To read or not to read

Mol, S.E.

4

Interactive Book Reading in Early Education: A Tool to Stimulate Print Knowledge as well as Oral Language

Abstract

This meta-analysis examines to what extent interactive storybook reading stimulates two pillars of learning to read: vocabulary and print knowledge. We quantitatively reviewed 31 (quasi)experiments (N = 2,049 children) in which educators were trained to encourage children to be actively involved before, during, and after joint book reading. A moderate effect size of d=.54 (CI = .33, .74) was found for oral language skills, implying that both quality and quantity of book reading in classrooms are important. Although teaching print-related skills is not part of interactive reading programs, 7% of the variance in kindergarten children's alphabetic knowledge could be attributed to the intervention. The study also shows that findings with experimenters were simply not replicable in a natural classroom setting. Further research is needed to disentangle the processes that explain the effects of interactive reading on children's print knowledge and the implementation strategies that may help transferring intervention effects from researchers to children's own teachers.

Based on:
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,* 979-1007. Doi: 10.3102/0034654309332561

Introduction

Exposure to books is considered to be a major source for developing one of the pillars of learning to read: vocabulary (e.g., Bus, Van IJzendoorn, & Pellegrini, 1995; Juel, 2006; Scarborough & Dobrich, 1994; Sénéchal & LeFevre, 2002). The nature of a text, the quality of the reading style, and the number of times a book is reread seem to be important contributors to young children's vocabulary development (e.g., Dickinson & Smith, 1994; McKeown & Beck, 2006; Reese & Cox, 1999). It is under discussion to what extent book reading also fosters the second pillar, print knowledge. There is ample evidence for the hypothesis that children spontaneously ignore the print during storybook reading (Bus & Van IJzendoorn, 1988; Evans & Saint-Aubin, 2005; Justice, Skibbe, Canning, & Lankford, 2005; Yaden, Smolkin, & Conlon, 1989). On the other hand, as children grow older and become more proficient in print knowledge, they may feel more attracted to letters and sounds in books (e.g., Mason, 1992; Morris, Bloodgood, Lomax, & Perney, 2003; Roy-Charland, Saint-Aubin, & Evans, 2007). In other words, even though evocative techniques, informative feedback, and sensitivity to a child's abilities include incentives mostly for children's oral language skills and not for print knowledge (e.g., Justice & Ezell, 2002; Whitehurst, Epstein, et al., 1994), the story text itself might trigger children's attention to the print, perhaps resulting in interactions with grown-ups that go beyond story understanding.

A recent meta-analysis of the effects of interactive book reading experiments in the family showed that 4% of the variance in vocabulary growth was explained by the additional effects of Whitehurst's (Whitehurst, Falco, Lonigan, & Fischel, 1988) "*Dialogic Reading*"- technique (Mol, Bus, De Jong, & Smeets, 2008). We hypothesized that Dialogic Reading provides a venue for focused language exchange, enabling responses to children's utterances and thinking processes as well as exposure to more formal adult language (e.g., Raikes et al., 2006). The data provided by this set of studies did not enable us to test the extent to which print knowledge was affected by the intervention, however, as there were hardly any studies that included measures such as alphabet knowledge, phonological sensitivity, or orthographic awareness. Strikingly, two subgroups did not appear to benefit from the intervention: The oral language skills of kindergarten children as well as children at risk for language and literacy impairments benefited less from interactive parent-child book reading (Mol et al., 2008). Because dialogue during shared reading is hardly observed in families at risk (e.g., Bus & Van IJzendoorn, 1995; Heath, 1982; Ninio, 1980), low-educated parents might have experienced difficulty with incorporating the trained techniques. On the other hand, expectations and methods may be pitched too low for older children. Too much talking might have a depressing effect on learning in more advanced groups. As teachers appear to provide more cognitively demanding talk about books than parents, the literacy environment at school might be more stimulating for these groups of children (Dickinson & Smith, 1994). Covering book reading research in school settings until 1995, Blok (1999) tested the effects of book reading frequency on language and reading development in 2.5-to 7.5-year-old children, but the studies he included in his meta-analysis did not provide sufficient information about the quality of book reading such as the reading style of the teachers. Therefore, the current meta-analysis elaborates on the gaps in these previous meta-analyses by exploring to what extent Dialogic Reading – taking the form of (a) the use of evocative techniques that encourage the child to talk about pictured materials; (b) informative feedback that highlights the differences between what the child has said and what he or she might have said; and (c) an adaptive adult who is sensitive to the child's developing abilities (e.g., Whitehurst et al., 1988) – before, during, and after reading storybooks affects children's language acquisition as well as print knowledge.

We selected studies in which teachers and/or graduate students were instructed to implement an interactive reading intervention in preschool or kindergarten classrooms. Insofar as the interventions did not use Whitehurst and colleagues' (1988) techniques of Dialogic Reading, teachers and experimenters were trained in applying similar reading techniques: to prompt child responses by asking open-ended questions or making comments, and to support children's enthusiasm and learning opportunities by providing positive reinforcement or relating the story text to their real life experiences. Mostly, teachers or experimenters received handouts that summarized the learned techniques as well as (suggestions for) storybooks. Alternatively, scripted questions or comments were added to storybooks in order to promote the use of similar interactive prompts and responses in each classroom (Brabham & Lynch-Brown, 2002; Kertoy, 1994; Mautte, 1990; Van Kleeck, Vander Woude, & Hammett, 2006). Some teachers were repeatedly observed in the classroom and coached by the researcher to ensure that they applied the reading strategies in various situations (Aram, 2006; Droop, Peters, Aarnoutse, & Verhoeven, 2005; Wasik & Bond, 2001; Wasik, Bond, & Hindman, 2006). In several studies, interactive book reading was accompanied by kits with materials that focused on book-related vocabulary, games with rhyme or letters, and painting or dramatizing the stories (Aram, 2006; Aram & Biron, 2004; Droop et al., 2005; Karweit, 1989; Wasik & Bond, 2001; Wasik et al., 2006). As these additional activities were integrated in the classroom environment, they were expected to foster vocabulary and print knowledge beyond the interactive reading sessions. For children at risk in particular, repeatedly interacting with storybooks might be an important extra stimulant. That is, we do not expect that reading in the classroom can add to the rich home literacy environment that children who are not at risk are likely to experience (e.g., Adams, 1990; Hart & Risley, 2003).

We anticipated greater gains for experiments in which experimenters read to the children than for interventions that were executed by the children's own teachers. Compared to researchers who are well informed about literacy acquisition, teachers may be less familiar with ways to promote literacy and with theories behind interventions. Besides, teachers may have less time and energy to invest in a program that has to be combined with everyday responsibilities (Aram, 2006). It is of great significance, however, that intervention effects induced by researchers can be transferred or generalized to classroom conditions (e.g., Fuchs et al., 2001; Shernoff & Kratochwill, 2007). It seems that teachers are more likely to implement innovations when the programs are well specified, include attractive and user-friendly materials, and are accompanied with training and technical assistance such as coaching or personalized consultation prior to and during the implementation phase (Rohrbach, Grana, Sussman, & Valente, 2006).

Group size may be another important moderator. That is, it can be questioned whether it is possible to engage all participants in group conversations that are challenging as well as comprehensible to children (Dickinson & Sprague, 2001). Sessions involving the entire class require a level of attention that at-risk youngsters are more likely to lack due to fewer opportunities to practice focused attention in other settings (Bodrova & Leong, 2006; Diamond, Barnett, Thomas, & Munro, 2007). On the other hand, as group sessions offer ample opportunity to observe and interact with more literate peers, we expected that the oral language skills of children at risk may improve from shared reading in (small) group sessions. Morrow and Smith (1990) showed that reading to children in small groups offered as much interaction as one-to-one reading, and led even to greater gains in story comprehension than individual sessions. Teachers seem to provide more positive comments and spend more time redirecting the discussion to the story when they are reading to small groups (Karweit & Wasik, 1996). In addition, children's receptive vocabulary is especially thought to improve as a result of repeatedly reading the same storybook because of the additional opportunities to encode, associate, and store novel information due to several exposures (Biemiller & Boote, 2006; Moschovaki & Meadows, 2004; Nielsen, 1993; Sénéchal, 1997).

In sum, by quantitatively and systematically summarizing (quasi-)experiments that examined the effects of interactive reading in educational settings, we addressed the following research questions:

1) Does trained interactive teacher behavior as a part of book reading improve young children's language and print-related skills, or does this behavior not add anything to the effects of joint book reading? We expected that children in the experimental groups would learn more than control-group children who were read to without a special focus on interaction.
2) Are effect sizes of interactive reading as great for print knowledge as oral language? We expected oral language skills to show greater gains than print-related skills in younger and hence less proficient children, whereas we hypothesized that print knowledge would be affected more in kindergartners.

3) Which conditions benefit the efficacy of an interactive reading intervention in the classroom? First, are interventions carried out by experimenters more effective than those implemented by teachers? Second, is reading in small groups more effective than whole-group reading or individual sessions? Third, is there support for the assumption that extra opportunities to use book vocabulary during play, art, or drama activities add to the effects of book reading, as Karweit and Wasik (1996) suggest? Fourth, are at-risk groups especially susceptible to interactive reading interventions, taking into account that they receive fewer incentives at home (Raikes et al., 2006)?

Possible methodological confounders, such as publication status, year of publication, design, and experiment fidelity, were examined as well.

Method

Search Strategy and Selection Criteria

This meta-analysis examines the effect of interactive book reading on the oral language and print knowledge of children not yet reading conventionally. To obtain eligible studies, social science research databases (PsycINFO, ERIC, Dissertation Abstracts International, WebSPIRS, C2-SPECTR, and the Best Evidence Encyclopedia) were searched up to December 2007, using different combinations of the keywords: *(dialogic/interactive) read*, intervention/program, teacher, classroom, early education, daycare, preschool,* and/or *kindergarten,* with *vocabulary, language acquisition/growth, story comprehension, (early/emergent) literacy, (print/alphabet*/letter) knowledge,* and *phon* / orthograph* awareness/ sensitivity* as dependent variables. We also used the so-called snowball method by identifying eligible studies within the references of the collected articles.

