To read or not to read

Mol, S.E.

3

Added Value of Dialogic Parent-Child Book Readings: A Meta-Analysis

Abstract

Book reading has been demonstrated to promote vocabulary. The current study was conducted to examine the added value of an interactive shared book reading format that emphasizes active as opposed to non-interactive participation by the child. Studies that included a Dialogic Reading intervention group and a reading-as-usual control group, and that reported vocabulary as an outcome measure were located. After extracting relevant data from 16 eligible studies, a meta-analysis was conducted to attain an overall mean effect size reflecting the success of Dialogic Reading in increasing children's vocabulary compared to typical shared reading. When focusing on measures of expressive vocabulary in particular ($k = 9$; $n = 322$), Cohen's d was .59 ($SE = .08$; *95%* CI = .44, .75, $p < .001$), which is a moderate effect size. However, the effect size reduced substantially when children were older (five- to six-years-old) or when they were at risk for language and literacy impairments. Dialogic Reading can change the home literacy activities of families with two- to four-year-old children but not those of families with children at greatest risk for school failure.

Based on:

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. Doi: 10.1080/10409280701838603

Introduction

Exposure to books is a major source for the development of one of the main areas in learning to read: vocabulary (Bus, van IJzendoorn, & Pellegrini, 1995). The importance of vocabulary appears from findings indicating that the vocabulary of children entering first grade predicts not only their word reading ability at the end of first grade but also reading comprehension later in their school career (Juel, 2006). Books stimulate vocabulary because they contain a wider vocabulary than occurs in ordinary conversations (Sulzby, 1985). Even simple stories for two- and three-year-olds such as the Maisy stories (by Lucy Cousins) include complex words and phrases such as 'starving' or 'right on time' with a much higher incidence in books than in daily communication. In particular as a result of repeated encounters with the same story, young children appear to expand their vocabulary (e.g., Verhallen, Bus, & de Jong, 2006).

Assuming a social-constructionist nature of book reading, books cannot be a source for acquiring new vocabulary unless children get intensive help and support from adults. Consequently, children may almost never encounter solely an oral rendering of the text. Instead, in most cases the words of the author are accompanied by the social interaction between the adult reader and child. The book reading paradigm assumes that whether or not children become interested in books and learn new vocabulary through shared book reading depends on this social context. Following this reasoning, parent-child storybook reading is hypothesized to be effective mainly as a result of tutorial behavior from mothers or fathers (Ninio & Bruner, 1978). It is assumed that children benefit most from book reading designed to increase child responding at the expense of simply listening to parents read (e.g., DeTemple & Snow, 2003; Raikes et al., 2006). That is, book reading may be most effective when parents seek to involve the child actively in verbal exchanges when reading aloud (Huebner & Meltzoff, 2005). In line with this hypothesis, a seminal study by Whitehurst and colleagues (Whitehurst, Falco, Lonigan, et al., 1988) followed by a series of replications in groups of different age and socioeconomic status (SES) experimentally tested the relations between adult tutoring during book reading and children's linguistic development. Assuming that meta-analyses can be applied most fruitfully within research programs in which studies with similar designs (and measures) have accumulated over the years (Bus & van IJzendoorn, 2004), we decided to take a meta-analytic approach to test the tenability of Whitehurst's conclusion that variations in reading to young children can have appreciable effects on language development.

When parents approach storybook reading with the intent of teaching language to their children, they may use techniques such as asking questions, giving feedback, and adjusting questions to the developmental level of the child. Because observational studies have suggested that most parents do not apply interactive reading techniques spontaneously (e.g., Britto, Brooks-Gunn, &

