonological awareness refers to the access to and the understanding of the sound structure of oral language (e.g., Caravolas, Lervåg, Defior, SeidlováMálková, & Hulme, 2013). Phonological awareness requires children to reflect consciously on the phonological segments of spoken words and to manipulate them in a systematic way. As such, phonological awareness depends on the capacity to focus attention on the perceptual representations of speech (Tallal, 2004). It can be assessed by tasks measuring segmentation, blending, and manipulation of speech sounds (Hogan, Catts, & Little, 2005; Vloedgraven & Verhoeven, 2007).

Research has shown that across languages children follow more or less the same sequence of phonological awareness development from access to larger phonological units to smaller units of sound (Anthony & Francis, 2005; Ziegler & Goswami, 2005). Indeed, it has been evidenced that the development of phonological awareness progresses from the syllable level and the onset-rime level to the phoneme level (Shankweiler & Liberman, 1989; Vloedgraven & Verhoeven, 2009). Awareness at the phoneme level, i.e., phonemic awareness, concerns the awareness of phonemes, the speech sounds that are used to build spoken words and to distinguish meanings (Goswami, 2000; Nagy & Scott, 2000). It is rather difficult for children because of the fact that phonemes are acoustically evanescent and thus hard to detect (Goswami, 2001). Letters may facilitate the perception of phonemes and are considered to be essential for the transfer to word reading and word spelling (NRP, 2000; Scarborough, 1998; Share, 2004). Therefore, it is important to consider the effects of training phonological awareness with and without the inclusion of letters.

Previous research has shown that the awareness of sounds (syllables, onsets, rimes) in a word is associated with children's literacy development (Melby-Lervåg, Lyster, & Hulme, 2012; Nagy & Scott, 2000). Numerous studies have indeed evidenced a relation between measures of phonological awareness administered in preschool or kindergarten and tests of word recognition among the same children in first or second grade (Blachman, 2000; Swanson, Trainin, Necochea, & Hammill, 2003). However, the relation can also be reversed in that literacy experiences facilitate phonological awareness (Lerner & Lonigan, 2015; Perfetti, Beck, Bell, & Hughes, 1987). Support has also been provided that lack of phonological skills can cause difficulties with the acquisition of reading and writing. Poor readers have been found to be less precise in phoneme discrimination and phoneme segmentation (Elbro & Scarborough, 2003a,b; Høien & Lundberg, 2000; Snowling, 2000). To conclude, research has abundantly shown that phonological awareness is related to later literacy outcomes and that later literacy problems are associated with persistent problems in phonological awareness. The predictive power of phonological awareness for later literacy has prompted curriculum specialists to develop phonological awareness interventions in preschool and kindergarten.

1.1. Intervention studies

Early interventions for phonological awareness have focused on explicit instruction in the perception of sound constituents in words and complementary drill and practice. In some cases, it was taught without making any connections to literacy, in other cases phonological awareness was taught in connection with letters.

Bus and van IJzendoorn (1999) conducted a meta-analysis on experimental training studies on phonological awareness in perspective of reading gains with all studies following a randomized or matched design. The overall effect sizes for phonological awareness and reading in these experimental training studies were 0.73 and 0.70, respectively. In long-term studies on the influence of phonological awareness training on reading the combined effect size was much smaller. There was also research evidence that phonemic awareness can be seen as a critical component in understanding the alphabetic principle. Programs that combined phonological and letter training obtained larger effects than purely phonological programs.

Another meta-analysis evaluating the effects of phonemic awareness instruction on learning to read and spell was conducted by the National Reading Panel (Ehri et al., 2001). There were 52 studies published in peer-reviewed journals, and these contributed 96 cases comparing the outcomes of treatment and control groups. Analysis of effect sizes revealed that the impact of instruction on helping children acquire phonological awareness was large and statistically significant ($d = 0.86$) and also exerted a moderate, statistically significant impact on later reading ($d = 0.53$) and spelling ($d = 0.59$). It helped various types of children: normally developing readers as well as at-risk and disabled readers; preschoolers, kindergartners, and first graders; low socioeconomic status children as well as mid-high SES. Importantly, phonological awareness instruction was more effective when it was taught in combination with letters and written words than with speech sounds and spoken words only.

More recently, the National Early Literacy Panel (NELP) examined the effectiveness of preschool and kindergarten interventions aimed at early literacy and conventional literacy in three separate meta-analyses. To begin with, Lonigan, Schatschneider, and Westberg (2008) evaluated the effectiveness of 78 code-based instructional practices to enhance early literacy skills. They found moderate to large effects for code-focused interventions designed to teach the alphabetic principle, usually in combination with phonological awareness. In addition, Lonigan, Shanahan, and Cunningham (2008) investigated the effects of 19 shared book reading studies focused on early literacy outcomes. They found moderate effects on print knowledge and oral language. Finally, Molfese and Westberg (2008) did a meta-analysis on 33 early literacy intervention studies. Overall, the interventions yielded substantial effects on reading readiness measures and small effects on spelling. Twelve studies also showed effects on oral language and nine studies on reading. The overall conclusion from NELP was that early literacy interventions are especially good in enhancing children's reading readiness and thus in preparing children for school entry and learning to read.

Aside from strict phonological awareness and/or letter training, storybook reading can be seen as an effective means to enhance young children's early literacy development. In an eye-tracking study, Evans, Williamson, and Pursoo (2008) aimed to show how pointing to words during storybook reading by the teacher may facilitate early literacy. They evidenced that teachers' pointing to the words increased print-looking time and print target recognition in four-year-olds. Bus, van IJzendoorn, and Pellegrini (1995) conducted a review of 29 studies on the effectiveness of shared reading in preschool or kindergarten. They concluded that individual differences in preschool exposure to shared reading explained approximately 8 percent of the variance in later formal reading outcomes studied. In a more recent meta-analysis, Mol et al. (2009) reviewed 31 experiments ($n = 2,049$ children) in which educators were trained to encourage children to be actively involved before, during, and after joint book reading. Although teaching print-related skills was not part of interactive reading programs, a substantial portion of the variance in kindergarten children's alphabetic knowledge could be attributed to the intervention.

1.2. Computer-supported early literacy interventions

Although there is a wide-spread belief that computers can support the early literacy development of children, the number of meta-analyses on the effects of computer-supported early literacy interventions is small. The meta-analysis by Blok, Oostdam, Otter, and Overmaat (2002) was the first with a focus on computer-assisted instruction in support of literacy throughout primary education. In the analysis, an overall positive effect over 42 different experimental comparisons was evidenced. Effect sizes were higher when the experimental group displayed an advantage at pretest and when the language of instruction was English. More recently, the effects of interactive storybooks and other contextualized methods were examined in a systematic way. In a meta-analysis on the effects of interactive storybook reading studies, Mol et al. (2009) evidenced a moderate effect size for oral language skills, along with an effect on kindergarten children's alphabetic knowledge but the role of technology-enhanced storybook reading was not emphasized in this analysis. It was concluded that the interaction with electronic storybooks may not only help children to learn that print carries meaning, but that it may also foster phonological awareness and literacy skills (see De Jong & Bus, 2002; Korat, 2009).

It is important to note that the methodological rigor of early literacy intervention studies in the literature has been challenged. In intervention studies, methodological flaws can be observed related to nonrandom assignment of participants to conditions, failure to control for Hawthorne effects by providing alternate interventions to control groups, insufficient or nonexistent assurance of fidelity of treatment, poor measurement sensitivity and inadequately described samples (see Troia, 1999). In a meta-analysis on the effectiveness of information and communication technology on the learning of written English, starting at age five (Andrews, Freeman, McGuinn, Robinson, & Zhu, 2007), only nine studies with rigorous design were selected showing mixed results. As regards the role of technology in interventions for young children, it was concluded that there is so much variation in the design of the interventions that no conclusions about outcomes could be made. Examining the effectiveness of technology use in classrooms from previous meta-analyses, Archer et al. (2014) concluded that only a small number of studies met most of the required criteria for adequate study design, that program focus and program outcome measures have not been clearly specified, and that most studies largely neglected fidelity of implementation.

Taken together, there is not yet clear research evidence on the effects of computer-based early literacy interventions in preschool and kindergarten. In the studies conducted so far, the scope of technology-enhanced interventions and their hypothesized effects on different outcome measures across different languages and writing systems have not always been clear. As a consequence, it is by no means clear to what extent interventions focusing on phonological awareness, letter knowledge and storybook reading lead to differential learning outcomes on phonological awareness and reading-related measures. It is also not clear what the impact is of study characteristics, related to the intervention or the child, or to the orthographic depth of the language being involved. Therefore, the present meta-analysis aimed to investigate the effects of computer-based early literacy training taking into account the focus of treatment and outcome measures and the role of study characteristics.

1.3. The present meta-analysis

The focus of the present meta-analysis was on the effects of computer-supported early literacy interventions in preschool and kindergarten on children's phonological awareness and reading-related skills in alphabetic languages, published during and after 1995. The year of 1995 was considered as a cut point since it was by that time that computers with higher quality software became available to a large proportion of schools (Zawacki-Richter & Latchem, 2018). In the present meta-analysis, only studies were selected that reported empirical research with comparisons between experimental and control conditions (both randomized trials and quasi-experimental studies) in which each condition had at least nine participants and that included a pretest and a posttest on the same outcome measure for both conditions. Three types of treatments (strict phonological awareness training, combined phonological awareness and letter training, and storybook reading using e-books) and two types of outcome measures, i.e., phonological awareness (syllabic awareness, word blending, rhyme, phoneme awareness) and reading-related skills (concept about print, letter knowledge, decoding, spelling) were distinguished. Although in some early literacy models language is seen as constituent part, in our analysis specific treatments on language abilities (comprehension, vocabulary) were left out for reasons of parsimony.

In addition to focus of treatment, a number of relevant study characteristics were taken into account. To begin with, we were interested in the effect of the treatment being part of an integrated learning system. In an ILS, the educational technology materials are fully integrated into the reading curriculum which makes the use of educational technology consistent with the reading curriculum of the kindergarten classroom (Cassady & Smith, 2003), and adaptive to the educational needs of the children (Becker & Hativa, 1994). Next to the classification of a study as ILS, the effect of treatment duration in total period of time and intensity of treatment was examined. Two participant characteristics were also considered: age and at-risk status. Next, the role of orthographic transparency was highlighted by making a distinction between language with high and low orthographic transparency. Previous research has shown that the addition of letters to phonological awareness training is more facilitative in transparent orthographies (Landerl et al., 2019). Furthermore, the effect of research design characteristics was investigated as regards the randomization of samples, the presence of an active control group, and the use of follow-up (retention) measures. Finally, the role of criterion variables was further explored by taking into account whether the measure required letter knowledge, whether speed was involved, whether it involved transfer, and whether there was a match between the criterion and target variable.

In the present study, an attempt was made to find an answer to the following questions:

1. How large were the effects of computer-assisted early literacy treatments in preschool and kindergarten across different types of treatments (strict phonological awareness training, combined phonological awareness and letter training, and storybook reading), and different types of outcome measures (phonological-awareness and reading-related skills)?
2. To what extent do study characteristics (type of treatment, integrated learning system or not, treatment duration, participants, orthographic transparency, research design, criterion variables) explain the differences in effect sizes?

To answer these questions two sets of multilevel analyses were performed. The first set included an overall analysis using all available data (all treatments and all outcomes) plus analyses for each of the three types of treatments, across all outcomes; the second set included separate analyses for the two types of outcomes measures, across all treatment types.

2. Method

2.1. Selection of studies

Studies published in the years between 1995 and 2017 were selected in three steps. In the first step, a keyword search was conducted in the Psycinfo and ERIC databases, using the following keyword combination: (multimedia OR technology OR computer) AND (reading OR literacy OR spelling OR dyslexia OR "specific reading disorder") AND (instruction OR intervention OR treatment). Only empirical studies published in the English language in a peer-reviewed journal were eligible for the meta-analysis. To be included, studies should focus on the effectiveness of computer-supported early literacy interventions (phonological awareness, letter knowledge, storybook reading) at the preschool and kindergarten level. Effectiveness of interventions was assessed in terms of phonological awareness and reading-related measures (e.g. letter knowledge and word decoding, not reading comprehension). The use of modern media technology in classrooms started to rise in the early nineties. Therefore, we started our meta-analysis with studies published in 1995. The restriction to articles in peer-reviewed journals was based on the idea that such articles satisfy minimum standards of methodological quality.

In the second step, some of the most important education-related peer-reviewed journals in the field of literacy research were browsed in order to ensure that all major publications in the field would be included. The following ten journals were searched between 1995 and 2017: *Annals of Dyslexia*, *British Journal of Educational Psychology*, *Journal of Educational Computing Research*, *Journal of Educational Psychology*, *Journal of Computer Assisted Learning*, *Journal of Research in Reading*, *Reading Psychology*, *Reading Research Quarterly*, *Reading & Writing*, and *Scientific Studies of Reading*.

In the third step, the references of identified articles and references in previously published meta-analyses (Blok et al., 2002; Bus & van IJzendoorn, 1999; National Reading Panel, 2000; Ehri et al., 2001; Mol et al., 2009; Troia, 1999) were used to check our selection and to further identify relevant studies.

In addition to this stepwise selection procedure, the following criteria were used to include studies in our data base:

Table 1
Characteristics of the studies organized by type of treatment.