Studies were included when they met the following criteria: (a) the study used an interactive, shared reading intervention with open-ended questions, prompts, comments, and positive reinforcement in encouraging children to become actively involved before, during, and after storybook reading; (b) the program was implemented in day care centers, preschool, kindergarten, or first-grade classrooms, and was not part of a larger intervention that specifically targeted the teaching of literacy concepts such as phonological sensitivity or orthographic awareness; (c) teachers, teacher aides, and/or research assistants were trained in using interactive reading techniques with individual or groups of children; (d) participants had no mental, physical, or sensory handicaps and were pre-conventional readers; (e) outcome variables included at least one objective measure of vocabulary or story comprehension; (f) a (quasi-)experimental design was applied, randomly assigning children to either an experimental or control group on individual, school, or classroom level; (g) children in the control group attended the regular school program, not including interactive reading; and (h) articles were published or unpublished, as far as the language of the article could be interpreted and a sufficient amount of statistical information was reported to determine effect sizes. Combined home and school interventions were eligible when separate data for experimental children in a single teacher group were presented.

Studies were excluded when the reading sessions were not the main focus, but book reading motivated teaching vocabulary (e.g., Ard & Beverly, 2004; Beck & McKeown, 2007; Biemiller & Boote, 2006; Collins, 2005; Elley, 1989; Justice, Meier, & Walpole, 2005; Penno, Wilkinson, & Moore, 2002; Sénéchal, Thomas, & Monker, 1995) or reading strategies and print concepts (e.g., Justice & Ezell, 2002; McCormick & Mason, 1986; L. M. Phillips, Norris, Mason, & Kerr, 1990). Furthermore, some studies could not be included because relevant data for separate interventions were not reported (e.g., Hargrave & Sénéchal, 2000; Whitehurst, Epstein, et al., 1994; Whitehurst, Zevenbergen et al., 1999) or because no control group was included (e.g., Morrow & Smith, 1990; Reese & Cox, 1999; Van Elsäcker & Verhoeven, 1997). The study of Brabham and Lynch-Brown (2002) was only partly included. That is, we excluded data regarding third graders and an experimental group in which children were not allowed to interact during the storybook sharing. Of the two intervention groups in Kertoy's study (1994), we included only the experimental group in which adults asked questions because these techniques reflected our inclusion criteria best. Because Mautte (1990) presented composite scores for print knowledge outcomes instead of separate scores, we included only her oral language measure.

Coding Process

As study or methodological characteristics, we coded publication year, publication status (1. published, 2. unpublished), sample size, design (1. experiment, 2. quasi-experiment[1]), and experiment fidelity (check up on reading techniques and frequency, in experimental and control groups). To test which populations benefited most from the intervention, we coded the language of the shared reading sessions (1. English, 2. other), school type (1. preschool, 2. kindergarten), and risk status (1. at risk for language and literacy impairments, 2. not at risk) of the participating children. Intervention characteristics were coded as another group of moderators, among which were the following: characteristics of the adult who carried out the intervention (1. teacher, 2. experimenter), size of the groups in which book reading took place (1. individual, 2. small group [max. 5 children], 3. large group), type of intervention program (1. Dialogic Reading in accordance with Whitehurst et al. (1988), 2. interactive reading without extra activities, 3. interactive reading with extra book-related classroom activities), information

[1] A study was coded as an experiment when each individual child was randomly assigned to a control or experimental group. Studies that randomized on classroom- or school-level but reported results on a subject level were treated as quasi-experiments.

about activities in the control group (1. intervention, such as reading the same books as the experimental group without interaction, 2. no intervention (such as stimulating play or the standard curriculum), and the duration of the intervention (in weeks, and the number of, recommended and/or mean, interactive reading sessions).

To calculate Cohen's *d* effect sizes, we gave preference to computing those with the help of posttest means and standard deviations for oral language measures (receptive vocabulary, expressive vocabulary, story comprehension, and syntax) and print-related skills (alphabetic knowledge, phonological sensitivity and orthographic awareness). Because some studies presented only *F*-values (Wasik & Bond, 2001), means corrected for pretest scores (Morrow, 1988; Morrow, 1989), or gain scores (Aram, 2006) instead of "raw" posttest data, an extra moderator was created to test the effect of positively biased outcomes. Analyses showed that the mean effect sizes slightly decreased when studies with the adjusted scores were excluded. However, in broad lines, the outcomes were similar to those found for the whole sample. We decided therefore to estimate mean effect sizes with all included studies.

When authors reported two independent experiments within one article (Aram, 2006; Droop et al., 2005; Karweit, 1989; Lonigan & Whitehurst, 1998), both studies were treated and coded separately[2]. When studies presented two parallel comparisons with one control group and all groups met the inclusion criteria, we split the study and adapted the sample sizes without adjusting the outcome values. For example, Lamb (1986), Lonigan, Anthony, Bloomfield, Dyer, and Samwel (1999), and Mautte (1990) included one experimental and two control conditions: a control group in which children were read to from the same books as the experimental group without interaction, next to a control group that attended solely the standard preschool curriculum between the pre- and posttest. To compare the effect of the intervention with both control groups, we divided the sample size of the experimental group by two and treated outcomes of the comparison with the control groups as two separate studies. On the other hand, Morrow (1988) was interested in the effect of repeated versus onetime book reading, so she included two experimental groups and one control group. Therefore, we split the sample size of the control group and included all the children who were read to interactively. In all four studies in which the sample sizes were adjusted, the means and standard deviations remained unchanged.

[2] Droop, Peters, Aarnoutse, and Verhoeven (2005) reported two independent experiments in which they trained teachers to read interactively to children classified as at risk and not at risk. Although all participating children were read to in large groups, the authors treated children at risk and not at risk as separate groups in their analyses – without reporting overall means and standard deviations for the control and experimental groups. We included all five comparisons as independent studies as analyses in which we excluded these samples showed that mean effect sizes were not affected more than could be expected by a decrease in power.

As a control, we excluded one of the four samples of each study that we split to analyze only independent studies with original sample sizes. The main effects did not differ from the outcomes that included the complete set of studies.

Two independent coders both coded all derived studies. We agreed completely on the studies that did not meet the inclusion criteria and had to be excluded. Of the 31 included studies, the agreement across study characteristics and moderators ranged from .67 to 1.00, resulting in a *mean* κ of .93 (M = 98%). The experiment fidelity scale that we calculated consisted of a sum of four items: whether the researchers checked (by means of self-reports or audiotaped reading sessions) (a) the use of trained interactive reading techniques in the experimental group, (b) the quality of reading sessions within the control group, and (c) the frequency of book reading in the experimental and (d) in the control group. Because not all studies reported this information clearly, this scale was more difficult to code reliably but the level of agreement was still satisfactory (M = 85%; κ = .71, *range* = .59 – .86). All discrepancies between coders were discussed and corrected. Authors were contacted when some uncertainties could not be clarified after carefully reading their article.

Statistical Analysis

We quantified the added value of interactive reading on young children's language and literacy development by using the Comprehensive Meta-Analysis program (Version 2.2; Borenstein, Hedges, Higgins, & Rothstein, 2005). To give greater weight to studies with larger sample sizes, the standardized differences between the means were determined by using weights based on the inverse of the variance (Lipsey & Wilson, 2001). Per study, effect sizes were first computed for receptive and expressive vocabulary outcomes separately. We combined these aggregated vocabulary scores with syntax and story comprehension measures to estimate a composite effect size for oral language. For print knowledge, we calculated separate effect sizes for alphabetic knowledge (e.g., tests measuring concepts of print, letter names), phonological sensitivity (e.g., rhyme, alliteration, blending, elision), and orthographic awareness (e.g., name/word writing). A positive sign in the Cohen's d column of Appendix 4.1 indicates a favorable outcome for the intervention program, with a d of .20 interpreted as a small, .50 as a moderate and .80 as a large effect size (Cohen, 1992). Put differently, a d of .50 indicates that the interactive reading has moved a child to the 69th percentile on average, compared to a child in the control group (Bus & Van IJzendoorn, 2004). The binominal effect size display, computed from the formula .50 ± (r / 2), indicates to what extent the prediction of children's language and literacy outcomes is enhanced by the intervention.

Overall effect sizes and 95% confidence intervals (CIs) around the point estimates were based on random effects models because such an approach is a conservative solution to deal with heterogeneity (Ioannidis, Patsopoulos,

& Evangelou, 2007; Viechtbauer, 2007). In such a model it is assumed that the variability beyond subject-level sampling error is derived from random differences among studies whose sources cannot be identified (Lipsey & Wilson, 2001). Therefore, another random component (reflected by τ) is included in addition to subject-level sampling error, resulting in wider CIs. For moderator analyses, the random effects model was applied as well. Contrasts of methodological, population, and intervention characteristics were tested and presented only when all cells within the subset contained at least four studies (Bakermans-Kranenburg, Van IJzendoorn, & Juffer, 2003). A significantly different effect size in study outcome was determined by a significant $Q_{between}(df)$-value.

Because studies with non-significant findings or small sample sizes are less likely to be published and we have not located any unpublished reports except for dissertations, we examined this potential publication bias graphically by funnel plot analysis. We also calculated fail-safe numbers (*N*fs), indicating the number of (unpublished) studies that are needed to overturn a significant result (Lipsey & Wilson, 2001).

Preliminary Testing of Potential Biases

For all six effect size composites, no outliers were identified on the basis of standardized *z*-values larger than 3.26 or smaller than -3.26 ($p < .001$), neither did we find evidence for any methodological biases. That is, publication status did not bias the effects of the intervention ($k_{dissertation} = 5$, $k_{published\ study} = 26$; $Q(1) = .21$, $p > .05$). The effect sizes for all language and print knowledge outcomes did not significantly differ for true experiments ($k = 19$) and quasi-experimental studies ($k = 12$; $Q(1) = 1.69$, $p > .05$). The fidelity scale that we developed to examine whether more controlled experiments would reveal stronger results than studies that hardly checked the content and/or frequency of book reading sessions in the experimental and control conditions revealed a statistically significant difference between quasi-experimental and experimental designs ($t(30) = 3.37$, $p < .01$). That is, experiments received significantly higher experiment fidelity scores than quasi experiments. When we entered all sum scores of the experiment fidelity scale ($M = 3.81$; $SD = 2.63$; *range* = 1 – 8) into a meta-regression analysis (unrestricted maximum likelihood), there was no evidence for significant regression models, implying that children's outcomes are not affected by the experiment fidelity. Due to inadequate descriptive data, we were unable to opt for a more elegant solution by accounting for the nonindependence of observations within schools or classes (Hedges, 2007). Furthermore, a cumulative analysis did not reveal a decreasing effect of the intervention with an increase in publication year.