Griffin, 2006; Laakso, Poikkeus, & Lyytinen, 1999; Silvén, Ahtola, & Niemi, 2003), it seems reasonable that parents should receive training in this. Whitehurst et al. (1988) were the first to report promising results on vocabulary as a result of an intervention based on a stimulation package called "Dialogic Reading", which was designed according to the following three principles: (a) the use of evocative techniques by the parent that encourage the child to talk about pictured materials; (b) informative feedback by incorporating expansions, corrective modeling, and other forms that highlight differences between what the child has said and what he might have said; and (c) an adaptive parent sensitive to the child's developing abilities. In sum, during typical shared reading the adult reads and the child listens, but in Dialogic Reading the child learns to become the storyteller; Dialogic Reading involves reading with rather than to children (Fielding-Barnsley & Purdie, 2003). A positive demonstration of this intervention could set the stage for more fine-grained analyses of particular processes mediating effects, but so far most follow-up intervention studies include this broad type of Dialogic Reading. Because replications of Whitehurst's influential study also have included children that have differed in age and risk status from the participants in Whitehurst's study, the set of studies enabled us to test which variables may moderate outcomes of the intervention.

Using an exhaustive set of studies investigating the effects of Dialogic Reading, we tested whether variations in parental reading affect children's language development. The experimental groups participated in a treatment program that instructed parents to alter the frequency and timing of various aspects of their child-directed speech during story time. Control families were assumed to read to their children but were not told to change any behaviors. Based on this set of studies, we aimed to test whether there is evidence for the hypothesis that Dialogic Reading causes effects beyond those of the typical format of book reading. At the time of the first meta-analysis of parent-child storybook reading (Bus et al., 1995), there were insufficient studies available to include dialogue between parents and children as a moderator variable and test the additional effects of the variations in book reading described above on learning outcomes. The frequency of parent-child storybook reading explained no less than 8% of the variance in language and literacy measures (Bus et al., 1995). By treating research group as a moderator variable, we tested whether the strong effect sizes found in the pioneering studies by Whitehurst and his colleagues could be replicated by subsequent research (Bus & Van IJzendoorn, 2004).

Focusing on vocabulary gains, the main indicator of results in all studies, we expected to find stronger effects for expressive than receptive vocabulary. That is, Dialogic Reading is particularly known for its role in stimulating active verbal involvement by the child and is therefore especially thought to support expressive language. We also expected Dialogic Reading to be more effective in younger age groups as their active participation and learning are highly dependent on

the parental ability to bridge the discrepancy between the child's world and the world of the book. Older and more experienced children may rely less on external support in understanding and enjoying the story and may be more inclined to initiate dialogue when they do not grasp the story content.

Lastly, we expected that studies including more low-educated families would reveal less pronounced effects than studies including higher educated families. Whitehurst et al.'s (1988) study focused on intact middle-class families living on suburban Long Island, New York. In line with Ninio (1980), who examined social class differences when mothers read to two-year-old children, it is to be expected that low-educated mothers may be less likely to engage in a number of potentially instructive behaviors during story time. In the same vein, Heath (1982) found that high-SES mothers pose more "why" questions than low-SES mothers do. Bus and van IJzendoorn (1995) observed that a high-SES mother is inclined to guide her three-year-old child's understanding of the story plot by asking questions and offering help to answer questions, whereas the low-SES mother, who reads infrequently, just explains parts of the story without any attempt to involve the child in thinking about the event. There is also evidence that when adults do not find reading books a source of pleasure for themselves, then activities such as book reading may not be embedded in family practice, and parents may not know how to engage children in reading sessions (Bus, Leseman, & Keultjes, 2000). In the context of these findings we wonder whether the technique of Dialogic Reading works for both children developing typically and those children at risk for language and literacy impairments (Hargrave & Sénéchal, 2000).