Treatments Focused on Phonological Awareness								
Study	Treatment	Hours	Language	Students	N	Comparisons	Outcomes	Average d (min, max)
Anthony (2016)	Phonological awareness (Earobics Step 1)	31.5	English (US)	Low-SES, ethnic-minority kindergarteners (5–6.7 years old)	247	Earobics vs. Building blocks (math)	Phonological awareness, graphophonemic knowledge, word reading, and vocabulary	0.05 (–0.10, 0.14) n = 7
Kartal and Terziyan (2016)	Phonological awareness (JerenAli)	3.33	Turkish	Low-SES kindergarteners (53–72 months)	20	JerenAli vs. control	Phonological awareness subtests and total score	0.39 (–0.36, 1.09) n = 14
Kartal et al. (2016)	Phonological awareness (Phonics-based CAI, JerenAli)	3.4	Turkish	Kindergarteners (4.5–6.2 years old)	53	Exp. vs. treated and untreated control at posttest and follow-up	Seven subtests of phonological awareness	0.35 (–0.70, 1.57) n = 28
Lonigan et al. (2003)	Phonological Awareness (DaisyQuest, Daisy's Castle)	5.46	English (US)	At risk (Head Start) preschoolers (44–64 months)	45	Exp. vs. control	8 phonological sensitivity measures	0.77 (–0.37, 2.03) n = 8
Macaruso and Rodman (2011)	Phonological Awareness (Phonics-based CAI, Lexia Early Reading)	8.0	English (US)	Preschoolers (5-year olds)	38	Lexia vs. control	GRADE-P: Phonological awareness, visual skills, concepts	0.35 (–0.30, 1.28) n = 10
Macaruso and Walker (2008)	Phonological Awareness (Phonics-based CAI and letter-sound correspondences (Lexia Early Reading)	19.0	English (US)	Kindergarteners (mean age 5.5 years)	71	Lexia vs. control for whole sample and for sub-sample of low performers	Letter naming fluency, phoneme segmentation fluency (DIBELS); phonological awareness, letters and letter-sound correspondences, concepts	0.30 (–0.17, 1.02) n = 10
Mitchell and Fox (2001)	Phonological Awareness (DaisyQuest + Daisy's Castle)	5.0	English (US)	At risk for reading failure (kindergarten and first grade)	72	DaisyQuest vs. two control conditions (teacher-delivered and computer control)	about print (Gates-MacGinitie Reading Test, Level PR)	0.15 (–0.28, 0.92) n = 10
O'Callaghan, McIvor, A., McVeigh, C., & Rushe (2016)	Phonological Awareness (Phonics-based CAI, Lexia)	10.79	English (UK)	Below-average scoring kindergarteners (4–6 years old)	98	Lexia vs. control	Blending words, phoneme segmentation, non-word reading (PhAB-2)	0.36 (0.23, 0.47) n = 3
Reitsma and Wesseling (1998)	Phonological Awareness (CAI for blending letter sounds into words)	3.33	Dutch	Kindergarteners (62–83 months)	98	Exp. vs. active (vocabulary training) and untreated control.	Blending, segmentation, word reading, spelling.	0.81 (0.44, 1.33) n = 4
Van de Sande et al. (2016)	Phonological Awareness: CAI for early literacy skills with or without support of executive functioning	3.33	Dutch	Kindergarteners (4.79–6.77 years)	101	Follow-ups after 10 and 18 weeks in Grade 1. Computer games with or without executive functioning support vs. control	Syllable and phoneme segmentation, blending words, letter knowledge, and total score	0.11 (–0.28, 0.60) n = 20
Wild (2009)	Phonological Awareness: CAI for practising phonological awareness skills ('Rhyme and Analogy' program)	4.0	English (UK)	Year 1 children (5–6 years old)	127	Two posttests Phonological awareness training with CAI vs. (1) the same training without CAI, and (2) computer math games	Phonological Assessment Battery, Marie Clay Dictation	0.40 (0.29, 0.55) n = 10
Treatments Focused on Letter Knowledge and Phonology								
Study	Treatment	Hours	Language	Students	N	Comparisons	Outcomes	Average d (min, max)

(continued on next page)

Table 1 (continued)

Study	Treatment	Hours	Language	Students	N	Comparisons	Outcomes	Average d (min, max)
Bauserman, Cassidy, Smith, and Stroud (2005)	Letters and Phonology: PLATO, Beginning Reading for the Real World, Level A	5.5	English (US)	Low-SES kindergarteners (5–6 years old)	88	Exp. school vs. control school	Concepts about Print, Phonological Awareness	0.45 (0.43, 0.47) n = 2
Both-de Vries and Bus (2014)	Letters: Illustrations in electronic alphabet books	0.25	Dutch	Kindergarteners (4–5 years old)	44	Alphabet book vs. control	Letter knowledge	0.31 n = 1
Brabham, Murray, and Bowden (2006)	Letters and Phonology: Practicing with alphabet books on CD-ROM	?	English (US)	Kindergarteners (age 5; 1 to 6; 8)	152	CD-ROM vs. Audio, for two instructional emphases (meanings or phonemes)	Vocabulary, letter names, phonetic cue reading, and phoneme identities	-0.21 (-0.92, 0.29) n = 8
Cassady and Smith (2003)	Letters and Phonology (Waterford Early Reading Program)	30.0	English (US)	Kindergarteners (5–6 years old)	88	Exp. school vs. control school	Concepts about Print and Phonological Awareness Test	0.20 (-0.43, 0.83) n = 2
Cviko, McKenney, and Voogt (2012)	Letters: Emergent literacy (PictoPal)	1.67	Dutch	Kindergarteners (4–6 years old)	168	PictoPal vs. control	Emergent Literacy Test	0.45 n = 1
Davidson et al. (2009)	Letters and Phonology: Emergent literacy (Ready, Set, Leap! curriculum)	150.0	English (US)	Low-SES Preschoolers (4–5 years old)	254	Exp. vs. control	Phonological awareness, letter naming, letter identification, picture vocabulary	0.14 (-0.10, 0.35) n = 6
De Graaff, Bosman, Hasselman, and Verhoeven (2009)	Letters and Phonology: Systematic phonics instruction	3.75	Dutch	Kindergarteners (5–7 years old)	93	Systematic and unsystematic phonics instruction vs. control	Letter knowledge, phonemic awareness, word reading and spelling	0.41 (-0.09, 0.70) n = 8
Hecht and Close (2002)	Letters and Phonology (Waterford Early Reading Program)	21.0	English (US)	Low SES kindergarteners (mean age 67 months)	76	Waterford Early Reading Program vs. control school	Segmenting and blending (CTOPP), letter name knowledge, letter sound, letter writing, word reading, invented spelling, vocabulary, CAP	0.47 (-0.33, 2.03) n = 9
Huffstetter, King, Onwuegbuzie, Schneider, and Powell-Smith (2010)	Letters and Phonology (Headsprout Early Reading)	20.0	English (US)	Preschoolers at risk for reading failure (Headstart, 4–5 years old)	62	Headsprout vs. control	Early reading skills (TERA-3)	1.00 n = 1
Johnson, Perry, and Shamir (2010)	Letters and Phonology: Three presentation methods for CAI (learner-controlled menu, linear sequencer, adaptive sequencer)	8.67	English (US)	Preschoolers and kindergarteners	183	Three experimental conditions each versus control	Phonological awareness and early reading skills (DIBELS)	0.21 (0.01, 0.41) n = 3
Kegel, van der Kooy-Hofland, and Bus (2009)	Letters and Phonology (Living Letters)	2.5	Dutch	Low-performing kindergarteners (5 years old)	90	Living Letters vs. Living Books	Phoneme recognition, word dictation, and an aggregate score	0.17 (-0.01, 0.38) n = 3
Kegel, Bus, and Van Lijendoorn (2011)	Letters and Phonology (Phoneme skills, Living Letters)	2.5	Dutch	Kindergarteners (48–63 months)	182	Living Letters with or without feedback vs. control	Early Literacy Skills	0.43 (0.01, 1.04) n = 4
Kegel and Bus (2012)	Letters and Phonology (Living Letters)	2.5	Dutch	Low-SES kindergarteners (4 years old)	312	Living Letters vs. control for carriers and non-carriers of the DRD4 7-repeat allele	Letter knowledge, phonemic sensitivity, developmental spelling, and an aggregate score	0.14 (-0.04, 0.38) n = 8
Kreskey and Truscott (2016)	Letters and Phonology: Headsprout Early Reading	11.37	English (US)	Kindergarteners at risk for reading failure (mean age 5.4)	102	Headsprout vs. control	Letter naming fluency (DIBELS)	-0.46 n = 1

(continued on next page)

Table 1 (continued)

Study	Treatment	Hours	Language	Students	N	Comparisons	Outcomes	Average d (min, max)
Lovio, Halttunen, Lyytinen, Näätänen, and Kujala (2012)	Letters and Phonology: GraphoGame	3.0	Finnish	Kindergarteners with reading difficulties (mean age 6 yrs)	29	Exp. vs. control	Phonological processing, letter knowledge and recognition, reading and writing syllables and non-words; word reading.	0.26 (0.00, 0.46) n = 6
Macaruso and Rodman (2011) Study 2	Letters and Phonology: (Lexia Early Reading + Primary Reading)	18.75	English (US)	Low-performing kindergarteners (mean age 5.5 years)	66	Lexia vs. control	GRADE-K: phonological awareness, early literacy skills	0.29 (-0.27, 1.12) n = 10
McKenney & Voogt, 2009, Study 3	Letters: CAI for emergent literacy (PictoPal)	2.67	Dutch	Kindergarteners (4- and 5-year olds)	79	PictoPal vs. treated control	Early Literacy Skills Test	0.41 n = 1
McKenney & Voogt, 2009, Study 4	Letters: CAI for emergent literacy (PictoPal)	2.67	Dutch	Kindergarteners (4- and 5-year olds)	37	PictoPal vs. treated control	Early Literacy Skills Test	-0.52 n = 1
Mioduser, Tur-Kaspa, and Leitner (2000)	Letters and Phonology: CAI for early reading skills	?	Hebrew	Kindergarteners at high risk for learning disabilities (5-6 years old)	46	One experimental condition vs. treated and untreated control	Phonological awareness, word recognition, letter recognition	0.66 (0.40, 0.91) n = 6
Piquette, Savage, and Abrami (2014)	Letters and Phonology: Web-based early literacy program (ABRACADABRA)	20.0	English (Canada)	Kindergarteners and first-graders	203	ABRA vs. control	Blending words, letter-sound knowledge, Fry words, word reading (GRADE)	0.32 (-0.40, 0.65) n = 4
Plak, Kegel, and Bus (2015)	Letters: Alphabetic knowledge (Living Letters) and text comprehension (Living Books)	3.875	Dutch	Low-performing kindergarteners (5-year olds)	257	Living Letters & Living Books vs. control for carriers and non-carriers of the DRD4 7-repeat allele	Cito Literacy Test	0.32 (-0.09, 0.62) n = 6
Plak, Merkelbach, Kegel, van Ljzendoorn, and Bus (2016)	Letters: Alphabetic knowledge (Living Letters) and text comprehension (Living Books)	3.0	Dutch	Literacy-delayed kindergarteners (5-year olds)	307	Living Letters & Living Books vs. control for carriers and non-carriers of the DRD4 7-repeat allele	Cito Literacy Test	0.20 (0.01, 0.74) n = 6
Regvoort and van der Leij (2007)	Letters: CAI for phonemic awareness and letter-sound relationships	12.0	Dutch	Kindergarteners of dyslexic parents (mean age 5.81 years)	73	Exp. vs. control at risk, and Exp. vs. control not at risk Follow-ups after 5, 10, and 20 months (Grade 1 and 2).	Phonemic awareness, letter knowledge, naming speed, reading fluency (words and pseudowords), spelling words	0.67 (0.14, 1.18) n = 4
Savage et al. (2013)	Letters and Phonology: Web-based early literacy program (ABRACADABRA)	20.0	English (Canada)	Kindergarteners, first and second graders (mean age 73.69 months)	74 ^a	ABRA vs. control	Letter-sound knowledge, word reading, blending words (CITOPP), phonemic segmentation, reading fluency and reading retell (DIBELS), sentence and passage comprehension (GRADE)	0.10 (-0.34, 0.42) n = 5
Segers and Verhoeven (2005)	Letters and Phonology: Discovery games and learning games	8.5	Dutch	Native Dutch kindergarteners (mean age 5; 6)	38	Exp. vs. control	Rhyming, phonemic segmentation, auditory blending, grapheme knowledge	-0.03 (-0.24, 0.21) n = 4
Tracey and Young (2007)	Letters and Phonology: Emergent literacy skills (Waterford Early Reading Program, Level 1)	44.0	English (US)	Kindergarteners at risk for reading failure	265	Exp. classrooms vs. control classrooms	TERRA-2, Waterford Inventory, Lindamood Auditory	1.12 (0.07, 2.10) n = 3
Van Daal and Reitsma (2000) Study 1	Letters and Phonology: Pre-reading skills, letter knowledge, decoding, spelling (Circus of Reading)	3-5	Dutch	Reading-ready kindergarteners (5.5-6.5 years)	21	Exp. vs. control	Conceptualization Test	0.38 (0.33, 0.43) n = 2

(continued on next page)

Table 1 (continued)