Results

The final set of intervention studies targeting interactive reading in educational settings comprised 31 studies. Sixteen studies tested at least one print-related skill next to an oral language outcome. In sum, 2,025 children ($N_{Experimental\ Group} = 1{,}016$; $N_{Control\ Group} = 1{,}009$) were studied, with a mean sample size of 65 children (*SD* = 56.10, *range* = 13 – 248). Specifically, 1,030 participants attended day care or preschool programs, and 995 were in kindergarten, of which 1,501 were read to by their teachers and 524 by experimenters. Insofar as articles provided school details (less than half), it appeared that children attended the educational setting at least half a day. Children were exposed to an average of 42.3 interactive reading sessions (*SD* = 33.31, *range* = 4 – 66). For specific study characteristics and unweighted effect sizes per outcome, see Appendix 4.1.

In the first subsection, we examined the additional effects of interactive reading on oral language and print knowledge. Second, we tried to explain the variability in effect sizes on the basis of intervention characteristics. As regards the participants, it should be noted in advance that 27 out of 31 studies targeted students classified as at risk ($n = 1{,}515$), including children qualified for public subsidy of day care costs; attending Head Start, Title I, or similarly funded classrooms; and/or scoring below national norms on early literacy tasks. When we left out the four studies with children not at risk, mean effect sizes were similar to the analyses that comprised all samples (see Table 4.1).

Table 4.1
Mean Effect Sizes of All Six Outcome Measures in the Overall and the At-Risk Sample

	Overall Sample				*At-Risk Sample*			
	k	*d*	95% CI	*N*fs	*k*	*d*	95% CI	*N*fs
Oral Language	31	0.54***	.33, .74	1724	27	0.57***	.36, .78	1448
Expressive Vocabulary	20	0.62***	.29, .95	503	17	0.72***	.33, 1.10	518
Receptive Vocabulary	23	0.45***	.22, .68	463	20	0.48***	.12, .67	335
Print Knowledge								
Alphabet Knowledge	13	0.39**	.16, .62	112	11	0.40**	.12, .67	61
Phonological Sensitivity	13	0.43***	.25, .62	332	11	0.52***	.29, .76	246
Orthographic Awareness	9	0.41***	.20, .62	83	7	0.36**	.10, .62	27

** $p < .01$, ***$p < .001$

Oral Language and Print Knowledge Outcomes

To examine to what extent children's oral language would benefit more from an interactive reading intervention than print-related skills, separate meta-analyses on each outcome were conducted.

The interactive reading intervention had a moderate effect on the oral language skills, explaining 6% of the variance ($d = .54$, $p < .001$; *95%* CI = .33, .74). Expressive vocabulary was especially affected by interactive reading ($k = 20$, $n = 1{,}350$; $d = .62$, $p < .001$; *95%* CI = .29, .95). However, this effect size did not differ significantly from receptive vocabulary because the 95% CIs showed considerable overlap ($k = 23$; $n = 1{,}765$; $d = .45$, $p < .001$; *95%* CI = .22, 68). To overturn these results into non-significant effect sizes a substantial number of missing studies have to be located or executed: For oral language outcomes the fail-safe number equaled 1,724; for expressive vocabulary, 503; and for receptive vocabulary, 463, respectively. In Figure 4.1, the effect sizes for all studies are displayed graphically.

Print knowledge was split into three subcategories: alphabetic knowledge ($k = 13$, $n = 1{,}170$), phonological sensitivity ($k = 13$, $n = 1{,}105$), and orthographic awareness ($k = 9$; $n = 880$). The additional effects of interactive reading on these skills explained about 4% to 5% of the variance and can be interpreted as modest, varying from $d = .39$ for alphabetic knowledge ($p < .01$; *95%* CI = .16, .62; *N*fs = 112) to $d = .43$ for phonological sensitivity ($p < .001$; *95%* CI = .25, .62; *N*fs = 332) and $d = .41$ for orthographic awareness ($p < .001$; *95%* CI = .20, .62; *N*fs = 83).

When the precision (1/SE) was plotted against the standardized difference in means, symmetry around the point estimate appeared to be present for oral language, expressive and receptive vocabulary, alphabetic knowledge, and orthographic awareness. However, a publication bias was detected in the plot reflecting phonological sensitivity outcomes. That is, four studies in the bottom left-hand corner had to be added in order to find symmetry around the point estimate. We used the trim-and-fill method to calculate the effect of this potential data censoring: Trimmed studies were replaced, and their missing counterparts were imputed as mirror images of the trimmed outcomes (Duval & Tweedie, 2000a, 2000b). Adding four studies that appeared to be missing in order to obtain symmetry resulted in an adjusted effect size of $d = .25$ (*95%* CI = .06, .45), suggesting the effect size to be closer to .25 than .43.

In sum, we can accept our first hypothesis: Children's language and print-related skills seem to improve as a result of interactive reading interventions. Second, we wondered whether all skills would be affected equally. Because the 95% CIs of the effect sizes of all outcome measures showed overlap, we had to reject our hypothesis that oral language skills – and expressive vocabulary in particular – would gain the most from the interaction. Although adults were not instructed to comment on letters, phonemes, or writing concepts during storybook reading, the results suggest that being read to interactively can be seen as an incentive for improving both oral language and print knowledge.

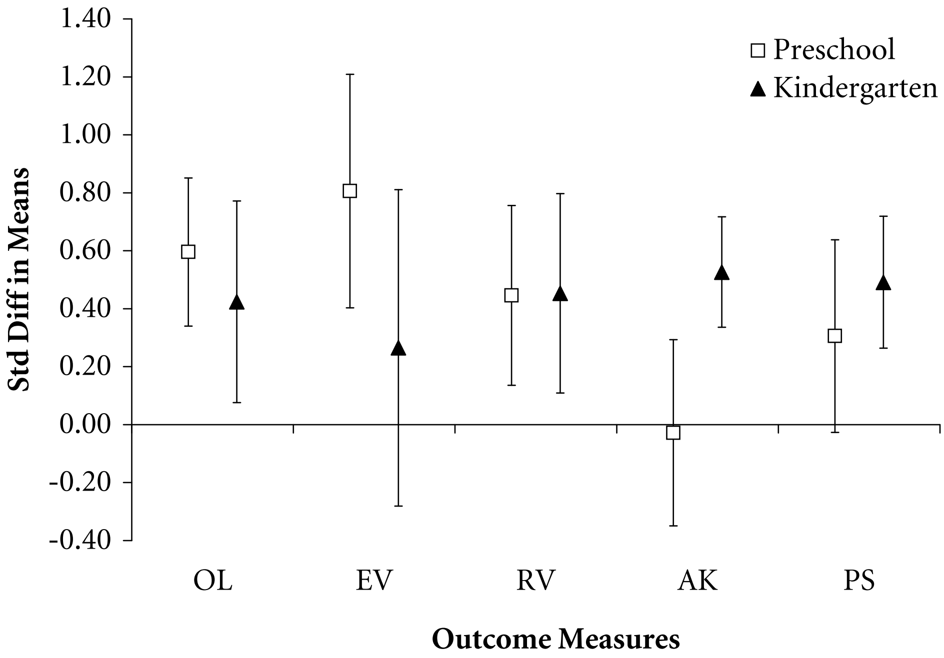
Specifically, we expected that the intervention would not affect all children's print-related skills to the same extent. That is, as children grow older, they might spontaneously pay more attention to print. To test this hypothesis, we compared older kindergarten groups with younger preschool groups. Twenty-one studies

Figure 4.1
Stem-and-Leaf Display of the Effect Sizes per Study at the Posttest on all Outcome Measures

Stem	OL	EV	RV	AK	PS	OA
2.7	1	1				
2.6						
2.5						
2.4						
2.3						
2.2	7	7				
2.1						
2.0		6				
1.9						
1.8						
1.7		5				
1.6						
1.5	7		2			
1.4	6		1		8,8	
1.3	5					
1.2			1			
1.1	8					
1.0	2	7	8		4	
0.9	3	3	2	2		
0.8	4					5
0.7		1	8	3		1
0.6	1,6,9	0	2	0,0,1,2,4,6	6,7	
0.5	3,4	4	2,3,4		3	2,9
0.4	6,7,7	5	6,7		7	
0.3	4,5	4			7	5,9
0.2	0,5	4	4	5	8,9	9
0.1	0,9	5	5,6,6		4	
0.0	4	4,6,9	3,6	7,8	2	
-0.0	8	7	1,4			
-0.1	3,3,3	6,8	0,9			0
-0.2	1			6		1
-0.3	6			3	0	
-0.4						
-0.5	8		1			
-0.6		1				

Note. OL = Oral Language (*k* = 31 studies, *n* = 2,025 children); *EV* = Expressive Vocabulary (*k* = 20, *n* = 1,350); *RV* = Receptive Vocabulary (*k* = 23; *n* = 1,765); *AK* = Alphabet Knowledge (*k* = 13, *n* = 1,170); *PS* = Phonological Sensitivity (*k* = 13, *n* = 1,105); *OA* = Orthographic Awareness (*k* = 9; *n* = 880) implemented the interactive reading intervention in preschools, versus ten in kindergarten classrooms. No statistically significant age differences were detected on oral language and phonological sensitivity outcomes ($Q_{Oral\ Language}(1) = .61$, $p > .05$; $Q_{Expressive\ Vocabulary}(1) = 2.44$, $p > .05$; $Q_{Receptive\ Vocabulary}(1) = .001$, $p > .05$; $Q_{Phonological\ Sensitivity}(1) = .82$, $p > .05$). Interestingly, a moderate effect of the intervention was found for the alphabetic knowledge of children in kindergarten classrooms ($Q(1) = 8.47$, $p < .01$; $k = 8$, $d = .53$, *95%* CI = .34, .72), whereas children in preschool showed no growth at all ($k = 5$, $d = -.03$, *95%* CI = -.35, .29). In Figure 4.2, this contrast is displayed graphically. When the sample was restricted to at-risk groups, we again found a significant age effect for alphabetic knowledge ($k_{\text{kindergarten}} = 6$, $k_{\text{preschool}} = 5$; $Q(1) = 22.25$, $p < .001$) but not for the other outcome variables. In short, only kindergarten children showed a statistically significant growth in alphabetic knowledge.

Figure 4.2
Paired Comparison Chart of Preschool versus Kindergarten Children on all Outcome Measures, with Significant Age Differences on Alphabet Knowledge.