To sum up, this meta-analysis tested the following hypotheses:

1) Does Dialogic Reading intensify the effects of parent-child picture storybook sharing, and how strong is the additional effect of Dialogic Reading? We expected that Dialogic Reading would add to the effects of typical book reading.
2) Does Dialogic Reading affect expressive language skills more strongly than receptive language skills? Expressive language skills may be particularly affected because those skills are emphasized within the Dialogic Reading format.
3) Is the strength of the association between Dialogic Reading and outcome measures related to the age at which the intervention started? We hypothesized that older children with more linguistic skills would be less dependent on the qualities of book reading sessions than younger children with inferior linguistic and comprehension skills.
4) Is Dialogic Reading as strongly related to outcome measures in samples of young children at risk for language and literacy impairments as in samples of children not at risk? It was hypothesized, for instance, that at-risk parents may be less responsive to training in book-sharing skills because reading is not a source of pleasure for many parents in this group.

5) Were stronger effect sizes revealed by the group who started this line of research than other researchers replicating the study, as is apparent in other research domains (Bus & van IJzendoorn, 2004)? If this is the case, studies carried out by Whitehurst and colleagues should be more effective than studies carried by other research groups.

Method

Search Strategy

An extensive literature search was conducted across Psychological Abstracts Online (PsycINFO), the Education Resources Information Center (ERIC), Dissertations Abstracts Online and SilverPlatter's Information Retrieval System for the Web (WebSPIRS) to identify all eligible studies from 1988 to March 2007. The following keywords were used during this computer search: *(shared / interactive / dialogic) book reading, (early) intervention, reading intervention, home literacy environment, parent, parent-child interaction, achievement, language development, vocabulary, preschool, kindergarten,* and *young children*. Additional studies were identified by manual search in reference lists of previous meta-analyses and primary studies. To retrieve unpublished documents or clarify uncertainties, several authors were contacted.

Selection Criteria

To be included in the present meta-analysis, studies had to describe original data and meet the following criteria: (a) involve Dialogic Reading programs in which parents were trained to read interactively with their child; (b) include pre-conventional reading participants with no mental, physical, or sensory handicaps and are pre-conventional readers; (c) contain outcome variables that were objective measures of expressive and/or receptive vocabulary; (d) involve a (quasi)experimental design, that included a control group in which parents were asked to read as usual; (e) be reported in English (although no restriction was put on the country and therefore the language in which the study was conducted); and (f) be published or unpublished.

We excluded studies if the intervention involved teacher- and/or stranger-child rather than parent-child shared book reading (e.g., Valdez-Menchaca & Whitehurst, 1992; Wasik & Bond, 2001), or if the intervention was a combined school- and home-based program in which no separate data for parent-child interaction were reported (Hargrave & Sénéchal, 2000; Lonigan, Anthony, Bloomfield, Dyer, & Samwel, 1999; Whitehurst, Arnold, et al., 1994; Whitehurst, Epstein, et al., 1994; Whitehurst et al., 1999). Studies were also excluded when the focal children suffered from physical handicaps such as hearing impairment (Fung, Chow, & McBride-Chang, 2005). Furthermore, one study was eliminated from the initial sample because children in the control group received a language-

related intervention (Dale, Crain-Thoreson, Notari-Syverson, & Cole, 1996). Finally, some studies were excluded because no data were presented (e.g., Gormley & Ruhl, 2005; Whitehurst & Lonigan, 1998; Zevenbergen & Whitehurst, 2003) or a control group was missing (Huebner, 2000a, Huebner & Meltzoff, 2005). Unfortunately, one relevant study (Park, 2006) focusing on 20 Korean, two- to four-year-old children learning English in the United States, could not be located. The abstract suggests that a positive but not significant effect for the Dialogic Reading group was present in this study.

Coding of the Studies

We coded all relevant studies using a standardized data extraction tool into which the following study characteristics could be entered: (a) *bibliographic reference*: publication type, year of publication; (b) *sample descriptors*: number of participants per group, mean age and school type at the start of the study (preschool or kindergarten), country of origin (Asia, Australia, Europe, the United States), language(s) used in the intervention, risk status (at risk, not at risk); (c) *research design descriptors:* design (experimental, quasi-experimental), type of Dialogic Reading training (video instruction, group session, individual training), duration of the intervention (in weeks), check of treatment integrity (use of Dialogic Reading techniques and the frequency of book reading in the experimental and the control group; scale = 0-8); and (d) *outcome measures:* test(s) used to measure vocabulary (receptive and/or expressive), posttest data (mean and standard deviation, *t*-test, *F*-test, *p*-value, sample size per test).