Treatments Focused on Phonological Awareness

Study	Treatment	Hours	Language	Students	N	Comparisons	Outcomes	Average d (min, max)
Van der Kooy-Hofland, Bus, and Roskos (2012)	Letters and Phonology (Phoneme skills, Living Letters)	3.0	Dutch	Kindergarteners at risk for reading failure (5–6 years)	98	Living Letters vs. Living Books	Letter knowledge, phonological skills, invented spelling, word recognition, decoding, word reading fluency, pseudoword reading fluency, spelling, reading comprehension	0.32 (0.06, 0.55) n = 5
Wolgemuth et al. (2013)	Letters and Phonology: Web-based early literacy program (ABRACADABRA)	34.67	English (Australia)	Kindergarteners, first and second graders	308	ABRA vs. control	Phonological awareness, word reading (GRADE-K and PIDS-BLA)	0.19 (0.10, 0.32) n = 5
Storybook Reading Study								
Chera and Wood (2003)	Storybooks: phonological awareness (e-book)	1.67	English (UK)	Preschoolers and kindergarteners (3–6 years old)	30	Exp. vs. control Follow-up after 2 weeks	Phonological awareness (6 tasks), BAS word reading	0.35 (–0.24, 0.95) n = 14
De Jong and Bus (2002)	Storybooks: Emergent literacy (E-book with or without games)	1.5	Dutch	Kindergarteners (4–6 years old)	48	E-book with or without games vs. reading book by adult, and vs. untreated control	Accuracy of word recognition with or without icons	–0.32 (–1.99, 1.06) n = 8
Gong and Levy (2009)	Storybooks: Print concepts and early reading (Storybook reading)	1.25	English (Canada)	Preschoolers (4–5 years old)	96	Three exp. conditions (ways to draw attention to print) vs one active control	Reading letters (WRAT-3)	0.25 (0.07, 0.42) n = 3
Homer et al. (2014)	Storybooks: Digital storybook (Kinect) and interactive storybook (with game activities)	0.25	English (US)	Kindergarteners and first-graders (5–7 years old)	39	Kinect storybook with or without games vs. book reading with an adult	Reading vocabulary, definitions.	0.02 (–0.19, 0.17) n = 10
Ihmeideh (2014)	Storybooks: Emergent literacy (E-book)	10.0	Arabic	Kindergarteners (age 4; 7 to 5; 2)	92	E-book vs. printed book	Phonological awareness, print awareness, alphabetic knowledge, and vocabulary	1.13 (–0.22, 2.73) n = 5
Korat (2009)	Storybooks: Three or five readings of an e-book	1.5	Hebrew	Low-SES preschoolers (4–5 years old) and kindergarteners (5–6 years old)	214	Three or five e-book readings vs. control, separately for preschool and kindergarten	Phonological awareness, word reading, word meaning	0.32 (–0.58 – 1.21) n = 12
Korat (2010)	Storybooks: Word reading and oral language (E-book)	1.8	Hebrew	Kindergarteners (age 5; 2–6; 3)	90	E-book versus control for kindergarten only	Word reading, vocabulary	1.13 (0.47, 1.79) n = 2
Korat and Blau (2010)	Storybooks: Emergent literacy (E-book with three or five readings)	1	Hebrew	Preschoolers and kindergarteners (4–6 years old)	247	Two versions of e-book vs. control, separately for preschool and kindergarten, and separately for low and middle SES	Phonological awareness, word reading, word meaning	0.38 (–0.38, 1.68) n = 24
Korat and Shamir (2007)	Storybooks: Emergent literacy (E-book)	1.5	Hebrew	Kindergarteners (5–6 years old)	128	E-book vs. reading book by adult, E-book vs. untreated control	Phonological awareness, word recognition, vocabulary	0.26 (–0.02, 1.06) n = 6
Korat and Shamir (2008)	Storybooks: Supporting emergent literacy (E-book reading)	1.5	Hebrew	Low and Middle SES kindergarteners (5–6 years old)	149	Three e-book modes (read story only, read with dictionary, read and play) vs. control	Word meaning, word recognition, phonological awareness	0.23 (–0.18, 1.44) n = 9

(continued on next page)

Table 1 (continued)

Treatments Focused on Phonological Awareness								
Study	Treatment	Hours	Language	Students	N	Comparisons	Outcomes	Average d (min, max)
Korat, Shamir, and Heibal (2013)	Storybooks: Shared book reading (E-book)	1.67	Hebrew	Low-SES preschoolers (4.6–5 yrs old)	90	E-book vs. printed book, and e-book vs. control	Vocabulary, syllabic awareness, sub-syllabic awareness	0.49 (0.09, 0.93) n = 6
Moody, Skribbe, Patker, and Jones (2014)	Storybooks: Storybook reading with print referencing	3.0	English (US)	Preschoolers (4–5 years old)	20	E-book vs. control	Concept of print, letter knowledge, early decoding	0.52 (0.06, 1.38) n = 3
Segal-Drori, Korat, Shamir, and Klein (2010)	Storybooks: Emergent reading (E-book)	1.0	Hebrew	Low-SES kindergarteners (5–6 years old)	128	Two e-book conditions (with or without adult instruction) each vs. printed book read by adult and untreated control	Phonological awareness, word reading, concept about print	0.69 (–0.38, 3.03) n = 12
Shamir (2009)	Storybooks: Emergent literacy (E-book)	1.75	Hebrew	Low-SES kindergarteners (5–6 years)	96	E-book vs. control	Vocabulary, syllabic and sub-syllabic phonological awareness	0.33 (0.09, 0.71) n = 3
Shamir, Korat, and Barbi (2008)	Storybooks: Emergent literacy (Paired-peer and individual use of e-book)	1.0	Hebrew	Low-SES kindergarteners (mean age 5.64 years)	110	Three e-book conditions (tutor, tutee, individual) vs. control	Phonological awareness, word recognition, concept about print	0.36 (–0.43, 1.26) n = 27
Shamir, Korat, and Fellah (2012)	Storybooks: Phonological awareness and vocabulary (E-book)	2.0	Hebrew	Kindergarteners at risk for learning disabilities (5–7 years)	110	E-book vs. active and untreated control condition	Phonological awareness, concept about print, oral vocabulary	0.71 (0.13, 1.93) n = 6
Shamir and Lifshitz (2013)	Storybooks: Phonological awareness (E-book)	2.0	Hebrew	Kindergarteners at risk for learning disabilities (4.5–7 years)	77	E-book with or without metacognitive guidance vs. control	Rhyming	–0.08 (–0.31, 0.15) n = 2
Shamir and Shlafer (2011)	Storybooks: Emergent literacy (E-book)	3.0	Hebrew	Kindergarteners (5–7 years old)	136	E-book. vs. control for children at risk and for typically developing children	Phonological awareness, concept about print	0.41 (0.09, 0.54) n = 4
Willoughby, Evans, and Nowak (2015)	Storybooks: Electronic alphabet books	5.33	English (Canada)	Preschoolers (3–4 years old)	92	Electronic alphabet book vs. paper alphabet book and vs. storybook	Letter name, letter sound, and same-sound	–0.13 (–0.31, 0.06) n = 6

^a Number of classrooms; effect sizes computed at the classroom level.

- The intervention should be a program developed for the purpose of enhancing children's early literacy and should not solely consist of the use of computers in general.
- The study must use both a pre-test and a control group; the control group received no-treatment or is an active control group receiving either a traditional early literacy treatment without computers or a non-literacy computer treatment. At pre-test the control group should be approximately equivalent to the experimental group (i.e., no major differences in age or socio-economic background).
- Case studies and studies with very small samples (less than 9 students per condition) were excluded.
- Studies written in English but may deal with other languages.
- Studies specifically dealing with bilingual readers or with second-language learners were excluded.
- Studies specifically dealing with students with specific language impairments or with other disabilities (e.g., autism, ADHD) were excluded; but studies on students at-risk for dyslexia were included.
- Available electronic or paper copy of the article.
- Availability of means and standard deviations or enough other information to calculate effect sizes.

These criteria were checked on the basis of the abstracts or, if necessary, on the basis of reading the whole article. In total 59 studies from 57 articles out of 32 scientific journals were identified for inclusion in the current meta-analysis, resulting in 339 effect sizes. In two articles, two studies were identified for inclusion and treated as separate studies. [Table 1](#) lists the included studies with their main characteristics.

Multiple effect sizes per study resulted from the use of multiple criterion variables and multiple contrasts between experimental conditions within a study. Contrasts used were contrasts between a treatment condition and a control condition. A study could have more than one treatment condition or more than one control condition. The 59 studies involved a total of 6786 students. The mean age of the students varied per study from 49.0 to 75.8 months, with an average of 65.03 months. The students were at preschool or kindergarten level; in most studies ($n = 36$) the children were on average between five and six years old, and in 12 studies between four and five; in six studies grade-1 students were also present. In 16 studies students at risk for reading problems were dealt with, two studies involved both regular students and students at risk. In 41 studies only regular students were involved. Student characteristics are shown for each study in [Table 1](#). The included articles mainly concerned students with English as their mother tongue (25 of the 59 articles). The other languages were Dutch ($n = 17$), Hebrew ($n = 13$), Turkish ($n = 2$), Arabic ($n = 1$), and Finnish ($n = 1$). The language is shown for each study in [Table 1](#).

The interventions studied in the articles could be classified into three categories: treatments focusing on phonology (11 studies with 104 effect sizes), on letter knowledge and phonology (28 studies with 103 effect sizes), and storybooks (20 studies with 132 effect sizes). [Table 1](#) has been organized by these three types of treatment and holds short descriptions of the treatments.

2.2. Coding of studies

All effects were individually coded on program, outcome variable and design factors. The coded program variables included type of intervention (phonological awareness, letter knowledge and phonological awareness, storybook reading), use of integrated learning system (ILS) or not, treatment duration (program length in number of weeks, intensity of treatment in minutes per week). The coded outcome variables were characterized by dummy variables (letter knowledge required or not, speeded or non-speeded test, test of trained or untrained items (transfer), does the criterion variable match the skills taught in the treatment or not). Finally, the coded study design factors involved participant characteristics (average age of participants in months, students at risk for reading problems or not), orthographic transparency (transparent or not), and research design characteristics (random allocation to conditions or not, active control condition or not, follow-up measurement or immediate posttest).

ILS was considered to be an important program characteristic. A dummy variable was constructed to indicate whether or not a program was considered to be an ILS. Non-ILS interventions included interventions specifically directed at phonological skills and e-book reading. In an ILS, the educational technology materials are fully integrated into the reading curriculum with instructional activities for each student based on recorded achievements of the student on previously assigned activities. In only seven of the 59 studies the treatment was identified as an ILS, for instance the Waterford Early Reading Program ([Cassady & Smith, 2003](#); [Hecht & Close, 2002](#)). The ABRA web-based literacy system ([Savage et al., 2013](#)) was also treated as an ILS. With ABRA, though, the teacher and not the computer program selects activities appropriate for the student.

The treatments lasted on average for 10.3 weeks, but there was a broad variation in treatment duration from 1 to 35 weeks. The intensity of treatment was on average 50.9 min per week, varying from 3.5 to 160 min and with one outlying value of 300 min. Treatment duration is shown for each study in [Table 1](#) as the average hours of treatment. Quality and quantity of treatment has been described in most studies only as intended not as actually delivered. We found only limited attention for fidelity of implementation (exceptions included [Davidson, Fields, & Yang, 2009](#); [Savage et al., 2013](#); [Anthony, 2016](#)).

The languages were categorized according to orthographic depth. One dummy variable was constructed: relatively transparent orthographies (33 studies, mainly Dutch and Hebrew) versus relatively opaque orthographies (26 studies, English).

Research design characteristics varied not only between studies but also between comparisons within studies. Random assignment to conditions was present in 47 studies (269 contrasts). So, in almost 80% of the cases some form of randomization was used. We classified as random both random assignment of individual students and random assignment of whole classrooms to experimental conditions. In about 20% of the cases with random assignment the conditions were assigned at the classroom level. In only three studies matching was used, which was counted as non-random. In about 50% of the contrasts the control condition received no

special treatment; in these cases, the regular classroom teaching was used as control. In the 169 comparisons with an active control condition, the alternative treatment for the controls was in 59% of the cases a reading treatment. In 52% of the cases the active control involved the use of a computer. For type of control condition, we used only one dummy variable in the models: active control versus regular classroom teaching used as control. Table 1 summarizes the contrasts used in each study. As final research design variable the post-treatment measures were divided into immediate posttests and follow-up tests. A dummy variable was created to indicate whether a post-test was a follow-up test or an immediate posttest. Only 29 follow-up tests were identified, which amounts to only 8.6% of the contrasts. Moreover, follow-up measurements occurred in only three studies (Chera & Wood, 2003; Kartal, Babür, & Erçetin, 2016; Van de Sande, Segers, & Verhoeven, 2016). The follow-ups that occurred were mostly timed not far from the first posttest. In 14 of the 29 cases the follow-up was five months after the first posttest. In the other cases the follow-up was after one month or after only two weeks. It seems that not much attention was given to long-term effects of computer-assisted instruction.

The criterion variables (see brief descriptions in Table 1) were classified into nine types of outcomes (see Table 3). These variables fell into two broad categories: phonological-awareness measures and reading-related measures. The first set of variables included syllabic or sub-syllabic awareness, word blending, rhyme, phoneme awareness and global phonological awareness. This last category included tests concerning various phonological skills. The phonological variables applied to 184 of the 339 contrasts. The reading-related variables included concept about print, tests of early literacy and letter knowledge, word decoding and spelling accuracy. This applied to 155 contrasts.

The criterion variables were further characterized using some dummy variables. A dummy variable indicated whether or not letter knowledge was required for the test (yes: 133, no: 206). This is almost equal to the distinction between phonological measures and reading-related measures. The only difference concerns the tests of concept of print; these do not require letter knowledge but were considered as reading-related tests. The next dummy variable indicated whether the test was speeded or not. In very few cases ($n = 10$) the test was speeded. Furthermore, a distinction was made between tests with trained or untrained items. The latter was interpreted as transfer and this occurred in 71.7% of the contrasts. The last dummy variable indicated whether or not the criterion variable matched the target level of the treatment. In 83.2% of the contrasts a match was observed. For instance, when the treatment was targeted at phonemic awareness and the criterion variable measured phonemic awareness then the corresponding contrast was coded as a match. But when in such case the criterion variable was word reading then there was no match. Some studies reported total scores of tests as well as subtest scores; in such cases only the subtest scores were included in the analyses and not their total score.

All studies were coded by one of the authors. Inter-coder reliability was initially assessed for two coders on a sample of 10 studies with 65 contrasts between a treatment condition and a control condition on a single criterion variable. For all categorical variables used in the analyses the proportion agreement between coders was between 80 and 100%. For the treatment duration variables, the correlations between coders were 0.86 and 0.91. Disagreements between coders were discussed and occasionally the coding scheme was adapted.