Note. OL = Oral Language; *EV* = Expressive Vocabulary; *RV* = Receptive Vocabulary; *AK* = Alphabet Knowledge *; PS* = Phonological Sensitivity

Intervention Characteristics as Moderators

To explain the variability in effect sizes, moderator analyses were conducted for five intervention characteristics: The adult who carried out the intervention, group size, type of intervention program, activities in the control group, and duration of the intervention.

First, we tested what was more effective: to be read to by a teacher or an experimenter. The random effects models were significant for the oral language composite ($Q(1) = 4.24$, $p < .05$) and expressive vocabulary ($Q(1) = 6.02$, $p < .05$), implying that experimenters such as the researcher or a graduate student were more effective than teachers (Oral Language: $k_{teacher} = 16$, $d = .35$, *95%* CI = .08, .62; $k_{experimenter} = 15$, $d = .79$, *95%* CI = .47, 1.10; Expressive Vocabulary: $k_{teacher} = 11$, $d = .28$, *95%* CI = .14, .69; $k_{experimenter} = 15$, $d = 1.10$, *95%* CI = .59, 1.56). It should be noted, however, that all but one study reported about teachers who interacted with children in small ($k = 5$, $n = 184$) or large groups ($k = 10$; $n = 1{,}295$), whereas experimenters read to children one-to-one ($k = 5$; $n = 145$) or in small ($k = 6$; $n = 299$) and large groups ($k = 4$; $n = 80$). When we examined both moderators simultaneously to find out which combination would be most effective ($Q(4) = 12.36$, $p < .05$), experimenters interacting with individual children seemed to have the strongest impact on children's oral language skills ($d = 1.38$, *95%* CI = .86, 1.89), differing significantly from the small to moderate effects that teachers revealed by reading to small groups ($Q(1) = 15.61$, $p<001$; $d = .15$, *95%* CI = -.24, .54) and large groups ($Q(1) = 8.69$, $p < .01$; $d = .48$, *95%* CI = .19, .78), as well as experimenters in large groups ($Q(1) = 5.36$, $p < .05$; $d = .34$, *95%* CI = -.28, .95). Insofar as sufficient studies were available with expressive vocabulary as a dependent measure, results were similar (see Table 4.2). No statistically significant differences were found for phonological sensitivity ($k_{teacher\text{-}large\ groups} = 6$, $k_{experimenter\text{-}small\ groups} = 5$; $Q(1) = 4.59$, $p > .05$).

Furthermore, the studies could be divided into three categories: 8 studies ($n = 260$) implemented Dialogic Reading (DR) as developed by Whitehurst et al. (1988), 11 studies ($n = 411$) tested the effects of similar techniques without referring to the specific Dialogic Reading-format and were coded as interactive reading (IR), and another 12 studies ($n = 1{,}354$) included extra classroom activities to support the interactive reading sessions (IR+). For the oral language composite, the random effects analysis revealed a statistically significant difference among the groups ($Q(2) = 11.38$, $p < .01$). DR was the least effective program ($d = .24$, *95%* CI = -.17, .64), differing significantly from the rather strong effect revealed by IR ($Q(1) = 7.54$, $p < .01$; $d = 1.01$, *95%* CI = .64, 1.39). IR also differed significantly from the IR+ programs which had a surprisingly low average effect size ($Q(1) = 7.95$, $p < .01$; $d = .38$, *95%* CI = .10, .66). As can be seen in Table 4.2, the patterns were similar for expressive and receptive vocabulary. In contrast to Karweit and Wasik's hypothesis (1996), the additional classroom activities offered by IR+ programs did not improve children's language skills more than single interactive reading sessions. A new perspective opened up, however, when the adult that carried out the intervention was taken into account as an additional moderator. First, IR+ programs were implemented by teachers in all but 1 study, whereas children in the IR condition were read to by experimenters in 10 out of 11 studies. Second, in the DR program, half of the studies were executed by teachers ($k_{teacher} = 4$, $n = 128$; $k_{experimenter} = 4$, $n = 132$), who tended to be less effective than experimenters ($Q(1) = 3.06$, $p = .08$): Experimenters seemed to be moderately effective in eliciting gains in oral language skills whereas teachers trained in dialogic reading techniques did not reveal effects ($d_{teacher} = -.08$, *95%* CI = -.59, .43; $d_{experimenter} = .58$, *95%* CI = .04, 1.11). When only experimenters were selected ($k = 14$, $n = 465$), the effect size differences between DR and IR were no longer significant ($Q(1) = .78$, $p > .05$). Finally, statistically significant differences were found in experiment fidelity scores across programs ($F(2, 28) = 10.79$, $p < .001$), with IR scoring significantly higher in fidelity than the studies that tested the effects of DR and IR+ ($M_{\text{IR}} = 6.00$, $SD = 2.37$; $M_{\text{DR}} = 3.38$, $SD = 2.45$; $M_{\text{IR+}} = 2.08$, $SD = 1.31$), implying that the better controlled experiments were implemented by experimenters.

Overall, the country and/or language did not explain any variability in the effects ($k_{English/US} = 21$, $n = 1{,}125$; $Q_{Oral\ Language}(1) = 1.14$, $p > .05$; $Q_{Alphabet\ Knowledge}(1) = .95$, $p > .05$; $Q_{Phonological\ Sensitivity}(1) = .001$, $p > .05$). Unfortunately, all experiments with control groups that received an intervention were conducted in the United States and in English ($k = 11$, $n = 675$), whereas studies that included control group children who received the standard school program were conducted in both English ($k = 10$, $n = 450$) and other languages such as Dutch, Hebrew, Portuguese, or Spanish ($k = 10$; $n = 900$). Significant group differences in the activity by the control group were present for all oral language outcomes ($Q_{Oral\ Language}(1) = 9.82$, $p < .01$; $Q_{Expressive\ Vocabulary}(1) = 9.08$, $p < .01$; $Q_{Receptive\ Vocabulary}(1) = 8.42$, $p < .01$). As can be seen in Table 4.2, studies that included a control group that was only pre- and posttested revealed significantly lower effect sizes for the oral language outcomes than studies in which the control-group children were part of a non language-related intervention. This suggests that more elegantly designed studies with a higher fidelity score revealed higher effect sizes ($t(30) = 5.02$, $p < .001$).

We used a 16-week cutoff to be close to an intervention of at least half a 10-month school year as the contrasts could not have been tested when we split at 5 months. Interestingly, children's oral language and alphabetic knowledge did not seem to be influenced by the duration of the interactive reading intervention ($Q_{Oral\ Language}(1) = 1.53$, $p > .05$; $Q_{Expressive\ Vocabulary}(1) = 1.67$, $p > .05$; $Q_{Recpetive\ Vocabulary}(1) = .27$, $p > .05$; $Q_{Alphabet\ Knowledge}(1) = .06$, $p > .05$), whereas phonological sensitivity skills significantly improved as the duration of the intervention increased ($Q(1) = 4.85$, $p < .01$). That is, interventions that were implemented during a short period ($M_{weeks} = 11.33$, $SD = 5.16$; $M_{sessions} = 27.17$, $SD = 6.40$) had a smaller effect on phonological sensitivity than interventions that were spread over 4 months to a school year ($k_{short} = 6$, $d = .21$, *95%* CI = -.04, .46; $k_{long} = 7$, $d = .60$, *95%* CI = .36, .83).

Table 4.2
Meta-Analytic Results per Outcome Measure

	k	*n*	*d*	95% CI	*Q* [a]	*p*
Oral Language						
Total Set	**31**	**2,025**	**.54*****	**.33, .74**	**318.15**	**< .001**
Design					1.69	.19
Experiment	19	671	.66***	.38, .93		
Quasi-Experiment	12	1,354	.38*	.07, .69		
School Type					.61	.44
Preschool	21	1,030	.60***	.34, .85		
Kindergarten	10	995	.42*	.08, .77		
Experimenter * Group size					12.36	.02
Experimenter-Individual	5	145	1.38***	.86, 1.89		
Experimenter-Small Group	6	299	.59*	.13, 1.05		
Experimenter-Large Group	4	80	.34	-.28, .95		
Teacher-Small Group	5	184	.15	-.24, .54		
Teacher-Large Group	10	1,295	.48***	.19, .78		
Type of Intervention Program					11.38	.003
DR	8	260	.24	-.17, .64		
IR	11	411	1.01***	.64, 1.39		
IR+	12	1,354	.38**	.10, .66		
Program Type * Experimenter					3.06	.08
DR-Teacher	4	128	-.08	-.59, .43		
DR-Experimenter	4	132	.58*	.04, 1.11		
Activity Control Group					9.82	.002
Intervention	11	675	.95***	.63, 1.27		
No intervention	20	1,350	.34**	.12, .55		
Duration Intervention					1.53	.22
Short (<16 weeks)	20	1,002	.64***	.38, .89		
Long	11	1,023	.37*	.03, .70		
Expressive Vocabulary						
Total Set	**20**	**1,350**	**.62*****	**.29, .95**	**212.00**	**.00**
Design					.41	.52
School Type					2.44	.12
Experimenter * Group size					8.02	.046
Experimenter-Individual	5	145	1.40**	.56, 2.23		
Experimenter-Small Group	4	223	.73	-.16, .62		
Teacher-Large Group	7	1,011	.47	-.08, 1.02		
Type of Intervention Program					9.29	.01
DR	7	243	.20	-.30, .70		
IR	6	253	1.36***	.78, 1.94		
IR+	7	1,011	.47	-.01, .95		
Activity Control Group					9.08	.003
Intervention	7	537	1.26***	.74, 1.77		
No intervention	13	970	.29	-.07, .65		
Duration Intervention					1.67	.197

	k	*n*	*d*	*95% CI*	*Q* [a]	*p*
Receptive Vocabulary						
Total Set	**23**	**1,765**	**.45*****	**.22, .68**	**121.91**	**< .001**
Design					.08	.78
School Type					.001	.98
Experimenter * Group size					1.15	.29
Type of Intervention Program					6.32	.04
DR	7	243	.17	-.24, .58		
IR	5	173	.98***	.49, 1.48		
IR+	11	1,295	.42**	.15, .69		
Activity Control Group					8.42	.004
Intervention	5	444	.92***	.56, .128		
No intervention	18	1,267	.31**	.11, .51		
Duration Intervention					.27	.61
Alphabet Knowledge						
Total Set	**13**	**1,170**	**.39****	**.16, .62**	**40.28**	**< .001**
Design					.02	.88
School Type					8.47	.004
Preschool	5	269	-.03	-.35, .29		
Kindergarten	8	901	.53***	.34, .72		
Duration Intervention					.06	.81
Phonological Sensitivity						
Total Set	**13**	**1,105**	**.43*****	**.25, .62**	**77.92**	**84.60**
Design					.81	.37
School Type					.82	.37
Experimenter * Group size					4.59	.10
Duration Intervention					4.85	.028
Short (<16 weeks)	6	413	.21	-.04, .46		
Long	7	699	.60***	.36, .83		
Orthographic Awareness						
Total Set	**9**	**880**	**.41****	**.20, .62**	**19.95**	**.01**

Note. k = number of studies; n = total number of participants; *95%* CI = confidence interval; [a]Q for subset stands for homogeneity ($df = k - 1$); Q for moderator stands for effects of contrasts (df = number of subsets – 1). Contrasts were not tested when $k<4$ studies. Except for the Oral Language Composite, point estimates were not presented when the $Q_{between}$ was not significant ($p > .05$); * $p < .05$, ** $p < .01$, ***$p < .001$

Discussion

This meta-analysis tested the effects of an intervention that is thought to enhance the quality of adult-child storybook reading in early education and is expected to foster children's language and literacy development as a consequence. Results showed that children's oral language as well as print knowledge benefited from interaction before, during and after shared reading sessions. That is, about 6% of the growth in oral language skills could be explained by an interactive reading intervention in an educational setting (r = .25). Because the program focuses on techniques such as eliciting and reinforcing verbal responses by the child, it seems likely that children's expressive vocabulary skills will benefit most. Indeed, a moderate effect size was found for expressive vocabulary, explaining 8% of the variance (r = .28). These results indicate that the quality of book reading is important in addition to its frequency (Bus et al., 1995; Scarborough & Dobrich, 1994). The findings could not be attributed to differences in design characteristics or publication biases. When translated into a binominal effect size display or a change in success ratio, the oral language of children exposed to an interactive reading program gained 28% more than their peers in a control group, meaning that with interaction 64% improved in oral language, compared to 36% of children who were not part of the intervention. As in the medical domain when drugs are prescribed to millions of people because of a difference as small as 3% between the control and intervention groups (see Bus, 2001; Bus & Van IJzendoorn, 2004), it can be argued that interactive reading in early education warrants implementation.

Although adults were not trained to refer to print, 7% of the variance in kindergarten children's alphabetic knowledge could be explained by the interactive reading program. As expected, older children were able to significantly expand their emergent alphabetic knowledge, whereas younger children's print knowledge hardly benefited from interactive storybook encounters. Phonological sensitivity improved from interventions that were spread over a longer period of time, such as a school year. Following Lonigan (2006), we may define this result as a "dissociation effect" (p. 85). He demonstrated that an oral language intervention significantly affected measures of rhyme and blending. Albeit the studies do not provide data related to the qualities of the interactions during reading, it is conceivable that kindergarten teachers made more references to print than preschool teachers and/or that children with some knowledge of print may have elicited discussion of print features. Alternatively, a storybook itself might emphasize print and enhance print knowledge by varying font types and sizes, displaying some utterances in text balloons, or using rhyme and alliterations (Justice & Lankford, 2002). Unfortunately, hardly any information was provided about print-salient features within the storybooks that were used in the intervention studies. We speculate that children's ability to divide their attention between an adult and a book increases with growing experience in comprehending and interpreting a story's content. As children grow older, they might have control of skills to explore and process other features of the printed text, such as single letters, while listening to and interacting with an adult at the same time, whereas younger children need to invest all efforts in understanding the story. This assumption seems to contradict eye-tracking research by Evans and Saint-Aubin (2005) and Justice et al. (2005), but more recent studies by the same group of authors (Justice, Pullen, & Pence, 2008; Roy-Charland et al., 2007) show that the degree to which children learn about print during book reading depends on the extent that adults (non) verbally refer to print and the materials' characteristics. Assuming that the input of children and their environment affect each other reciprocally, we propose a transactional model of book reading to explain that not only does the interactive reading style of the teacher affect learning but the child's role is important as well (Sameroff & Fiese, 2000).

Differential effects of reading experience might also be reflected in the number of repetitions of the same book. Familiarity with the story content due to repeated readings of one story may create new opportunities to shift children's attention to other features of the text, as G. Phillips and McNaughton (1990) suggested based on a series of case studies. Unfortunately, these hypotheses could not be tested in the current meta-analysis, as almost all studies that included print-related measures reread a storybook at least once. Next to quality-related explanations, it seems likely that both language skills and print knowledge affect each other reciprocally (e.g., NICHD, 2005; Poe, Burchinal & Roberts, 2004; Samuelsson et al., 2007; Speece, Ritchey, Cooper, Roth, & Schatschneider, 2004; Storch & Whitehurst, 2002). The lexical restructuring hypothesis assumes that remembering and recognizing words in smaller segments become more efficient as children acquire more and more words via spoken language experiences (Metsala & Walley, 1998). On the other hand, the so-called Jabberwocky effect implies that phonological sensitivity stimulates vocabulary knowledge. As in Lewis Carroll's poem, phonemes carry meaning: Children who know more about units of words can use that knowledge to tune to parts of new words that have meaning for them and expand their vocabulary (Dickinson, McCabe, Anastasopoulos, Peisner-Feinberg, & Poe, 2003). Further research is needed, though, to understand the underlying processes that might explain the additional effect of interactive reading on print knowledge.

Based on a meta-analysis of 31 experiments that study different designs and populations, we can make some propositions for interventions that work best (Lipsey & Wilson, 2001). First, most explicit effect sizes appeared to be present in experiments that were highly controlled and executed by examiners. Teachers seemed to have difficulty with fostering the same growth in young children's language and literacy skills as researchers. It can be hypothesized that teachers were not successful in incorporating and internalizing the novel strategies in line with the intentions of the program developers (Dickinson & Sprague, 2001). Studying the factors that promote the transfer of evidence-based interventions to realworld settings is important, though. A cost-effective potential for generalization to future cohorts as well as the integration within the regular curriculum will widen the scope of the intervention (Aram, 2006). To promote treatment fidelity, it may be critical that the social component of the implementation process is emphasized with several opportunities for feedback and positive reinforcement next to the training in more technical aspects such as the theory behind the intervention (Rohrbach et al., 2006; Shernoff & Kratochwill, 2007). Even though the techniques developed by Whitehurst et al. (1988) are most tried and improved, the lowest effects were found for this program. It should be noted here that Dialogic Reading was implemented by showing and discussing a videotape in a session before and halfway through the program. Less standardized interactive reading interventions may incorporate more opportunities to coach teachers, discuss and solve concrete problems, and adapt the program to the needs of a specific classroom. Furthermore, in the current meta-analysis, teachers participated mostly in interventions that affected all kinds of classroom activities. It seems plausible that investing in play, art, or drama activities might have distracted teachers from giving as much attention to the interactive storybook reading as the researchers had anticipated, or as is evident for children who are part of a single interactive reading program. However, the above speculations provide only a partial solution to the interpretation of the effect sizes. The hypotheses can only be confirmed when experimenters and teachers are contrasted across program types. For instance, it should be tested whether lower effect sizes will also be found when experimenters implement IR+ programs or when teachers carry out an IR-only program. Varying the intensity of the intervention will be interesting as well, as will be exploring the frequency with which a book should be repeated to benefit optimally from the interaction. In sum, the current findings raise important questions to investigate through further research.

Second, we did not find support for Karweit and Wasik's (1996) conclusion that teachers should read to small groups when it is feasible to do so; neither did the results confirm Dickinson and Sprague's (2001) assumption that it is too challenging for teachers to tune to children's developmental level while reading to whole classrooms. In fact, this meta-analysis demonstrated that children's skills improved when their teachers engaged them in whole-group interactive reading sessions. Teachers might feel more inclined to focus their questions on events that are directly related to the book in order to keep control over the reading session and help the children to focus on the story's meaning and vocabulary as a consequence (Dickinson & Smith, 1994; Karweit & Wasik, 1996). In small group settings, the amount of extraneous talk and the opportunity for each child to elaborate on his or her own experiences might be distracting. Taken together, the quantitative summary of the research base thus far shows that interactive reading is worth implementing in classroom settings. To bridge the gap between research and practice, however, more research is needed to investigate the most promising implementation strategies.

Third, as groups at risk will be more susceptible to and in need of stimulation at school (Mol et al., 2008), they seem to be a promising group to invest in. All but four studies included children at risk for language and literacy impairments.

Cautions and Limitations

The result of a meta-analysis can be only as good as the studies that are included. Due to a lack of detail in the description of the intervention and its application in some studies, it was hard to extract data on all the characteristics that we were interested in. For instance, researchers reported the number of sessions they intended the teachers to offer to the children, but they did not seem to have observed whether this intensity was actually realized. We could not disentangle whether the duration of the intervention was a real confounder because the interactive reading programs with additional classroom activities were spread over a longer period than the dialogic reading interventions. Future studies should consider Lonigan and Whitehurst's (1998) approach: They presented separate and significantly different data for day care centers that reported to have held frequent and infrequent reading sessions, categorized as high- versus low-compliant centers. Overall, most studies seemed to lack control over the quality and frequency of book reading in control groups, especially when control-group children did not receive an intervention. This might restrict our conclusions regarding the unique effects of interactive reading in the classroom.

Because only half of the studies included a measure of print knowledge, most of the moderators could be analyzed only for oral language outcomes. Besides, studies that presented not raw posttest means and standard deviations but scores that took into account the pretest scores revealed stronger effects. Although none of the studies reported that their groups differed on the pretest, it seems tenable that a child's initial skills affect and account for his or her learning potential. Finally, the current findings are not completely independent from the meta-analysis on dialogic reading in parent-child dyads (see Mol et al., 2008). For two studies (Crain-Thoreson & Dale, 1999; Lonigan & Whitehurst, 1998), we included the same control-group posttest data as Mol et al. (2008).