2.3. Calculation of effect sizes

In order to make the results from different studies comparable a standardized measure of effect size was calculated using the following formula:

$$(1) \text{ Effect size } g = (\text{Gain Exp} - \text{Gain Control}) / \text{pooled standard deviation of pre-test means.}$$

Because the effect size g has a small upward bias, it was transformed in the unbiased effect size d using formula 2 (Hedges, 1982).

$$(2) d = (1 - 3/4N - 9)g \text{ where } (N = \text{total sample size of the study})$$

The distribution of the effect sizes has been summarized per study in Table 1.

To be included studies needed to have a pretest as well as a posttest. Nevertheless, we included in the data some contrasts with a missing pretest. This happened when the pretest was thought too difficult given the age of the children. In such cases the pretest was fixed at a value of zero in computing the effect size. In such studies other pretest measures were available to check the equivalence of the experimental conditions.

2.4. Multilevel analysis

A multilevel approach was used to analyse the data, distinguishing variance between studies and variance within studies (Borenstein, Hedges, Higgins, & Rothstein, 2009; Hox, 2010). This approach allows for the inclusion of multiple effect sizes from the same study. In the present dataset, we had variation between effect sizes within a single study because of the use of multiple outcome measures, the use of multiple experimental groups, and the use of multiple control groups. In the multilevel approach the data form a hierarchical structure with the studies at the top level (level 3) and the effect sizes within studies at level 2. In addition, at level 1 the standard error of effect sizes was included as a predictor in the random part with a coefficient fixed equal to 1 (Hox, 2010; Marsh, Bornmann, Metz, Daniel, & O'Mara, 2009). The sampling variance of d was assumed to be known from statistical theory and was calculated using the following formula (see Borenstein et al., 2009):

Table 2
Distribution of Effect Sizes for Code-Related Skills per Type of Treatment (based on 59 studies).

Type of Treatment	Mean	S.D.	Minimum	Maximum	N
Phonology	0.33	0.491	-0.70	2.03	104
Letter knowledge	0.29	0.472	-0.92	2.10	103
Storybook reading	0.25	0.458	-0.80	2.10	132
Total	0.29	0.473	-0.92	2.10	339

Note: Vocabulary measures and total scores excluded. Extreme values winsorized.

$$(3) \text{ Sampling Variance} = (n_{\text{exp}} + n_{\text{control}})/(n_{\text{exp}} n_{\text{control}}) + d^2/2(n_{\text{exp}} + n_{\text{control}})$$

The standard error of the effect size was calculated by taking the square root of the sampling variance.

The variance at level 1 was constrained; there was only variance to be estimated at levels 2 and 3. In this way, it became possible to estimate the variance of the effect sizes at two levels, between and within studies, and the effect of explanatory variables could be estimated separately for each level to explain differences between studies and differences between comparisons within studies. The explanatory variables included variables defined at the study level (characteristics of the studies) and variables that characterized the contrasts within studies.

MLwiN 2.36 (Rasbash, Charlton, Browne, Healy, & Cameron, 2016) was used applying Restricted Maximum Likelihood (RML) to estimate the variances in the model. RML rather than ML was used because of the fact that the sample sizes were relatively small (Hox, 2010, p. 215).

3. Results

3.1. Descriptives of effect sizes

First, the descriptive statistics, means and standard deviations of the effect sizes are presented, as well as schematic graphs of the distributions. Second, the results of the multilevel analyses are shown to answer the research questions.

Table 2 shows the descriptive statistics of the effect sizes for code-related skills per type of treatment. The average effect size was 0.29. Treatments focused on phonological awareness had on average a somewhat higher effect size and storybook treatment a somewhat lower effect size, but mean differences were small and not significant. Standard deviations were all relatively high. In Fig. 1, a boxplot is given of the distribution of effect sizes for the three types of treatment. It is shown that there were some outliers in effect sizes for each type of treatment, especially for storybook treatments. Six extreme outliers, identified by graphical inspection of the effect sizes, have been winsorized so that the largest effect size became equal to 2.10. For phonology treatments the outliers were all at the positive side. The large negative effect sizes were mainly associated with comparisons of experimental treatments with active controls (specific reading treatments). After disregarding such comparisons, the average effect size increased from 0.29 to 0.34 (phonology 0.42, letter knowledge and phonology 0.36, storybook reading 0.28) for in total 240 comparisons.

In Table 3, the effect sizes for code-related skills per type of outcome are presented and Fig. 2 gives a boxplot of the distributions of effect sizes for the types of outcome. The mean effect sizes were moderately positive for all types of outcome measures, but for word decoding a rather low average effect size appeared.

In Table 4, the average effect sizes are presented as a function of type of treatment on the one hand and type of outcome on the other hand. Note that the n's per cell may be very low here. For instance, the high value of 0.88 for global phonological awareness was based on only one comparison. For phonological awareness treatments also concept about print, word decoding and spelling accuracy were only measured in a very few cases. Phonological awareness treatments resulted on average in moderately positive effect sizes on all types of criterion variables. For treatments focusing on both letters and phonology the effect sizes were also moderately positive,

Table 3
Distribution of Effect Sizes per Type of Outcome Variable (based on 59 studies).

Type of Variable	Mean	S.D.	Minimum	Maximum	N
<i>Phonological Awareness Measures</i>					
(Sub-)Syllabic awareness	0.27	0.462	-0.28	1.41	23
Word blending	0.38	0.362	-0.23	1.33	24
Rhyme	0.26	0.458	-0.57	1.81	36
Phoneme awareness	0.32	0.592	-0.92	2.03	54
Global phonological awareness	0.29	0.304	-0.38	0.91	47
<i>Reading-Related Measures</i>					
Concept about print	0.33	0.512	-0.52	1.46	22
Early literacy and letter knowledge	0.34	0.514	-0.46	2.10	58
Word Decoding	0.15	0.481	-0.80	2.10	63
Spelling accuracy	0.36	0.290	-0.03	1.11	12
Total	0.29	0.473	-0.92	2.10	339

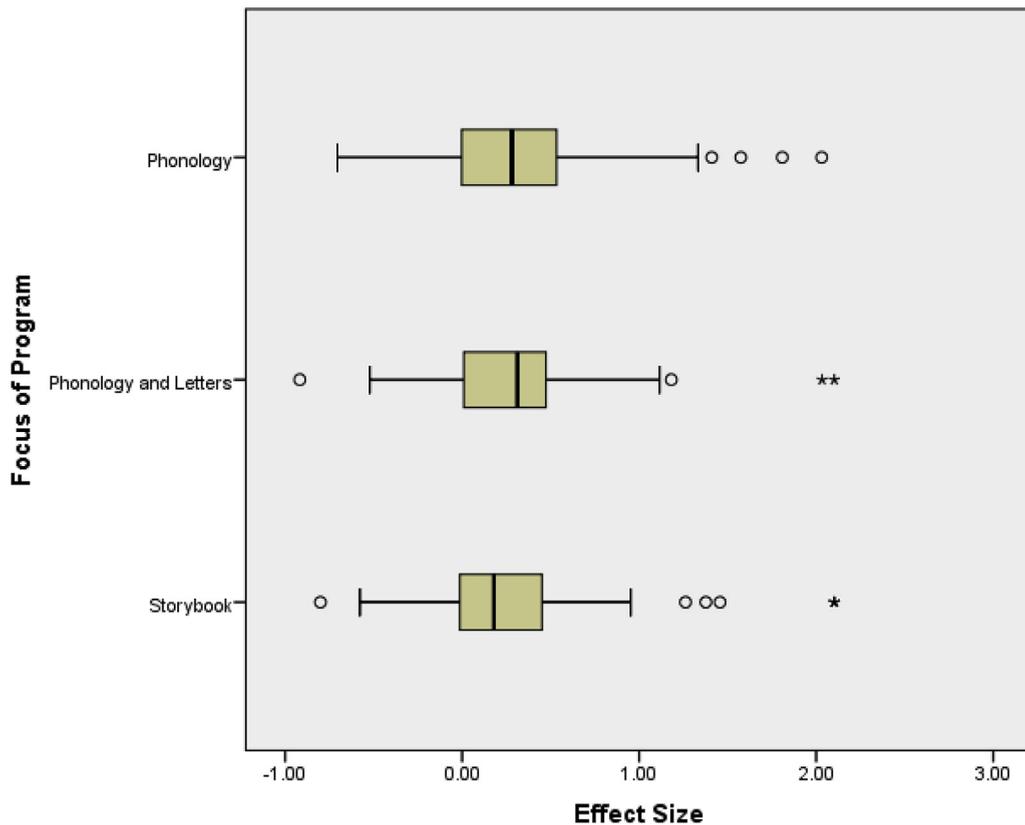


Fig. 1. Boxplot of the distribution of effect sizes per type of treatment.

except for rhyme and concept about print. The highest averages appeared for global phonological awareness, spelling accuracy, and early literacy and letter knowledge. Finally, for storybook reading the effects were generally moderate except for word decoding which showed only a low average. The clearest positive effect of storybook reading appeared on concept about print.

In general, moderately positive effect sizes have been observed with broad variability, ranging from negative to highly positive values. On average, differences between the three types of treatment appeared rather small and not significant. The same was true for the different types of outcome variables, except for word decoding because of its low average effect size. In studies about phonology awareness treatments, the outcome variables mostly were phonology-based but also reading-related variables were used (in 17% of the cases). The average difference in effect size between these two types of outcomes appeared small (0.33 for phonology and 0.30 for reading-related). In studies on treatments with a combined focus on phonology and letter knowledge more reading-related outcomes were assessed (in 65% of the cases). Average effect sizes for these two types of outcome were both equal to 0.29. Finally, in studies on electronic storybooks the two types of outcome appeared about equally frequent, with an average effect size for phonology outcomes equal to 0.27 and for reading-related outcomes equal to 0.23.

3.2. Association between effect size differences and study characteristics

Multilevel analyses were performed to answer the two research questions for different types of treatments and different types of outcome measures. The associations of effect sizes with various subsets of predictors were tested. Using all predictors simultaneously would have resulted in a too large number of predictors given the number of studies ($m = 59$). The analyses were done for the whole set of effect sizes ($n = 339$) as well as separately for the three types of treatments and the two types of outcome variables.

First, the overall analysis and the analyses by treatment type are presented. Tables 5a and 5b shows the results of the multilevel analyses for seven subsets of predictors: focus of criterion variable (the two types of criterion variables), ILS, treatment duration, participant characteristics, language (transparent or not), characteristics of the research design, and of the criterion variables. Table 5a holds the fixed-effects parameter estimates (the regression coefficients) and their standard errors (within parentheses). Statistically significant coefficients ($p < .05$) are in boldface type. Table 5b holds the random part of the models, the between- and within-studies variance components. In the column Total the results are given for the analyses across the total set of effect sizes.

The tables start with the results for the empty model, the model without any predictors. The empty model provides answers to the first research question, overall in the column Total and by treatment type in the other columns. In this model the average effect size across studies (intercept) is estimated. Overall, across all treatments and all comparisons, the average effect size was a small positive number, 0.28 with 95% CI (0.21, 0.35). In all cases the average effect size was statistically significant (compare the estimate with its

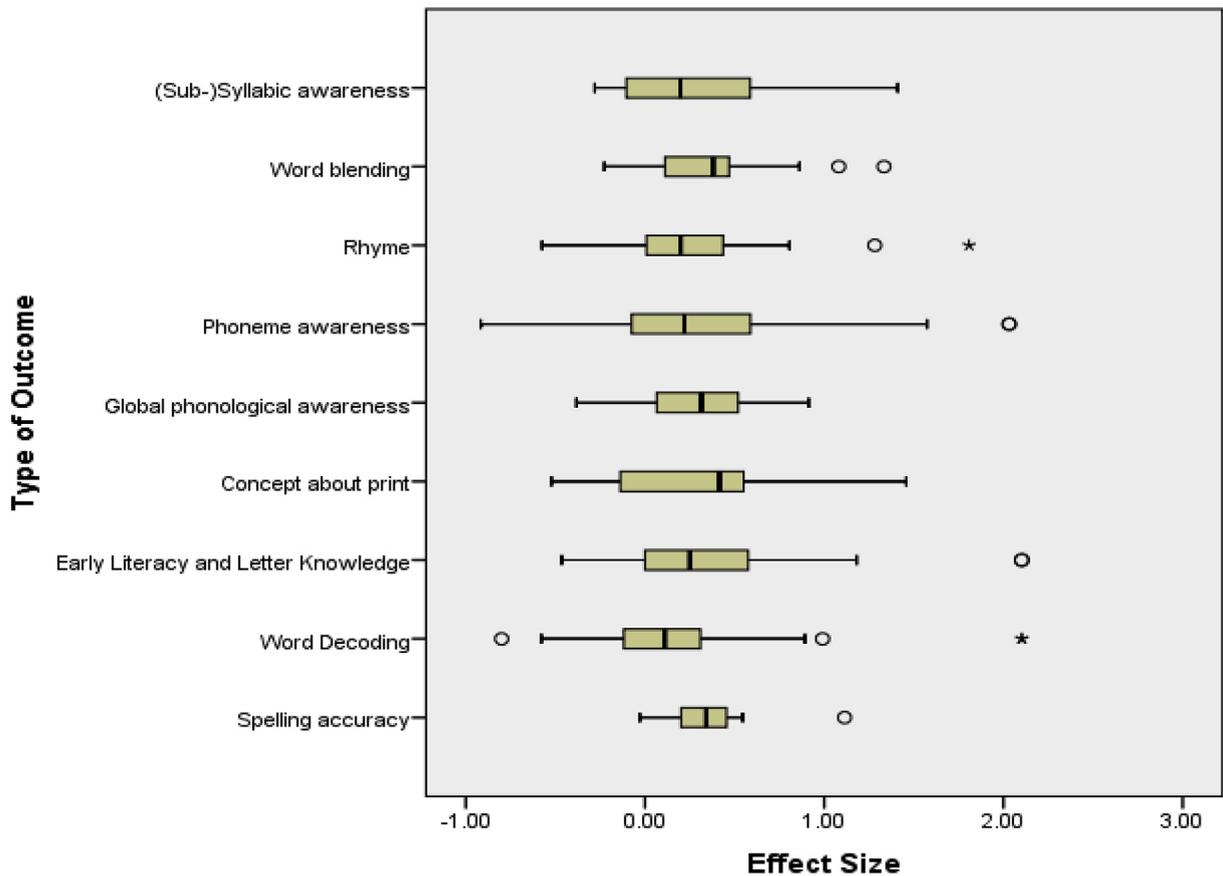


Fig. 2. Boxplot of the distribution of effect sizes per type of criterion variable.