Practical Implications

An interactive exposure to storybooks can be considered as an effective stimulant for the development of two pillars of learning to read: oral language and print knowledge. The current meta-analysis showed that interactive qualities of book reading in classrooms are effective supplements to book reading (Bus et al., 1995; Scarborough & Dobrich, 1994). Teachers who read to whole groups and accompanied the storybooks with extra activities knew to elicit moderate effects in oral language and print knowledge. The added value of interactive reading was reflected best in children who individually interacted with experimenters. Although the included studies did not provide enough details to grasp exactly what happened during the interactive reading sessions, it seems evident that children had a chance to learn about the story language as well as the written format of read-aloud texts. As program type and experimenter appeared to be interrelated in the current meta-analysis, more research is needed to disentangle the specific effects for interventions with and without additional activities that are implemented by experimenters versus teachers. Contrasting different implementation strategies to enhance effectiveness might be helpful as well. Compared to a short training by videotape, closely monitoring or coaching teachers might yield better opportunities to internalize a program's principles and adapt the trained techniques to the developmental level of the children in a classroom. Future observations are needed to explain our finding that interactive reading also affects children's print knowledge: To what extent do adults use the reading sessions to teach letters and sounds, do print-salient features attract attention, or do older children spontaneously pay attention to print and expand their skills by themselves when they grow more knowledgeable?

Appendix

Appendix 4.1

Study, Population, and Intervention Characteristics per Study Included in the Meta-Analysis.

First Author	Year	Publ. Status[a]	N_{tot} (n_{int} & n_{contr})[b]	Design[c]	Country (Language)	School Type[d]	At Risk	Interv. Program[e]	Intervention ContrGr	Adult ($n_{readers}$ in Int.Gr)	Group Size (n_{childr})	Duration in weeks ($n_{sessions}$)[f]	Outcome[g]	ES (d)[h]	SE
Aram-st1[i]	2006	Publ.	40 (20 & 20)	QE	Israel (Hebrew)	PreS	Yes	IR+	No	Teacher (6)	Small (4-6)	28 (M = 50)	*OL-RV*	0.47	0.23
													AK	0.07	0.23
													PS	0.67	0.32
													OA	-0.21	0.23
Aram-st2[j]	2006	Publ.	38 (17 & 21)	QE	Israel (Hebrew)	K	Yes	IR+	No	Teacher (6)	Small (4-6)	28 (M = 50)	*OL-RV*	0.53	0.24
													AK	0.61	0.24
													PS	1.48	0.37
													OA	0.59	0.24
Aram	2004	Publ.	59 (35 & 24)	QE	Israel (Hebrew)	PreS	Yes	IR+	No	Experimenter (2)	Small (4-6)	28 (M = 66)	*OL*	-0.36	0.27
													PS	0.37	0.27
													OA	0.39	0.19
Brabham	2002	Publ.	78 (39 & 39)	E	U.S. (English)	G1	No	IR	Yes	Teacher (5)	Large (6-9)	2 (R = 6)	*OL*	1.02	0.12
													RV	1.52	0.18
Crain-Thoreson	1999	Publ.	22 (13 & 9)	E	U.S. (English)	PreS	Yes	DR	No	Teacher (8)	Individual	8 (R = 32)	*OL*	-0.13	0.25
													EV	-0.18	0.31
													RV	-0.04	0.43
de Oliviera Fontes	2004	Publ.	38 (19 & 19)	E	Brasil (Portuguese)	K	Yes	IR	No	Experimenter (1)	Small (2-4)	16 (M = 13.8)	*OL*	0.84	0.17
													EV	0.60	0.33
													RV	0.92	0.34
													AK	0.25	0.33
													PS	0.66	0.33
													OA	-0.10	0.32
Droop-st1[k]	2005	Publ.	95 (42 & 53)	QE	Netherlands (Dutch)	K	Yes	IR+	No	Teacher (11)	Large (whole class)	16 (M = 32)	*OL*	0.66	0.15
													EV	0.71	0.21
													RV	0.62	0.21
													AK	0.66	0.21
													PS	0.29	0.09
													OA	0.85	0.22

Droop-st2[k]	2005	Publ.	168 (91 & 77)	QE	Netherlands (Dutch)	K	No	IR+	No	Teacher (11)	Large (whole class)	16 (M = 32)	*OL*	-0.13	0.11
													EV	-0.16	0.16
													RV	-0.10	0.15
													AK	0.08	0.15
													PS	0.02	0.07
													OA	0.35	0.16
Droop-st3[l]	2005	Publ.	248 (94 & 154)	QE	Netherlands (Dutch)	K	No	IR+	No	Teacher (17)	Large (whole class)	20 (M = 40)	*OL*	-0.13	0.09
													EV	-0.07	0.13
													RV	-0.19	0.13
													AK	0.64	0.13
													PS	0.28	0.06
													OA	0.71	0.13
Droop-st4[m]	2005	Publ.	81 (39 & 42)	QE	Netherlands (Dutch)	K	Yes	IR+	No	Teacher (17)	Large (whole class)	20 (M = 40)	*OL*	0.20	0.16
													EV	0.24	0.22
													RV	0.16	0.22
													AK	0.73	0.23
													PS	0.47	0.10
													OA	0.29	0.22
Droop-st5[k]	2005	Publ.	113 (62 & 51)	QE	Netherlands (Dutch)	K	Yes	IR+	No	Teacher (17)	Large (whole class)	20 (M = 40)	*OL*	0.35	0.14
													EV	0.15	0.19
													RV	0.54	0.19
													AK	0.60	0.19
													PS	0.53	0.09
													OA	0.52	0.19
Karweit-st1[i]	1989	Publ.	86 (43 & 43)	QE	U.S. (English)	PreS	Yes	IR+	No	Teacher (2)	Large (whole class)	28 (R = 56)	*OL*	0.47	0.11
													RV	0.52	0.22
Karweit-st2[j]	1989	Publ.	120 (60 & 60)	QE	U.S. (English)	K	Yes	IR+	No	Teacher (2)	Large (whole class)	28 (R = 56)	*OL*	0.46	0.09
													RV	0.24	0.18
													AK	0.62	0.19
													PS	1.04	0.19
Kertoy	1994	Publ.	16 (8 & 8)	E	U.S. (English)	K	No	IR	Yes	Experimenter (2)	Individual	3 (R = 4)	*OL-EV*	0.54	0.51

First Author	*Year*	*Publ. Status* [a]	N_{tot} (n_{int}&n_{contr})[b]	*Design*[c]	*Country (Language)*	*School Type*[d]	*At Risk*	*Interv. Program*[e]	*Interven tion ContrGr*	*Adult* ($n_{readers}$ *in Int.Gr*)	*Group Size* (n_{childr})	*Duration in weeks* ($n_{sessions}$)[f]	*Outcome*[g]	*ES* (d)[h]	*SE*
van Kleeck	2006	Publ.	30 (15 & 15)	E	U.S. (English)	PreS	Yes	IR	No	Experimenter (*N* not rep.)	Individual	8 (*R* = 16)	*OL*	1.18	0.23
													EV	1.07	0.28
													RV	1.41	0.41
Lamb-st1[n]	1986	Diss.	14 (5 & 9)	E	U.S. (English)	PreS	Yes	IR	Yes	Experimenter (1)	Large (10)	10 (*M* = 50)	*OL*	-0.21	0.40
													RV	-0.01	0.56
													AK	-0.26	0.56
Lamb-st2[o]	1986	Diss.	13 (5 & 8)	E	U.S. (English)	PreS	Yes	IR	No	Experimenter (1)	Large (10)	10 (*M* = 50)	*OL*	-0.08	0.41
													RV	0.16	0.57
													AK	0.92	0.60
Lonigan-st1[p]	1998	Publ.	27 (16 & 11)	E	U.S. (English)	PreS	yes	DR	No	Teacher (*N* not rep.)	Small (5)	6 (*M* = 6.8)	*OL*	-0.58	0.23
													EV	-0.61	0.28
													RV	-0.51	0.40
Lonigan-st2[q]	1998	Publ.	31 (15 & 16)	E	U.S. (English)	PreS	Yes	DR	No	Teacher (*N* not rep.)	Small (5)	6 (*M* = 17.4)	*OL*	0.25	0.21
													EV	0.34	0.26
													RV	0.06	0.36
Lonigan-st1[r]	1999	Publ.	46 (17 & 29)	E	U.S. (English)	PreS	Yes	DR	Yes	Experimenter (*N* not rep.)	Small (2-5)	6 (*M* = 18.5)	*OL*	0.19	0.18
													EV	0.06	0.22
													RV	0.46	0.31
													PS	-0.30	0.16
Lonigan-st2[s]	1999	Publ.	49 (17 & 32)	E	U.S. (English)	PreS	Yes	DR	No	Experimenter (*N* not rep.)	Small (2-5)	6 (*M* = 18.5)	*OL*	0.04	0.17
													EV	0.04	0.21
													RV	0.03	0.30
													PS	0.14	0.16
Mautte-st1[n]	1990	Diss.	29 (9 & 20)	E	U.S. (English)	PreS	Yes	IR	Yes	Experimenter (1)	Large (7)	20 (*M* = 60)	*OL*	0.34	0.40
Mautte-st2[o]	1990	Diss.	24 (9 & 15)	E	U.S. (English)	PreS	Yes	IR	No	Experimenter (1)	Large (7)	20 (*M* = 60)	*OL*	1.46	0.40
Morrow-st1[t]	1988	Publ.	39 (25 & 14)	E	U.S. (English)	PreS	Yes	IR	Yes	Experimenter (*N* not rep.)	Individual	10 (*M* = 10)	*OL-EV*	2.71	0.47
Morrow-st2[u]	1988	Publ.	40 (27 & 13)	E	U.S. (English)	PreS	Yes	IR	Yes	Experimenter (*N* not rep.)	Individual	10 (*M* = 10)	*OL-EV*	0.93	0.43

Morrow	1989	Publ.	90 (45 & 45)	E	U.S. (English)	PreS	Yes	IR	Yes	Experimenter (*N* not rep.)	Small (3)	11 (*M* = 11)	*OL-EV*	2.27	0.45
Murphy	2007	Diss.	17 (8 & 9)	E	U.S. (English)	PreS	Yes	DR	Yes	Experimenter (1)	Small (3)	8 (*M* = 23)	*OL*	0.69	0.35
													AK	0.60	0.50
													PS	1.48	0.27
Valdez-Menchaca	1992	Publ.	20 (10 & 10)	E	Mexico (Spanish)	PreS	Yes	DR	No	Experimenter (1)	Individual	7 (*M* = 30)	*OL*	1.57	0.40
													EV	1.75	0.26
													RV	1.21	0.50
Wasik	2001	Publ.	121 (61 & 60)	QE	U.S. (English)	PreS	Yes	IR+	Yes	Teacher (2)	Large (whole class)	15 (*R* = 45)	*OL*	1.35	0.49
													EV	2.06	0.12
													RV	1.08	0.30
Wasik	2006	Publ.	185 (124 & 61)	QE	U.S. (English)	PreS	Yes	IR+	Yes	Teacher (10)	Large (whole class)	36 (*R* = 72)	*OL*	0.61	0.14
													EV	0.45	0.16
													RV	0.78	0.11
													AK	-0.33	0.22
Whitehurst	1994	Publ.	48 (26 & 22)	E	U.S. (English)	PreS	Yes	DR	No	Teacher (*N* not rep.)	Small (5)	6 (*M* = 16.6)	*OL*	0.10	0.16
													EV	0.09	0.15
													RV	0.15	0.15