Table 4
Average Effect Sizes by Type of Treatment and Type of Outcome (based on 59 studies).

Type of Treatment					
Type of Outcome	Phonology	Letters	Storybook	Total	n
<i>Phonological Awareness Measures</i>					
(Sub-)Syllabic awareness	0.17		0.37	0.27	23
Word blending	0.37	0.43		0.38	24
Rhyme	0.29	0.00	0.32	0.26	36
Phoneme awareness	0.39	0.26	0.21	0.32	54
Global phonological awareness	0.88	0.38	0.23	0.29	47
Sub-total	0.33 (s = 0.525) n = 86	0.29 (s = 0.471) n = 36	0.27 (s = 0.334) n = 62	0.30 (s = 0.457) (n = 184)	
<i>Reading-Related Measures</i>					
Concept about print	0.48	-0.01	0.49	0.33	22
Early literacy and letter knowledge	0.28	0.36	0.36	0.34	58
Word Decoding	0.22	0.24	0.12	0.15	63
Spelling accuracy	0.30	0.38		0.36	12
Sub-total	0.30 (s = 0.287) n = 18	0.29 (s = 0.476) n = 67	0.23 (s = 0.547) n = 70	0.27 (s = 0.492) (n = 155)	
Total	0.33 (s = 0.491) n = 104	0.29 (s = 0.472) n = 103	0.25 (s = 0.458) n = 132	0.29 (s = 0.473) n = 339	

Note. The n's vary per cell between 0 and 44.

standard error). The largest average effect size was obtained for phonology awareness treatments, the smallest average was found for the storybook treatments, but the average differences between the three types of treatment were not statistically significant, see Model 1 in Table 6a. The variance of effect sizes was clearly larger within studies than between studies (see the empty model in Table 5b). The overall intraclass correlation was 0.20, thus for the whole set of effect sizes 20% of the variance was between studies. Table 5b shows that the distribution of the variances for phonology treatments differs from that of the other two types of treatments.

Table 5a

Relationships of between- and within-study characteristics with effect sizes, separately for focus of treatment.

Model/Effect	Phonology	Letters	Storybook	Total
	(n = 104)	(n = 103)	(n = 132)	(n = 339)
	(m = 11)	(m = 28)	(m = 20)	(m = 59)
0. Empty: Intercept	0.342 (0.076)	0.289 (0.058)	0.234 (0.049)	0.279 (0.034)
1. Focus of Criterion				
- Intercept	0.326 (0.084)	0.301 (0.084)	0.249 (0.063)	0.287 (0.042)
- Reading-related	0.093 (0.104)	-0.019 (0.095)	-0.029 (0.076)	-0.016 (0.051)
2. ILS or not				
- Intercept		0.241 (0.067)		0.262 (0.036)
- ILS	*	0.029 (0.023)	*	0.144 (0.103)
3. Treatment Duration				
- Intercept	0.367 (0.088)	0.298 (0.054)	0.224 (0.053)	0.278 (0.033)
- Program length	-0.003 (0.012)	0.004 (0.006)	0.020 (0.020)	0.004 (0.004)
- Intensity	-0.002 (0.003)	-0.001 (0.001)	0.004 (0.003)	-0.000 (0.001)
4. Participants				
- Intercept	0.328 (0.079)	0.297 (0.076)	0.221 (0.058)	0.268 (0.040)
- Age	-0.025 (0.013)	0.004 (0.009)	-0.002 (0.006)	-0.003 (0.005)
- At risk	0.079 (0.133)	-0.015 (0.124)	0.067 (0.128)	0.045 (0.074)
5. Language:				
- Intercept	0.337 (0.104)	0.280 (0.085)	0.243 (0.010)	0.288 (0.052)
- Transparent	0.024 (0.167)	0.018 (0.118)	-0.012 (0.116)	-0.014 (0.069)
6. Research Design				
- Intercept	0.647 (0.164)	0.520 (0.119)	0.411 (0.241)	0.512 (0.082)
- Random	-0.218 (0.172)	-0.148 (0.122)	-0.138 (0.252)	-0.184 (0.082)
- Active control	-0.184 (0.109)	-0.209 (0.103)	-0.111 (0.101)	-0.142 (0.056)
- Follow-up	-0.166 (0.113)	*	-0.138 (0.260)	-0.167 (0.107)
7. Criterion Variables				
- Intercept	0.348 (0.311)	0.449 (0.203)	0.177 (0.105)	0.213 (0.097)
- Letter knowledge	0.057 (0.105)	0.070 (0.102)	-0.081 (0.091)	0.014 (0.057)
- Speed	-0.495 (0.222)	-0.350 (0.208)	*	-0.378 (0.150)
- Transfer	0.010 (0.301)	-0.046 (0.119)	0.013 (0.097)	0.026 (0.068)
- Match DV-treatment	*	-0.166 (0.153)	0.133 (0.096)	0.068 (0.084)

Statistically significant effects ($p < .05$) are in boldface type; m = number of studies, n = number of effect sizes.

* No variation on predictor.

For phonology treatments the intraclass correlation was high, namely 0.41. The within-studies effect sizes seem clearly more homogeneous for phonology than for the other two types of treatments. This may be due to less diversity in outcome measures used. In Table 5b (and also 6b) results of significance tests are reported for the between-studies variances. Between-studies variances significantly (at the 5%-level) larger than 0 have been printed in boldface type. This is the test for random intercepts performed by comparing the deviance of the random intercept model with the deviance of the fixed intercept model (Snijders & Bosker, 2012, pp. 97–98).

The empty models provide answers to the first research question on the effects of computer-assisted early literacy intervention across types of treatments and outcome measures. All other models in the tables deal with aspects of the second research question to which we turn now, examining treatments effects as a function of study characteristics.

Model 1 tests the average difference between the two main categories of outcome variables, phonological awareness measures versus reading-related measures as distinguished in Table 3. One dummy variable was used to make this distinction. Table 5a shows only minimal and non-significant differences overall as well as for the three types of treatments separately. The estimated intercepts (Table 5a) and variance components (Table 5b) of Model 1 are highly similar to the estimates in the empty model. So the average effects of the treatments appeared nearly the same, regardless the focus of the outcome.

Model 2 concerns the nature of the treatment: an Integrated Learning System (ILS) or not. The ILS occurred only when the focus of the treatment involved both letter knowledge and phonological awareness. The ILS treatments achieved a somewhat higher effect size but the difference was not statistically significant (see Table 5a).

Model 3 is the model for treatment duration. Two variables were identified to represent treatment duration: program length and intensity of treatment. Program length was expressed as the number of weeks a program lasted. And intensity of treatment was the number of minutes the treatment took per week. Both variables were centered at their grand mean. None of them appeared in a statistically significant way related with the effect sizes (see Table 5a). Nevertheless, the between-studies variance reduced by 22% in the overall analysis (from 0.027 to 0.021, see Table 5b), but was still statistically significant.

In Model 4 participant characteristics were controlled for. The average age of the participants in a study, centered around the mean age, appeared not associated with the average effect size obtained in the study. In 18 of the 59 studies were participants identified as being at risk for reading problems. This difference, however, was not significantly associated with the average effect size of the study. Also, the estimates of average effect size (intercepts) only marginally changed after controlling for these participant

Table 5b
Estimated variance components for the multilevel models in Table 5a.

Model/	Phonology	Letters	Storybook	Total
Variance Component	(n = 104)	(n = 103)	(n = 132)	(n = 339)
	(m = 11)	(m = 28)	(m = 20)	(m = 59)
0. Empty				
- Between Studies	0.043 (0.027)	0.032 (0.024)	0.016 (0.015)	0.027 (0.012)
- Within Studies	0.063 (0.021)	0.137 (0.031)	0.101 (0.024)	0.108 (0.015)
1. Focus of Criterion				
- Between Studies	0.052 (0.031)	0.032 (0.024)	0.016 (0.015)	0.027 (0.012)
- Within Studies	0.061 (0.020)	0.139 (0.031)	0.102 (0.024)	0.109 (0.015)
2. ILS or not				
- Between Studies	^a	0.029 (0.023)	^a	0.025 (0.011)
- Within Studies	^a	0.137 (0.030)	^a	0.108 (0.015)
3. Treatment Duration				
- Between Studies	0.057 (0.032)	0.011 (0.020)	0.021 (0.017)	0.021 (0.011)
- Within Studies	0.062 (0.021)	0.154 (0.034)	0.099 (0.023)	0.111 (0.015)
4. Participants				
- Between Studies	0.030 (0.020)	0.039 (0.026)	0.021 (0.017)	0.029 (0.012)
- Within Studies	0.065 (0.021)	0.137 (0.031)	0.101 (0.024)	0.108 (0.015)
5. Language				
- Between Studies	0.052 (0.030)	0.035 (0.025)	0.019 (0.016)	0.028 (0.012)
- Within Studies	0.062 (0.021)	0.137 (0.031)	0.101 (0.024)	0.108 (0.015)
6. Research Design				
- Between Studies	0.029 (0.020)	0.015 (0.019)	0.024 (0.018)	0.019 (0.010)
- Within Studies	0.063 (0.021)	0.141 (0.031)	0.098 (0.023)	0.106 (0.015)
7. Criterion Variables				
- Between Studies	0.047 (0.028)	0.029 (0.023)	0.012 (0.013)	0.027 (0.011)
- Within Studies	0.058 (0.020)	0.137 (0.030)	0.104 (0.024)	0.106 (0.015)

m = number of studies, n = number of effect sizes.

Significance (at 5%-level) of the between-studies variances indicated in boldface type, based on a chi-square test (deviance test).

^a ILS did not occur.

characteristics (see Table 5a). The estimated variance components for the treatment duration model stayed about the same as those in the empty model (see Table 5b), except for a clear reduction in the case of phonology treatments. There the between-studies variance reduced to non-significance.

Model 5 concerns the role of orthographic depth of the language involved, which in our data was highly correlated with the difference between English and other languages (mainly Dutch and Hebrew). Only one language predictor in the form of a dummy variable was used (transparent or not). Table 5a shows that this predictor was hardly related with the effect sizes, not overall and not for any of the types of treatment. Consequently, the estimated variance components for the language model hardly differed from the variance components in the empty model (see Table 5b).

The previous models, except Model 1, concerned predictors at the study level. These predictors, of course, could not explain variation between effect sizes within studies. In the next two models, predictors are involved at the level of effects within studies. Model 6 in Table 5a concerns characteristics of the research design. The average effect size of a study appeared on average lower in studies with random assignment of participants to conditions ($b = -0.184, p = .025$). This effect was not statistically significant in the separate analyses for types of treatment mainly due to the smaller numbers of studies per type of treatment. A statistically significant effect was also observed for the type of control condition in comparisons within studies ($b = -0.184, p = .024$). When the experimental condition was compared with an active control condition the effect size tended to be lower than for comparisons with education as usual. But for the subsets of electronic storybooks and phonology treatments this effect was not statistically significant. In a few studies next to immediate posttests also follow-up tests were used. Effect sizes at follow-up appeared not significantly lower than at the immediate posttest. The average effect sizes appeared almost twice as large in cases of nonrandom assignment and comparisons with non-active control conditions (compare in Table 5a the intercepts of Model 6 with those of the empty model). The treatment model hardly reduced the within-studies variance (see Table 5b). The between-studies variance reduced by 30% in the overall analysis. This reduction was 53% for letters treatments and 33% for phonology treatments, but no reduction for storybooks.

Finally, Model 7 in Table 5a shows the results when characteristics of the criterion variables are used as (within-study) predictors. Four dummy variables were created to characterize the criterion variables: whether or not letter knowledge was required, whether or not the test had a speed character, whether or not the test materials differed from the training materials (transfer), and whether or not the dependent variable matched the goals of the treatment. Only the speediness of the test made a significant difference ($b = -0.378, p = .012$). On speeded tests lower effect sizes were obtained than on non-speeded tests. This effect, however, could only be found for treatments focusing on phonology. In studies on storybooks speeded tests did not occur. And for combined letters and phonology treatments this effect was not statistically significant. As shown in Table 5b the variance components in the overall analysis were hardly affected by the predictors in the criterion model.

Table 6a
Relationships of between- and within-study characteristics with effect sizes, separately for type of criterion variable.

Model/Effect	Phonological Awareness	Reading-related	Total
	(n = 184)	(n = 155)	(n = 339)
	(m = 45)	(m = 48)	(m = 59)
0. Empty: Intercept	0.288 (0.045)	0.261 (0.048)	0.279 (0.034)
1. Focus of Treatment			
- Intercept	0.327 (0.085)	0.263 (0.107)	0.333 (0.070)
- Phonology and Letters	-0.039 (0.111)	0.041 (0.130)	-0.046 (0.088)
- Storybook	-0.072 (0.111)	-0.049 (0.132)	-0.098 (0.090)
2. ILS or not			
- Intercept	0.254 (0.047)	0.248 (0.052)	0.262 (0.036)
- ILS	0.270 (0.132)	0.092 (0.135)	0.144 (0.103)
3. Treatment Duration			
- Intercept	0.294 (0.046)	0.257 (0.048)	0.278 (0.033)
- Program length	-0.001 (0.005)	0.008 (0.006)	0.004 (0.004)
- Intensity	0.000 (0.001)	-0.001 (0.001)	-0.000 (0.001)
4. Participants			
- Intercept	0.289 (0.055)	0.243 (0.054)	0.268 (0.040)
- Age	-0.003 (0.007)	0.002 (0.007)	-0.003 (0.005)
- At risk	0.008 (0.097)	0.081 (0.110)	0.045 (0.074)
5. Language:			
- Intercept	0.313 (0.070)	0.247 (0.070)	0.288 (0.052)
- Transparent	-0.043 (0.092)	0.026 (0.096)	-0.014 (0.069)
6. Research Design			
- Intercept	0.513 (0.108)	0.458 (0.126)	0.512 (0.082)
- Random	-0.135 (0.113)	-0.188 (0.130)	-0.183 (0.085)
- Active control	-0.212 (0.071)	-0.074 (0.090)	-0.145 (0.058)
- Follow-up	-0.151 (0.103)	-0.054 (0.245)	-0.164 (0.107)
7. Criterion Variables			
- Intercept	0.145 (0.146)	0.333 (0.134)	0.213 (0.097)
- Letter knowledge	^a	0.081 (0.118)	0.014 (0.057)
- Speed	-0.205 (0.217)	-0.409 (0.189)	-0.378 (0.150)
- Transfer	0.138 (0.134)	-0.044 (0.086)	0.026 (0.068)
- Match DV-treatment	0.036 (0.126)	-0.138 (0.093)	0.068 (0.084)

Statistically significant effects ($p < .05$) are in boldface type; m = number of studies, n = number of effect sizes.

^a No variation on predictor.

In addition, a second set of analyses was performed, distinguishing types of outcome across all treatments. Two types of outcomes were distinguished: phonological-awareness variables versus reading-related variables. The results of these analyses are presented in Tables 6a and 6b. Note that the column Total in these tables amounts to the same as the column Total in Tables 5a and 5b Only Model 1 differs between the two sets of analyses. Model 1 in Table 6a tests the average differences between the three types of treatment while in Table 5a it concerns the two types of outcome variables.

The empty model deals with the first research question. The empty model shows that the average effect sizes are practically the same for both types of outcomes (Model 0 in Table 6a, for the corresponding significance test see Model 1 in Table 5a). The variance components, however, differed (see Model 0 in Table 6b). For phonological-awareness outcomes a much higher intraclass correlation was obtained than for reading-related outcomes, namely 0.53 versus 0.21. This result parallels the estimates for phonology treatments versus the other two treatments. The phonological-awareness measures were more homogeneous within studies than the reading-related measures.

Next, we present the results for the second research question, separately for types of criterion variables. Model 1 in Tables 6a and 6b deals with the focus of the treatment, namely the three types of treatment distinguished before (see Tables 5a and 5b). As already noted, no statistically significant differences between the three types of treatment were found, and this appeared true for either type of criterion variable.

Model 2 again concerns the nature of the treatment: an ILS or not. ILS treatments showed higher average effect sizes, especially on phonological awareness where the difference was statistically significant ($z = 2.05$, $p = .040$). For phonological awareness the average effect size of ILS was more than twice as large, namely $0.254 + 0.270 = 0.524$, than for other treatments (see Model 2 in Table 6a). The between-studies variance in Model 2 was reduced for phonological awareness by 9.8% compared with the empty model (see Table 6b).

Treatment duration again showed no significant relationship with the average effect sizes, for both types of criterion variables (see Model 3 in Table 6a). The between-studies variance, however, reduced by 9.8% for phonological awareness and by 25% for reading-related variables compared with the empty model (see Model 3 in Table 6b).

Participant characteristics had no impact on average effect sizes (Model 4 in Table 6a) or on their variances (Table 6b). Effect sizes appeared on average largely the same whether the children were at risk for reading problems or not and were not associated with the

Table 6b
Estimated variance components for the multilevel models in Table 6a.

Model/Effect	Phonological Awareness	Reading-related	Total
	(n = 184)	(n = 155)	(n = 339)
	(m = 45)	(m = 48)	(m = 59)
0. Empty			
- Between Studies	0.051 (0.019)	0.036 (0.022)	0.027 (0.012)
- Within Studies	0.045 (0.014)	0.133 (0.027)	0.108 (0.015)
1. Focus of Treatment			
- Between Studies	0.054 (0.020)	0.035 (0.022)	0.029 (0.012)
- Within Studies	0.045 (0.014)	0.133 (0.027)	0.108 (0.015)
2. ILS or not			
- Between Studies	0.046 (0.018)	0.035 (0.022)	0.025 (0.011)
- Within Studies	0.045 (0.014)	0.133 (0.027)	0.108 (0.015)
3. Treatment Duration			
- Between Studies	0.046 (0.018)	0.027 (0.021)	0.021 (0.011)
- Within Studies	0.041 (0.013)	0.140 (0.028)	0.111 (0.015)
4. Participants			
- Between Studies	0.055 (0.020)	0.035 (0.022)	0.029 (0.012)
- Within Studies	0.045 (0.014)	0.133 (0.027)	0.108 (0.015)
5. Language			
- Between Studies	0.053 (0.019)	0.035 (0.022)	0.028 (0.012)
- Within Studies	0.045 (0.014)	0.134 (0.027)	0.108 (0.015)
6. Research Design			
- Between Studies	0.041 (0.016)	0.031 (0.021)	0.019 (0.010)
- Within Studies	0.043 (0.013)	0.131 (0.027)	0.106 (0.015)
7. Criterion Variables			
- Between Studies	0.052 (0.019)	0.043 (0.022)	0.027 (0.011)
- Within Studies	0.046 (0.014)	0.116 (0.025)	0.106 (0.015)

m = number of studies, n = number of effect sizes.

Significance (at 5%-level) of the between-studies variances indicated in boldface type, based on a chi-square test (deviance test).

children's age.

The results for the language model (Model 5) were the same as in the previous analyses. The average effect sizes seemed not affected by the orthographic depth of the language and the results were the same for the two types of criterion variables.

Model 6 concerns characteristics of the research design. Like in the first set of analyses the average effect size appeared on average lower in studies with random assignment of participants to conditions. This was the same for both types of criterion variables but not statistically significant for either of them, because of the increased estimates of standard errors. And also, when the experimental condition was compared with an active control condition the effect size tended to be lower than for comparisons with education as usual. This effect was statistically significant overall and in case of phonological awareness ($z = 2.99, p = .003$) but not in case of reading-related measures. Effect sizes at follow-up were not significantly lower than at the immediate posttest, the same for both types of criterion variables. The average effect sizes appeared considerably larger in cases of nonrandom assignment and in cases of comparisons with non-active control conditions, around 0.46 for reading-related measures and 0.51 for phonological-awareness measures.

Finally, Model 7 in Table 6a shows the results for characteristics of the criterion variables. The effect of speediness of the test appeared statistically significant only for the reading-related measures.

4. Discussion

The focus of the present meta-analysis was on the effects of computer-supported early literacy interventions in preschool and kindergarten in alphabetic languages, published during or after 1995. The first research question was about the size of the effect of computer-assisted early literacy intervention on phonological awareness and reading-related skills. This question was answered by multilevel modeling of 339 effect sizes from 59 studies. In this model the average overall effect size was estimated to be 0.28, a small positive value that was statistically significant ($z = 8.21, p < .001$). A similar (significant) average effect size was also evidenced separately for strict phonological training, combined phonological awareness and letter training and electronic storybook intervention. The average effects of these interventions on phonological-awareness and reading-related variables were about the same. The latter outcome corresponds to the finding in the literature that the development of phonological awareness and early reading is bidirectionally related (Snowling, 2000). With respect to characteristics of the criterion variables only one statistically significant effect was found. Using speeded tests resulted on average in a lower effect size than using non-speeded tests. This outcome can be explained from the fact that early literacy interventions by and large focus on accuracy not speed (see Ehri et al., 2001).

The average effect size in the present meta-analysis is small when compared to findings in earlier meta-analyses on phonological awareness interventions established by Bus and van IJzendoorn (1999), Ehri et al. (2001), and Troia (1999). However, it is important

to note that these interventions were mostly not computer-based but carried out by teachers. Interestingly, in our technology-enhanced treatment studies we found similar (moderate) outcomes for the three types of treatment, whereas in the non-technology studies higher main effects were found for specific phonological awareness training than for storybook reading, and a beneficial effect of combined phonological awareness and letter training was also evidenced. It can tentatively be concluded that teachers outperform electronic devices in facilitating young children's phonological awareness and insight in the alphabetic principle. This result is fully commensurate with previous findings on a broad variety of children's learning processes (cf. Zawacki-Richter & Latchem, 2018).

It is also interesting to note that the present average effect size is about the same as in the meta-analysis of computer-assisted interventions by Blok et al. (2002). In contrast to Ehri's meta-analysis, we found no advantage for the combined phonological awareness and letter training as compared with strict phonological awareness training. Neither did we find a difference between these two conditions and electronic storybook reading. This can tentatively be explained from the fact that since the past two decades an early literacy curriculum has become standard throughout preschool and kindergarten outweighing possible effects of strict phonological awareness and/or letter training (Neuman & Dickinson, 2011).

Next to the average effect size, we also looked at the variability of the effect sizes. Variability appeared very high; the effect sizes ranged from -0.92 to 2.10 . And this was true for most subcategories of effect sizes as well. The lowest amount of variability was found for effects on spelling accuracy and on global phonological awareness. In the multilevel models, the variances were decomposed into between-studies and within-studies variance. Overall, twenty percent of the variance was between studies which shows that there is high variability not only between studies but also within individual studies. Partly, this can be explained by the fact that different criterion variables and varying experimental conditions were introduced within studies which may have led to differential outcomes.

The second research question focused on possible moderators of the treatment effects. First, the impact of treatment characteristics and study design factors on the intervention effects was examined by estimating multilevel models for separate treatment characteristics and design factors in order to have models with not too many predictors. Focus of treatment (phonology, letters, storybooks) showed no evidence for differential intervention effects. Using an integrated learning system (ILS) appeared beneficial, especially evidenced for phonological awareness of the children. This can be seen as an interesting finding. It shows that technology-enhanced early literacy interventions obtain better outcomes as they are better integrated in the school curriculum. This insight has previously been articulated by Grabe and Grabe (2003) examining the general effectiveness of technology use in actual classrooms. However, our present finding should be evaluated with great caution since ILS was only involved in 12 percent of the studies. More research is needed to uncover the effects of the integration of computer-support in the standard early literacy curriculum.

Furthermore, no clear evidence was found for effects of treatment duration or participant characteristics (age of the students and being at risk for reading problems). However, the size of the intervention effects was found to be significantly associated with characteristics of the research design. In studies with random assignment to conditions on average a lower effect size was found. The same was true for comparing the intervention condition with an active control condition (including a special reading treatment) rather than education as usual. Importantly, it should be acknowledged that in the vast majority of cases technology-enhanced intervention studies lack data on treatment fidelity and implementation which was also concluded by Archer et al. (2014). It should also be mentioned that no effects of follow-up measurements were evidenced on the treatments in our analysis. This can primarily be seen as a shortcoming of the studies being conducted since retention measures were involved in less than eight percent of the effects being reported.

To examine the impact of the orthographic depth of the language on the intervention effects, both the mean effect sizes and their subdivisions across categories of programs and outcome measures were related to the distinction between studies focusing on English, on the one hand, and studies focusing on more transparent orthographies, on the other hand. The results showed no evidence for any effect of the transparency of the orthography on the average effect size, nor for the subdivisions of effect sizes. This result shows that notwithstanding the orthographic depth of the language under consideration, substantial effects of technology-enhanced early literacy interventions can be attained. It should be noted that the primary focus of early literacy programs is on the development of phonological awareness which turns out to be highly universal across languages (Ziegler & Goswami, 2005).

The present meta-analysis yields some important guidelines for teachers in early childhood education. Overall, the research in this meta-analysis lends support to the finding that computers can have a beneficial effect on young children's learning in the domain of early literacy. Substantial gains in the domains of phonological awareness, letter knowledge and early reading and spelling can be evidenced via computer-supported early literacy interventions. However, it should be noted that the intervention outcomes stay behind those of classroom-based early literacy interventions. Importantly, the present research indicated that the use of educational technology can be more effective when the materials are better integrated and consistent with the curriculum and provide ongoing scaffolding for each learner. Teacher training can be seen as highly relevant to help them integrating computer-based learning devices in early childhood education (cf. Cassidy & Smith, 2003; Neuman & Dickinson, 2011).

To conclude, the present meta-analysis shows that technology-enhanced early literacy interventions on average yield small positive effects across program types and outcome measures and across languages differing in orthographic depth. If the intervention is part of an integrated learning system in the classroom, the effects are even better. The present analysis also shows the importance of methodological rigor of study designs being used in that effect sizes were higher in cases of absence of randomization and comparisons with classroom teaching as usual which is in line with previous research (cf. Andrews et al., 2007; Troia, 1999).

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.edurev.2020.100325>.