Notes. [a] Publ. Status: Publ. = Published, Diss. = Dissertation; [b] Sample sizes were adapted if studies included more than one intervention or control group and all groups met the inclusion criteria. [c] Design: QE = Quasi-Experiment, E = Experimental Study; [d] School Type at the start of the study: PreS = Preschool, K = Kindergarten, G1 = Grade 1; [e] Type of Intervention Program: DR = Dialogic Reading, IR = Interactive Reading, IR+ = Interactive Reading with Additional Activities; [f] Duration intervention in weeks: the number of sessions is presented between brackets. M = Mean number of sessions as reported in the article; R = Recommended number of sessions as proposed in the article when no observed Mean is provided; In the analyses, interventions less than 16 weeks, were treated as "short"; [g] Outcome: OL = Oral Language Composite, RV = Receptive Vocabulary, OL-EV = Oral Language Measure represents Expressive Vocabulary, OL-RV = Oral Language Composite reflects Receptive Vocabulary, AK = Alphabetic Knowledge, PS = Phonological Sensitivity, OA = Orthographic Awareness; [h] Effect Sizes were based on posttest mean and standard deviations of an experimental and control group, weighted by sample size. A positive sign indicates a favorable outcome for the experimental group; [i] Preschool (sub)sample; [j] Kindergarten (sub)sample; [k] Second Language Learners, at risk; [l] First Language Learners, not at risk; [m] First Language Learners, at risk; [n] interactive reading experimental group versus a typical reading control group; [o] interactive reading experimental group versus a no-intervention control group; [p] Low-compliance centers; [q] High-compliance centers; [r] Dialogic Reading experimental group versus a typical reading control group; [s] Dialogic Reading experimental group versus a no-intervention control group; [t] repeated reading experimental group versus control group; [u] different books reading group versus control group.

References

References marked with an asterisk (*) indicate studies included in the meta-analysis.

Adams, M. J. (1990). *Beginning to read.* Cambridge, MA: MIT Press.

* Aram, D. (2006). Early literacy interventions: The relative roles of storybook reading, alphabetic activities, and their combination. *Reading and Writing, 19,* 489-515.

* Aram, D., & Biron, S. (2004). Joint storybook reading and joint writing interventions among low SES preschoolers: Differential contributions to early literacy. *Early Childhood Research Quarterly, 19,* 588-610.

Ard, L. M., & Beverly, B. L. (2004). Preschool word learning during joint book reading: Effect of adult questions and comments. *Communication Disorders Quarterly, 26,* 17-28.

Bakermans-Kranenburg, M. J., Van IJzendoorn, M. H., & Juffer, F. (2003). Less is more: Meta-analyses of sensitivity and attachment interventions in early childhood. *Psychological Bulletin, 129,* 195-215.

Beck, I. L., & McKeown, M. G. (2007). Increasing young low-income children's oral vocabulary repertoires through rich and focused instruction. *Elementary School Journal, 107,* 251-273.

Biemiller, A., & Boote, C. (2006). An effective method for building meaning vocabulary in primary grades. *Journal of Educational Psychology, 98,* 44-62.

Blok, H. (1999). Reading to young children in educational settings: A meta-analysis of recent research. *Language Learning, 49,* 343-371.

Bodrova, E., & Leong, D. J. (2006). Vygotskian perspectives on teaching and learning early literacy. In D. K. Dickinson & S. B. Neuman (Eds.), *Handbook of early literacy research* (Vol. 2, pp. 243-256). New York: The Guilford Press.

Borenstein, M., Hedges, L., Higgins, J., & Rothstein, H. (2005). *Comprehensive meta-analysis version 2* [Computer Software]. Englewood, NJ: Biostat.

* Brabham, E. G., & Lynch-Brown, C. (2002). Effects of teachers' reading-aloud styles on vocabulary acquisition and comprehension of students in the early elementary grades. *Journal of Educational Psychology, 94,* 465-473.

Bus, A. G., & van IJzendoorn, M. H. (1988). Mother-child interactions, attachment, and emergent literacy: A cross-sectional study. *Child Development, 59,* 1262-1272.

Bus, A. G., & van IJzendoorn, M. H. (1995). Mothers reading to their three-year-olds: The role of mother-child attachment security in becoming literate. *Reading Research Quarterly, 30*, 998-1015.

Bus, A. G., & van IJzendoorn, M. H. (2004). Meta-analysis in reading research. In N. Duke & M. Mallette (Eds.), *Literacy research methods* (pp. 227–252). New York: Guilford Press.

Bus, A. G. (2001). Parent-child book reading through the lens of attachment theory. In L. Verhoeven & C. E. Snow (Eds.), *Literacy and Motivation; Reading Engagement in Individuals and Groups* (pp. 39-53). Mahwah: Lawrence Erlbaum Associates.

Bus, A. G., van IJzendoorn, M. H., & Pellegrini, A. D. (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.

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112,* 155-159.

Collins, M. F. (2005). IRA Outstanding Dissertation Award for 2005: ESL preschoolers' English vocabulary acquisition from storybook reading. *Reading Research Quarterly, 40,* 406-408.

* Crain-Thoreson, C., & Dale, P. S. (1999). Enhancing linguistic performance: Parents and teachers as book reading partners for children with language delays. *Topics in Early Childhood Special Education, 19,* 28-39.

* de Oliviera Fontes, M. J., & Cardoso-Martins, C. (2004). Efeitos da leitura de histórias no desenvolvimento da linguagem de crianças de nível sócio-econômico baixo [Effects of story reading on low SES children's language development]. *Psicologia: Reflexao e Crítica, 17,* 83-94.

Diamond, A., Barnett, W. S., Thomas, J., & Munro, S. (2007). Preschool program improves cognitive control. *Science, 318,* 1387-1388.
Dickinson, D. K., McCabe, A., Anastasopoulos, L., Peisner-Feinberg, E.-S., & Poe, M. D. (2003). The comprehensive language approach to early literacy: The interrelationships among vocabulary, phonological sensitivity, and print knowledge among preschool-aged children. *Journal of Educational Psychology, 95,* 465-481.
Dickinson, D. K., & Smith, M. W. (1994). Long-term effects of preschool teachers' book readings on low-income children's vocabulary and story comprehension. *Reading Research Quarterly, 29,* 104-122.
Dickinson, D. K., & Sprague, K. E. (2001). The nature and impact of early childhood care environments on the language and early literacy development of children from low-income families. In S. B. Neuman & D. K. Dickinson (Eds.), *Handbook of Early Literacy Research* (Vol. 1, pp. 263-280). New York: The Guilford Press.
* Droop, M., Peters, S., Aarnoutse, C., & Verhoeven, L. (2005). Effecten van stimulering van beginnende geletterdheid in groep 2 [Effects of beginning literacy education in kindergarten]. *Pedagogische Studiën, 82,* 160-180.
Duval, S., & Tweedie, R. (2000a). A nonparametric "trim and fill" method for accounting for publication bias in meta-analysis. *Journal of the American Statistical Association, 95*, 89-98.
Duval, S., & Tweedie, R. (2000b). Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics, 56,* 455-463.
Elley, W. B. (1989). Vocabulary acquisition from listening to stories. *Reading Research Quarterly, 24,* 174-187.
Evans, M. A., & Saint-Aubin, J. (2005). What children are looking at during shared storybook reading - Evidence from eye movement monitoring. *Psychological Science, 16,* 913-920.
Fuchs, D., Fuchs, L. S., Thompson, A., Al Otaiba, S., Yen, L., Yang, N. J., Braun, M., & O'Connor, R. E. (2001). Is reading important in reading-readiness programs? A randomized field trail with teachers as program implementers. *Journal of Educational Psychology, 93,* 251-267.
Hargrave, A. C., & Sénéchal, M. (2000). A book reading intervention with preschool children who have limited vocabularies: The benefits of regular reading and dialogic reading. *Early Childhood Research Quarterly, 15,* 75-90.
Hart, B., & Risley, T. (2003). The early catastrophe: The 30 million word gap. *American Educator, 27,* 4-9.
Heath, S. B. (1982). What no bedtime story means: Narrative skills at home and school. *Language in Society, 11,* 49-76.
Hedges, L. V. (2007). Correcting a significance test for clustering. *Journal of Educational and Behavioral Statistics, 32,* 151-179.
Ioannidis, J., Patsopoulos, N., & Evangelou, E. (2007). Uncertainty in heterogeneity estimates in meta-analyses. *BMJ, 335,* 914-916.
Juel, C. (2006). The impact of early school experiences on initial reading. In D. K. Dickinson & S. B. Neuman (Eds.), *Handbook of early literacy research* (Vol. 2, pp. 410-426). New York: The Guilford Press.
Justice, L. M., & Ezell, H. K. (2002). Use of storybook reading to increase print awareness in at-risk children. *American Journal of Speech-Language Pathology, 11,* 17-29.
Justice, L. M., & Lankford, C. (2002). Preschool children's visual attention to print during storybook reading: Pilot findings. *Communication Disorders Quarterly, 24,* 9-19.
Justice, L. M., Meier, J., & Walpole, S. (2005). Learning new words from storybooks: An efficacy study with at-risk kindergartners. *Language, Speech, and Hearing Services in Schools, 36,* 17-32.
Justice, L. M., Pullen, P. C., & Pence, K. (2008). Influence of verbal and nonverbal references to print on preschoolers' visual attention to print during storybook reading. *Developmental Psychology, 44,* 855-866.

Justice, L. M., Skibbe, L., Canning, A., & Lankford, C. (2005). Pre-schoolers, print and storybooks: An observational study using eye movement analysis. *Journal of Research in Reading, 28,* 229-243.

* Karweit, N. (1989). The effects of a story-reading program on the vocabulary and story comprehension skills of disadvantaged prekindergarten and kindergarten students. *Early Education and Development, 1,* 105-114.