References

- References marked with an asterisk indicate studies included in the meta-analysis.
- Andrews, R., Freeman, A., Hou, D., McGuinn, N., Robinson, A., & Zhu, J. (2007). The effectiveness of information and communication technology on the learning of written English for 5- to 16-year-olds. *British Journal of Educational Technology*, 38(2), 325–336. <https://doi.org/10.1111/j.1467-8535.2006.00628.x>.
- *Anthony, J. L. (2016). For which children of economic disadvantage and in which instructional contexts does Earobics Step 1 improve kindergartners' literacy? *Journal of Research on Educational Effectiveness*, 9(1), 54–76. <https://doi.org/10.1080/19345747.2015.1055637>.
- Anthony, J. L., & Francis, D. J. (2005). Development of phonological awareness. *Current Directions in Psychological Science*, 14, 255–259. <https://doi.org/10.1111/j.0963-7214.2005.00376.x>.
- Anthony, J. L., & Lonigan, C. J. (2004). The nature of phonological awareness: Converging evidence from four studies of preschool and early grade school children. *Journal of Educational Psychology*, 96(1), 43–55. <https://doi.org/10.1037/0022-0663.96.1.43>.
- Archer, K., Savage, R., Sanghera-Sidhu, S., Wood, E., Gottardo, A., & Chen, V. (2014). Examining the effectiveness of technology use in classrooms: A tertiary meta-analysis. *Computers & Education*, 7, 140–149. <https://doi.org/10.1016/j.compedu.2014.06.001>.
- *Bauserman, K. L., Cassady, J. C., Smith, L. L., & Stroud, J. C. (2005). Kindergarten literacy achievement: The effects of the PLATO integrated learning system. *Reading Research & Instruction*, 44, 49–60. <https://doi.org/10.1080/19388070509558436>.
- Becker, H. J., & Hativa, N. (1994). History, theory and research concerning integrated learning systems. *International Journal of Educational Research*, 21(1), 5–12.
- Blachman, B. A. (2000). Phonological awareness. In M. L. Kamil, P. B. Mosenthal, P. D. Pearson, & R. Barr (Vol. Eds.), *Handbook of reading research: Vol III*, (pp. 483–502). Mahwah, NJ: Lawrence Erlbaum Associates.
- Blok, H., Oostdam, R., Otter, M. E., & Overmaat, M. (2002). Computer-assisted instruction in support of beginning reading instruction: A review. *Review of Educational Research*, 72(1), 101–130. <https://doi.org/10.3102/00346543072001101>.
- Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. Chichester: Wiley.
- *Both-de Vries, A. C., & Bus, A. G. (2014). Visual processing of pictures and letters in alphabet books and the implications for letter learning. *Contemporary Educational Psychology*, 39(2), 156–163. <https://doi.org/10.1016/j.cedpsych.2014.03.005>.
- *Brabham, E. G., Murray, B. A., & Bowden, S. H. (2006). Reading alphabet books in kindergarten: Effects of instructional emphasis and media practice. *Journal of Research in Childhood Education*, 20(3), 219–234. <https://doi.org/10.1080/02568540609594563>.
- Bryant, P. E., & Bradley, L. (1985). *Children's reading problems*. Oxford, England: Blackwell.
- 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(3), 403–414. <https://doi.org/10.1037/0022-0663.91.3.403>.
- Bus, A., van IJzendoorn, M., & Pellegrini, A. (1995). Joint book reading makes for success in learning to read: A meta-analysis on intergenerational transmission of literacy. *Review of Educational Research*, 65, 1–21. <https://doi.org/10.3102/00346543065001001>.
- 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. <https://doi.org/10.1177/0956797612473122>.
- *Cassady, J. C., & Smith, L. L. (2003). The impact of a reading-focused integrated learning system on phonological awareness in kindergarten. *Journal of Literacy Research*, 35(4), 947–964. https://doi.org/10.1207/s15548430jlr3504_2.
- Castles, A., Rastle, K., & Nation, K. (2018). Ending the reading wars: Reading acquisition from novice to expert. *Psychological Science in the Public Interest*, 19, 5–51. <https://doi.org/10.1177/1529100618772271>.
- *Chera, P., & Wood, C. (2003). Animated multimedia 'talking books' can promote phonological awareness in children beginning to read. *Learning and Instruction*, 13, 33–52. [https://doi.org/10.1016/S0959-4752\(01\)00035-4](https://doi.org/10.1016/S0959-4752(01)00035-4).
- *Cviko, A., McKenney, S., & Voogt, J. (2012). Teachers enacting a technology-rich curriculum for emergent literacy. *Educational Technology Research & Development*, 60(1), 31–54. <https://doi.org/10.1007/s11423-011-9208-3>.
- *Davidson, M. R., Fields, M. K., & Yang, J. (2009). A randomized trial study of a preschool literacy curriculum: The importance of implementation. *Journal of Research on Educational Effectiveness*, 2(3), 177–208. <https://doi.org/10.1080/19345740902770028>.
- *De Graaff, S., Bosman, A. M. T., Hasselman, F., & Verhoeven, L. (2009). Benefits of systematic phonics instruction. *Scientific Studies of Reading*, 13(4), 318–333. <https://doi.org/10.1080/10888430903001308>.
- *De Jong, M. T., & Bus, A. G. (2002). Quality of book-reading matters for emergent readers: An experiment with the same book in a regular or electronic format. *Journal of Educational Psychology*, 94(1), 145–155. <https://doi.org/10.1037/0022-0663.94.1.145>.
- Ehri, L. (2005). Learning to read words: Theory, findings, and issues. *Scientific Studies of Reading*, 9, 167–188. https://doi.org/10.1207/s1532799xssr0902_4.
- Ehri, L. (2014). Orthographic mapping in the acquisition of sight word reading, spelling memory, and vocabulary learning. *Scientific Studies of Reading*, 18, 5–21. <https://doi.org/10.1080/10888438.2013.819356>.
- Ehri, L. C., Nunes, S. R., Willows, D. M., Valeska Schuster, B., Yaghub-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(3), 250–287. <https://doi.org/10.1598/RRQ.36.3.2>.
- Elbro, C., & Scarborough, H. S. (2003a). Early identification. In P. Bryant, & T. Nunes (Eds.), *Handbook of children's reading* (pp. 339–359). Dordrecht: Kluwer Academic Publishers.
- Elbro, C., & Scarborough, H. S. (2003b). Early intervention. In P. Bryant, & T. Nunes (Eds.), *Handbook of children's reading* (pp. 361–381). Dordrecht: Kluwer Academic Publishers.
- Evans, M. A., Williamson, K., & Pursoo, T. (2008). Preschoolers' attention to print during shared book reading. *Scientific Studies of Reading*, 12(1), 106–129. <https://doi.org/10.1080/10888430701773884>.
- *Gong, Z., & Levy, B. A. (2009). Four year old children's acquisition of print knowledge during electronic storybook reading. *Reading and Writing*, 22, 889–905. <https://doi.org/10.1007/s11145-008-9130-1>.
- Goswami, U. (2000). Phonological and lexical processes. In M. L. Kamil, P. B. Rosenthal, P. D. Pearson, & R. Barr (Vol. Eds.), *Handbook of reading research: Vol 3*, (pp. 251–268). Mahwah, NJ: Erlbaum.
- Goswami, U. (2001). Early phonological development and the acquisition of literacy. In S. Neuman, & D. Dickinson (Eds.), *Handbook of research in early literacy for the 21st century* (pp. 111–125). New York: Guilford Press.
- Grabe, M., & Grabe, C. (2003). *Integrating technology for meaningful learning*. Belmont, CA: Cengage Learning.
- *Hecht, S. A., & Close, L. (2002). Emergent literacy skills and training time uniquely predict variability in responses to phonemic awareness training in disadvantaged kindergartners. *Journal of Experimental Child Psychology*, 82, 93–115. [https://doi.org/10.1016/S0022-0965\(02\)00001-2](https://doi.org/10.1016/S0022-0965(02)00001-2).
- Hedges, L. V. (1982). Estimation of effect size from a series of independent experiments. *Psychological Bulletin*, 92(2), 490–499. <https://doi.org/10.1037/0033-2909.92.2.490>.
- Hogan, T. P., Catts, H. W., & Little, T. D. (2005). The relationship between phonological awareness and reading: Implications for the assessment of phonological awareness. *Language, Speech, and Hearing Services in Schools*, 36(4), 285–293.
- Høien, T., & Lundberg, I. (2000). *Dyslexia: From theory to intervention*. Dordrecht, The Netherlands: Kluwer Academic Publishers.
- *Homer, B. D., Kinzer, C. K., Plass, J. L., Letourneau, S. M., Hoffman, D., Bromley, M., et al. (2014). Moved to learn: The effects of interactivity in a Kinect-based literacy game for beginning readers. *Computers & Education*, 74, 37–49. <https://doi.org/10.1016/j.compedu.2014.01.007>.
- Hox, J. J. (2010). *Multilevel analysis: Techniques and applications* (2nd ed.). New York and Hove: Routledge.
- *Huffstetter, M., King, J. R., Onwuegbuzie, A. J., Schneider, J. J., & Powell-Smith, K. A. (2010). Effects of a computer-based early reading program on the early reading and oral language skills of at-risk preschool children. *Journal of Education for Students Placed at Risk*, 15(4), 279–298. <https://doi.org/10.1080/10824669.2010.532415>.
- *Ihmeideh, F. M. (2014). The effect of electronic books on enhancing emergent literacy skills of pre-school children. *Computers & Education*, 79, 40–48. <https://doi.org/10.1016/j.compedu.2014.07.008>.

- *Johnson, E. P., Perry, J., & Shamir, H. (2010). Variability in reading ability gains as a function of computer-assisted instruction method of presentation. *Computers & Education*, 55, 209–217. <https://doi.org/10.1016/j.compedu.2010.01.006>.
- *Kartal, G., Babür, N., & Erçetin, G. (2016). Training for phonological awareness in an orthographically transparent language in two different modalities. *Reading & Writing Quarterly*, 32, 550–579. <https://doi.org/10.1080/10573569.2015.1065213>.
- *Kartal, G., & Terziyan, T. (2016). Development and evaluation of game-like phonological awareness software for kindergarteners: JerenAli. *Journal of Educational Computing Research*, 53(4), 519–539. <https://doi.org/10.1177/0735633115608397>.
- *Kegel, C. A., & Bus, A. G. (2012). Online tutoring as a pivotal quality of web-based early literacy programs. *Journal of Educational Psychology*, 104, 182–192. <https://doi.org/10.1037/a0025849>.
- *Kegel, C. A. T., Bus, A. G., & Van IJzendoorn, M. H. (2011). Differential susceptibility in early literacy instruction through computer games: The role of the dopamine D4 receptor gene (DRD4). *Mind, Brain, and Education*, 5(2), 71–78. <https://doi.org/10.1111/j.1751-228X.2011.01112.x>.
- *Kegel, C. A. T., van der Kooy-Hofland, V. A. C., & Bus, A. G. (2009). Improving early phoneme skills with a computer program: Differential effects of regulatory skills. *Learning and Individual Differences*, 19(4), 549–554. <https://doi.org/10.1016/j.lindif.2009.07.002>.
- *Korat, O. (2009). The effects of CD-ROM storybook reading on children's emergent literacy as a function of age group and repeated reading. *Education and Information Technologies*, 14(1), 39–53. <https://doi.org/10.1007/s10639-008-9063-y>.
- *Korat, O. (2010). Reading electronic books as a support for vocabulary, story comprehension and word reading in kindergarten and first grade. *Computers & Education*, 55, 24–31. <https://doi.org/10.1016/j.compedu.2009.11.014>.
- *Korat, O., & Blau, H. (2010). Repeated reading of cd-rom storybook as a support for emergent literacy: A developmental perspective in two SES groups. *Journal of Educational Computing Research*, 43(4), 445–466.
- *Korat, O., & Shamir, A. (2007). Electronic books versus adult readers: Effects on children's emergent literacy as a function of social class. *Journal of Computer Assisted Learning*, 23, 248–259. <https://doi.org/10.1111/j.1365-2729.2006.00213.x>.
- *Korat, O., & Shamir, A. (2008). The educational electronic book as a tool for supporting children's emergent literacy in low versus middle SES groups. *Computers & Education*, 50, 110–124. <https://doi.org/10.1016/j.compedu.2006.04.002>.
- *Korat, O., Shamir, A., & Heibal, S. (2013). Expanding the boundaries of shared book reading: E-Books and printed books in parent-child reading as support for children's language. *First Language*, 33, 504–523. <https://doi.org/10.1177/0142723713503148>.
- *Kreskey, D. D., & Truscott, S. D. (2016). Is computer-aided instruction an effective tier-one intervention for kindergarten students at risk for reading failure in an applied setting? *Contemporary School Psychology*, 20(2), 142–151. <https://doi.org/10.1007/s40688-015-0056-8>.
- Landerl, K., Freudenthaler, H., Heene, M., De Jong, P., Desrochers, A., Manolitsis, G., et al. (2019). Phonological awareness and rapid automatized naming as longitudinal predictors of reading in five alphabetic orthographies with varying degrees of consistency. *Scientific Studies of Reading*, 23(3), 220–234. <https://doi.org/10.1080/10888438.2018.1510936>.
- Lerner, M. D., & Lonigan, C. J. (2015). Bidirectional relations between phonological awareness and letter knowledge in preschool revisited: A growth curve analysis of the relation between two code-related skills. *Journal of Experimental Child Psychology*, 144, 166–183. <https://doi.org/10.1016/j.jecp.2015.09.023>.
- Liberman, I. Y., & Shankweiler, D. (1985). Phonology and the problems of learning to read and write. *Remedial and Special Education*, 6, 8–17. <https://doi.org/10.1177/074193258500600604>.
- Lonigan, C. J., Burgess, S. R., & Anthony, J. L. (2000). Development of emergent literacy and early reading skills in preschool children: Evidence from a latent-variable longitudinal study. *Developmental Psychology*, 36, 596–613. <https://doi.org/10.1037/0012-1649.36.5.596>.
- *Lonigan, C. J., Driscoll, K., Phillips, B. M., Cantor, B. G., Anthony, J. L., & Goldstein, H. (2003). A computer-assisted instruction phonological sensitivity program for preschool children at-risk for reading problems. *Journal of Early Intervention*, 25(4), 248–262. <https://doi.org/10.1177/105381510302500402>.
- Lonigan, C. J., Schatschneider, C., & Cunningham, A. (2008). Impact of shared-reading interventions on young children's early literacy skills. *Report of the national early literacy Panel: Vols 153–172*. Jessup, MA: The National Institute for Literacy.
- Lonigan, C. J., Schatschneider, C., & Westberg, L. (2008). Impact of code-focused interventions on young children's early literacy skills. *Report of the national early literacy Panel: Vols 107–152*. Jessup, MA: The National Institute for Literacy.
- *Lovio, R., Halttunen, A., Lyytinen, H., Näätänen, R., & Kujala, T. (2012). Reading skill and neural processing accuracy improvement after a 3-hour intervention in preschoolers with difficulties in reading-related skills. *Brain Research*, 1448, 42–55. <https://doi.org/10.1016/j.brainres.2012.01.071>.
- *Macaruso, P., & Rodman, A. (2011). Efficacy of computer-assisted instruction for the development of early literacy skills in young children. *Reading Psychology*, 32(2), 172–196. <https://doi.org/10.1080/02702711003608071>.
- *Macaruso, P., & Walker, A. (2008). The efficacy of computer-assisted instruction for advancing literacy skills in kindergarten children. *Reading Psychology*, 29, 266–287. <https://doi.org/10.1080/02702710801982019>.
- Marsh, H. W., Borrmann, L., Mutz, R., Dabiel, H., & O'Mara, A. (2009). Gender effects in peer reviews of grant proposals: A comprehensive meta-analysis comparing traditional and multilevel approaches. *Review of Educational Research*, 79(3), 1290–1326. <https://doi.org/10.3102/0034654309334143>.
- *McKenney, S., & Voogt, J. (2009). Designing technology for emergent literacy: The PictoPal initiative. *Computers & Education*, 52(4), 719–729. <https://doi.org/10.1016/j.compedu.2008.11.013>.
- Melby-Lervåg, M., Lyster, S., & Hulme, C. (2012). Phonological skills and their role in learning to read: A meta-analytic review. *Psychological Bulletin*, 138, 322–352. <https://doi.org/10.1037/a0026744>.
- *Mioduser, D., Tur-Kaspa, H., & Leitner, I. (2000). The learning value of computer-based instruction of early reading skills. *Journal of Computer Assisted Learning*, 16(1), 54–63. <https://doi.org/10.1046/j.1365-2729.2000.00115.x>.
- *Mitchell, M. J., & Fox, B. J. (2001). The effects of computer software for developing phonological awareness in low-progress readers. *Reading Research & Instruction*, 40(4), 315–332. <https://doi.org/10.1080/19388070109558353>.
- Moats, L. C. (2000). *Speech to print: Language essentials for teachers*. Baltimore, MD: Brookes.
- Mol, S. E., Bus, A. G., & de Jong, M. T. (2009). Interactive book reading in early education: A tool to stimulate print knowledge as well as oral language. *Review of Educational Research*, 79(2), 979–1007. <https://doi.org/10.3102/0034654309332561>.
- Molfese, V., & Westberg, L. (2008). Impact of preschool and kindergarten programs on young children's early literacy skills. *Report of the national early literacy Panel: Vols 189–210*. Jessup, MA: The National Institute for Literacy.
- *Moody, A., Skribbe, L., Parker, M., & Jones, A. (2014). Use of electronic storybooks to promote print awareness in preschoolers who are living in poverty. *The Journal of Literacy and Technology*, 15(3), 2–27.
- Nagy, E., & Scott, J. (2000). Vocabulary processes. In M. L. Kamil, P. B. Mosenthal, P. D. Pearson, & R. Barr (Vol. Eds.), *Handbook of reading research: Vol 3*, (pp. 269–284). Mahwah, NJ: Lawrence Erlbaum.
- National Institute of Child Health and Human Development (2000). *Report of the National Reading Panel. Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction: Reports of the subgroups (NIH Publication No. 00-4754)* Washington, DC: U.S. Government Printing Office. www.nichd.nih.gov/publications/nrp/report.htm.
- Neuman, S. B., & Dickinson, D. K. (2011). *Handbook of early literacy research*. New York: Guilford Press.
- *O'Callaghan, P., McIvor, A., McVeigh, C., & Rushe, T. (2016). A randomized controlled trial of an early-intervention, computer-based literacy program to boost phonological skills in 4- to 6-year-old children. *British Journal of Educational Psychology*, 86(4), 546–558. <https://doi.org/10.1111/bjep.12122>.
- Perfetti, C. A., Beck, I., Bell, L., & Hughes, C. (1987). Phonemic knowledge and learning to read are reciprocal: A longitudinal study of first grade children. *Merrill-Palmer Quarterly*, 33, 283–319.
- *Piquette, N. A., Savage, R. S., & Abrami, P. C. (2014). A cluster randomized controlled field trial of the ABRACADABRA web-based reading technology: Replication and extension of basic findings. *Frontiers in Psychology*, 5. <https://doi.org/10.3389/fpsyg.2014.01413> Article 1413.
- *Plak, R. D., Kegel, C. A. T., & Bus, A. G. (2015). Genetic differential susceptibility in literacy delayed children: A randomized controlled trial on emergent literacy in kindergarten. *Development and Psychopathology*, 27(1), 69–79. <https://doi.org/10.1017/S0954579414001308>.
- *Plak, R. D., Merkelbach, I., Kegel, C. A. T., van IJzendoorn, H. H., & Bus, A. G. (2016). Brief computer interventions enhance emergent academic skills in susceptible

- children: A gene-by-environment experiment. *Learning and Instruction*, 45(1), 1–8. <https://doi.org/10.1016/j.learninstruc.2016.06.002>.
- Rasbash, J., Charlton, C., Browne, W. J., Healy, M., & Cameron, B. (2016). *MLwiN version 2.36*. Centre for Multilevel Modelling, University of Bristol.
- *Regtvoort, A. G. F. M., & van der Leij, A. (2007). Early Intervention with children of dyslexic parents: Effects of computer-based reading instruction at home on literacy acquisition. *Learning and Individual Differences*, 17, 35–53. <https://doi.org/10.1016/j.lindif.2007.01.005>.
- *Reitsma, P., & Wesseling, R. (1998). Effects of computer-assisted training of blending skills in kindergartners. *Scientific Studies of Reading*, 2(4), 301–320. https://doi.org/10.1207/s1532799xssr0204_1.
- Rohde, L. (2015). *The comprehensive emergent literacy model: Early literacy in context*, Vol 5SAGE Open <https://doi.org/10.1177/2158244015577664>.
- *Savage, R. S., Abrami, P. C., Piquette, N., Wood, E., Deleveaux, G., Sukhbinder, S.-S., et al. (2013). A (Pan-Canadian) cluster randomized control effectiveness trial of the ABRACADABRA web-based literacy program. *Journal of Educational Psychology*, 105(2), 310–328. <https://doi.org/10.1037/a0031025>.
- Scarborough, H. (1998). Early identification of children at risk for reading disabilities. In B. K. Shaprio, P. J. Accardo, & A. J. Capute (Eds.). *Specific reading disability. A view of the spectrum* (pp. 75–119). Timonium, MD: York Press.
- *Segal-Drori, O., Korat, O., Shamir, A., & Klein, P. S. (2010). Reading electronic and printed books with and without adult instruction: Effects on emergent reading. *Reading and Writing*, 23, 913–930. <https://doi.org/10.1007/s11145-009-9182-x>.
- *Segers, E., & Verhoeven, L. (2005). Long-term effects of computer training of phonological awareness in kindergarten. *Journal of Computer Assisted Learning*, 21, 17–27. <https://doi.org/10.1111/j.1365-2729.2005.00107.x>.
- Sénéchal, M., LeFevre, J., Smith-Chant, B. L., & Colton, K. V. (2001). On refining theoretical models of emergent literacy: The role of empirical evidence. *Journal of School Psychology*, 39, 439–460.
- *Shamir, A. (2009). Processes and outcomes of joint activity with e-books for promoting kindergartners' emergent literacy. *Educational Media International*, 46(1), 81–96. <https://doi.org/10.1080/09523980902781295>.
- *Shamir, A., Korat, O., & Barbi, N. (2008). The effects of CD-ROM storybook reading on low SES kindergartners' emergent literacy as a function of learning context. *Computers & Education*, 51(1), 354–367. <https://doi.org/10.1016/j.compedu.2007.05.010>.
- *Shamir, A., Korat, O., & Fellah, R. (2012). Promoting vocabulary, phonological awareness and concept about print among children at risk for learning disability: Can e-books help? *Reading and Writing*, 25(1), 45–69. <https://doi.org/10.1007/s11145-010-9247-x>.
- *Shamir, A., & Lifshitz, I. (2013). E-books for supporting the emergent literacy and emergent math of children at risk for learning disabilities: Can metacognitive guidance make a difference? *European Journal of Special Needs Education*, 28(1), 33–48. <https://doi.org/10.1080/08856257.2012.742746>.
- *Shamir, A., & Shlafer, I. (2011). E-Books effectiveness in promoting phonological awareness and concept about print: A comparison between children at risk for learning disabilities and typically developing kindergartners. *Computers & Education*, 57(3), 1989–1997. <https://doi.org/10.1016/j.compedu.2011.05.001>.
- Shankweiler, D., & Liberman, I. (1989). *Phonology and reading disability*. Ann Arbor, MI: University Press.
- Share, D. L. (2004). Orthographic learning at a glance: On the time course and developmental onset of reading. *Journal of Experimental Child Psychology*, 87, 267–298. <https://doi.org/10.1016/j.jecp.2004.01.001>.
- Snijders, T. A. B., & Bosker, R. J. (2012). *Multilevel analysis: An introduction to basic and advanced multilevel modeling* (2nd ed.). Los Angeles: Sage.
- Snowling, M. J. (2000). Language and literacy skills: Who is at risk and why? In D. V. M. Bishop, & L. B. Leonard (Eds.). *Speech and language impairment in children: Causes, characteristics, interventions and outcome* (pp. 245–260). Hove, UK: Psychology Press.
- Swanson, H. L., Trainin, G., Necochea, D. M., & Hammill, D. D. (2003). Rapid naming, phonological awareness, and reading: A meta-analysis of the correlational evidence. *Review of Educational Research*, 73, 407–440. <https://doi.org/10.3102/00346543073004407>.
- Tallal, P. (2004). Improving language and literacy is a matter of time. *Nature Reviews Neuroscience*, 5, 721–728. <https://doi.org/10.1038/nrn1499>.
- Torgesen, J. K. (2002). The prevention of reading difficulties. *Journal of School Psychology*, 40, 7–26. [https://doi.org/10.1016/S0022-4405\(01\)00092-9](https://doi.org/10.1016/S0022-4405(01)00092-9).
- Torgesen, J. K., & Mathes, P. (2000). *A basic guide to understanding, assessing, and teaching phonological awareness*. Austin, TX: ProEd.
- *Tracey, D. H., & Young, J. W. (2007). Technology and early literacy: The impact of an integrated learning system on high-risk kindergartners' achievement. *Reading Psychology*, 28(5), 443–467. <https://doi.org/10.1080/02702710701568488>.
- Troia, G. A. (1999). Phonological awareness intervention research: A critical review of the experimental methodology. *Reading Research Quarterly*, 34(1), 28–52. <https://doi.org/10.1598/RRQ.34.1.3>.
- *Van Daal, V. H. P., & Reitsma, P. (2000). Computer-assisted learning to read and spell: Results from two pilot studies. *Journal of Research in Reading*, 23(2), 181–193. <https://doi.org/10.1111/1467-9817.00113>.
- *Van der Kooy-Hofland, V. A. C., Bus, A. G., & Roskos, K. (2012). Effects of a brief but intensive remedial computer intervention in a sub-sample of kindergartners with early literacy delays. *Reading and Writing*, 25(7), 1479–1497. <https://doi.org/10.1007/s11145-011-9328-5>.
- *Van de Sande, E., Segers, E., & Verhoeven, L. (2016). Supporting executive functions during children's preliteracy learning with the computer. *Journal of Computer Assisted Learning*, 32, 468–480. <https://doi.org/10.1111/jcal.12147>.
- Vloedgraven, J., & Verhoeven, L. (2007). Screening of phonological awareness in the early elementary grades: An IRT approach. *Annals of Dyslexia*, 57, 33–50. <https://doi.org/10.1007/s11881-007-0001-2>.
- Vloedgraven, J., & Verhoeven, L. (2009). The nature of phonological awareness throughout the elementary grades: An item response theory perspective. *Learning and Individual Differences*, 19, 161–169. <https://doi.org/10.1016/j.lindif.2008.09.005>.
- Whitehurst, G. J., & Lonigan, C. J. (Eds.). (2001). *Emergent literacy: Development from prereaders to readers*. New York: Guilford.
- *Wild, M. (2009). Using computer-aided instruction to support the systematic practice of phonological skills in beginning readers. *Journal of Research in Reading*, 32(4), 413–432. <https://doi.org/10.1111/j.1467-9817.2009.01405.x>.
- *Willoughby, D., Evans, M. A., & Nowak, S. (2015). Do ABC eBooks boost engagement and learning in preschoolers? An experimental study comparing eBooks with paper ABC and storybook controls. *Computers & Education*, 82, 107–117. <https://doi.org/10.1016/j.compedu.2014.11.008>.
- *Wolgemuth, J. F., Savage, R. S., Helmer, J., Harper, H., Lea, T., Abrami, P. C., et al. (2013). ABRACADABRA aids Indigenous and non-Indigenous early literacy in Australia: Evidence from a multisite randomized controlled trial. *Computers & Education*, 67, 250–264. <https://doi.org/10.1016/j.compedu.2013.04.002>.
- Zawacki-Richter, O., & Latchem, C. (2018). Exploring four decades of research in Computers & Education. *Computers & Education*, 122, 136–152.
- 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. <https://doi.org/10.1037/0033-2909.131.1.3>.