Karweit, N., & Wasik, B. A. (1996). The effects of story reading programs on literacy and language development of disadvantaged preschoolers. *Journal of Education for Students Placed at Risk, 1,* 319-348.

* Kertoy, M. K. (1994). Adult interactive strategies and the spontaneous comments of preschoolers during joint storybook readings. *Journal of Research in Childhood Education, 9,* 58-67.

* Lamb, H. A. (1986). *The effects of a read-aloud program with language interaction.* Unpublished doctoral dissertation, The Florida State University, Tallahassee.

Lipsey, M. W. & Wilson, D. B. (2001). *Practical meta-analysis*. Thousand Oaks, CA: Sage.

Lonigan, C. J. (2006). Conceptualizing phonological processing skills in prereaders. In D. K. Dickinson & S. B. Neuman (Eds.), *Handbook of early literacy research* (Vol. 2, pp. 77-89). New York: The Guilford Press.

* Lonigan, C. J., Anthony, J. L., Bloomfield, B. G., Dyer, S. M., & Samwel, C. S. (1999). Effects of two shared-reading interventions on emergent literacy skills of at-risk preschoolers. *Journal of Early Intervention, 22,* 306-322.

* Lonigan, C. J., & Whitehurst, G. J. (1998). Relative efficacy of parent and teacher involvement in a shared-reading intervention for preschool children from low-income backgrounds. *Early Childhood Research Quarterly, 13,* 263-290.

Mason, J. M. (1992). Reading stories in preliterate children: A proposed connection to reading. In P. B. Gough, L. C. Ehri, & R. Treiman (Eds.), *Reading acquisition* (pp. 215-241). Hillsdale, NJ: Lawrence Erlbaum.

* Mautte, L. A. (1990). *The effects of adult-interactive behaviors within the context of repeated storybook readings upon the language development and selected prereading skills of prekindergaten at-risk students.* Unpublished doctoral dissertation, University of South Florida, Tampa.

McCormick, C., & Mason, J. M. (1986). *Use of little books at home: A minimal intervention strategy that fosters early reading* (Techn Rep. No. 388). Urbana: University of Illinois, Center for the Study of Reading.

McKeown, M. G. & Beck, I. L. (2006). Encouraging young children's language interactions with stories. In D. K. Dickinson & S. B. Neuman (Eds.), *Handbook of early literacy research* (Vol. 2, pp. 281-294). New York: The Guilford Press.

Metsala, J. L., & Walley, A. C. (1998). Spoken vocabulary growth and the segmental restructuring of lexical representations: Precursors to phonemic awareness and early reading ability. In L. C. Ehri & J. L. Metsala (Eds.), *Word recognition in beginning literacy* (pp. 89-120). Mahwah, NJ: Lawrence Erlbaum.

Mol, S. E., Bus, A. G., De Jong, M. T., & Smeets, D. J. H. (2008). Added value of dialogic parent-child book readings: A meta-analysis. *Early Education and Development, 19,* 7-26.

Morris, D., Bloodgood, J. W., Lomax, R. G., & Perney, J. (2003). Developmental steps in learning to read: A longitudinal study in kindergarten and first grade. *Reading Research Quarterly, 38,* 302-328.

* Morrow, L. M. (1988). Young children's responses to one-to-one story readings in school settings. *Reading Research Quarterly, 23,* 89-107.

* Morrow, L. M. (1989). The effect of small group story reading on children's questions and comments. In: S. McCormick & J. Zuttell (Eds.), *Cognition and social perspectives for literacy research and instruction* (pp. 77-86). Chicago: 38th Yearbook of the National Reading Conference.

Morrow, L. M., & Smith, J. K. (1990). The effects of group size on interactive storybook reading. *Reading Research Quarterly,* 213-231.

Moschovaki, E., & Meadows, S. (2004). A short term longitudinal study of classroom book reading in Greek kindergarten schools. *Educational Studies in Language and Literature, 4,* 151-168.

Murphy, M. M. (2007). Enhancing print knowledge, phonological awareness, and oral language skills with at-risk preschool children in Head Start classrooms. Unpublished doctoral dissertation, University of Nebraska, Lincoln.

NICHD Early Child Care Research Network (2005). Pathways to reading: The role of oral language in the transition to reading. *Developmental Psychology, 41,* 428 - 442.

Nielsen, D. C. (1993). The effects of four models of group interaction with storybooks on the literacy growth of low-achieving kindergarten children. In: D. J. Leu & C. K. Kinzer (Eds.), *Examining central issues in literacy research, theory, and practice* (pp. 279-287). Chicago: 42nd Yearbook of the National Reading Conference.

Ninio, A. (1980). Picture-book reading in mother-infant dyads belonging to two subgroups in Israel. *Child Development, 51,* 587-590.

Penno, J. F., Wilkinson, I. A. G., & Moore, D. W. (2002). Vocabulary acquisition from teacher explanation and repeated listening to stories: Do they overcome the Matthew effect? *Journal of Educational Psychology, 94,* 23-33.

Phillips, G., & McNaughton, S. (1990). The practice of storybook reading to preschool children in mainstream New Zealand families. *Reading Research Quarterly*, 25, 196-212.

Phillips, L. M., Norris, S. P., Mason, J. M., & Kerr, B. M. (1990). Effect of early literacy intervention on kindergarten achievement. In: J. Zuttell & S. McCormick (Eds.), *Literacy theory and research: Analyses from multiple paradigms* (pp. 199-207). Chicago: 39th Yearbook of the National Reading Conference..

Poe, M. D., Burchinal, M. R., & Roberts, J. E. (2004). Early language and the development of children's reading skills. *Journal of School Psychology, 42,* 315-332.

Raikes, H., Pan, B. A., Luze, G., Tamis-LeMonde, C. S., Brooks-Gunn, J., Banks Tarullo, J. C., et al. (2006). Mother-child bookreading in low-income families: Correlates and outcomes during the first three years of life. *Child Development, 6,* 193–235.

Reese, E., & Cox, A. (1999). Quality of adult book reading affects children's emergent literacy. *Developmental Psychology, 35,* 20-28.

Rohrbach, L. A., Grana, R., Sussman, S., & Valente, T. W. (2006). Type II translation; Transporting prevention interventions from research to real-word settings. *Evaluation & the Health Professions, 29,* 302-333.

Roy-Charland, A., Saint-Aubin, J., & Evans, M. A. (2007). Eye movements in shared book reading with children from kindergarten to Grade 4. *Reading and Writing, 20,* 909-931.

Sameroff, A. J., & Fiese, B. H. (2000). Models of development and developmental risk. In C. H. Zeanah (Ed.), *Handbook of infant mental health* (pp. 3-19). New York: Guilford.

Samuelsson, S., Olson, R., Wadsworth, S., Corley, R., DeFries, J. C., Willcutt, E. et al. (2007). Genetic and environmental influences on prereading skills and early reading and spelling development in the United States, Australia, and Scandinavia. *Reading and Writing, 20,* 51-75.

Scarborough, H. S., & Dobrich, W. (1994). On the efficacy of reading to preschoolers. *Developmental Review, 14,* 245-302.

Sénéchal, M. (1997). The differential effect of storybook reading on preschoolers' acquisition of expressive and receptive vocabulary. *Journal of Child Language, 24,* 123-138.

Sénéchal, M., & LeFevre, J. (2002). Parental involvement in the development of children's reading skill: A five-year longitudinal study. *Child Development, 73,* 445-460.

Sénéchal, M., Thomas, E., & Monker, J. A. (1995). Individual differences in 4-year-old children's acquisition of vocabulary during storybook reading. *Journal of Educational Psychology, 87,* 218-229.

Shernoff, E. S., & Kratochwill, T. R. (2007). Transporting an evidence-based classroom management program for preschoolers with disruptive behavior problems to a school: An analysis of implementation, outcomes, and contextual variables. *School Psychology Quarterly, 22,* 449-472.

Speece, D. L., Ritchey, K. D., Cooper, D. H., Roth, F. P., & Schatschneider, C. (2004). Growth in early reading skills from kindergarten to third grade. *Contemporary Educational Psychology, 29,* 312-332.

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

* Valdez-Menchaca, M. C., & Whitehurst, G. J. (1992). Accelerating language development through picture book reading: A systematic extension to Mexican day care. *Developmental Psychology, 28,* 1106-1114.

van Elsäcker, W., & Verhoeven, L. (1997). Kleuters leren meer van voorlezen in kleine groepen [Children in kindergarten learn more from interactive reading in small groups]. *Pedagogische Studiën, 74,* 117-129.

* van Kleeck, A., Vander Woude, J., & Hammett, L. (2006). Fostering literal and inferential language skills in Head Start preschoolers with language impairment using scripted book-sharing discussions. *American Journal of Speech-Language Pathology, 15,* 85-95.

Viechtbauer, W. (2007). Confidence intervals for the amount of heterogeneity in meta-analysis. *Statistics in Medicine, 26,* 37-52.

* Wasik, B. A., & Bond, M. A. (2001). Beyond the pages of a book: Interactive book reading and language development in preschool classrooms. *Journal of Educational Psychology, 93,* 243-250.

* Wasik, B. A., Bond, M. A., & Hindman, A. (2006). The effects of a language and literacy intervention on Head Start children and teachers. *Journal of Educational Psychology, 98,* 63-74.

* Whitehurst, G. J., Arnold, D. S., Epstein, J. N., Angell, A. L., Smith, M., & Fischel, J. E. (1994). A picture book reading intervention in day care and home for children from low-income families. *Developmental Psychology, 30,* 679-689.

Whitehurst, G. J., Epstein, J. N., Angel, A. L., Payne, A. C., Crone, D. A., & Fischel, J. E. (1994). Outcomes of an emergent literacy intervention in Head Start. *Journal of Educational Psychology, 86,* 542-555.

Whitehurst, G. J., Falco, F. L., Lonigan, C. J., & Fischel, J. E. (1988). Accelerating language development through picture book reading. *Developmental Psychology, 24,* 552-559.

Whitehurst, G. J., Zevenbergen, A. A., Crone, D. A., Schultz, M. D., Velting, O. N., & Fischel, J. E. (1999). Outcomes of an emergent literacy intervention from Head Start through second grade. *Journal of Educational Psychology, 91,* 261-272.

Yaden, D. B., Smolkin, L. B., & Conlon, A. (1989). Preschoolers' questions about pictures, print convention and story text during reading aloud at home. *Reading Research Quarterly, 24,* 188-214.