As indicators of expressive vocabulary, we included the Expressive One-Word Picture Vocabulary Test (EOWPVT; Gardner, 1981), the expressive vocabulary subtest of the Illinois Test of Psycholinguistic Abilities (ITPA-VE; Kirk, McCarthy, & Kirk, 1968), and mean and/or total length of utterances by the child during videotaped reading sessions. As measures for receptive vocabulary we used the Peabody Picture Vocabulary Test (PPVT; Dunn & Dunn, 1997) and the Bracken Basic Concept Scale (BBCS; Bracken, 1984), which assesses conceptual development and tests concepts such as colors, materials, positions, sequences, and shapes.

If studies included more than one intervention or control group and all groups met the inclusion criteria, sample sizes were adapted. For example, Arnold, Lonigan, Whitehurst, and Epstein (1994) trained parents in the intervention group either by showing them a video or via group sessions. In this study, the sample size of the control group was split into two groups in order to include results from both Dialogic Reading-interventions. Chow, McBride-Chang, Cheung, and Chow (2007) included a typical reading control group and a no intervention control group in addition to an experimental group comprising parents trained to read dialogically. In the typical reading group, parents kept a log of their reading frequency and were asked to read as usual with their child when they received the

same picture storybooks as the parents in the Dialogic Reading group, whereas the children in the control group were only pre- and posttested. Because the intervention group could not be included twice, the sample size was divided by two. The means and standard deviations remained unchanged.

To assess inter-coder reliability, two coders independently coded all selected studies. Both coders agreed completely on including the same list of studies. The average percentage of agreement across study characteristics and moderators was 96% (κ = .94, *range* = .60 – 1.00). Discrepancies between coders were resolved by discussion until consensus was reached.

The coding process resulted in a final set of sixteen studies, of which eight studies reported measures of both children's receptive and expressive vocabulary. Of the remaining studies, seven tested only receptive vocabulary, whereas one focussed solely on expressive vocabulary. A total of 626 parent-child dyads ($N_{DR\text{-}intervention}$ = 313; $N_{Control}$ = 313) were studied. Participants' mean age ranged from 27.8 to 70.2 months. For further details about the characteristics of the studies included, see Appendix 3.1.

Meta-Analytic Procedures

The standardized difference between the mean of a Dialogic Reading-intervention group and a reading-as-usual control group at posttest was computed in order to quantify the additional value of Dialogic Reading on vocabulary. Because authors presented various statistics, we used Wilson's Effect Size Calculator (Wilson, 2001) and Comprehensive Meta-Analysis software (Version 2.2; Borenstein, Hedges, Higgins, & Rothstein, 2005) to calculate Cohen's *d* effect sizes for each vocabulary outcome. A positive effect size indicates a favorable outcome for the Dialogic Reading intervention group. A *d* of .20 is interpreted as a small, .50 as a moderate, and .80 as a large effect size (Cohen, 1992). Because studies with an increased sample size provide more reliable estimates of the population mean due to a smaller standard error, effect sizes were determined by weighting each outcome by the inverse of its variance (Cooper & Hedges, 1994; Lipsey & Wilson, 2001). However, some studies contained more than one outcome measure or effect size. To prevent unequal weighting of effect sizes due to the number of measurements in the study (e.g., some reported one receptive and two expressive vocabulary measures), effect sizes were aggregated within a domain before being averaged across outcomes within a study. Then, an overall combined Cohen's *d* was calculated per study before further analyses were conducted.

Bias due to the fact that studies with non-significant findings are less likely to be published was examined graphically by funnel plot analysis. For each study, effect sizes were plotted against precision as determined by the sample size or the inversed standard error, to detect a potential bias due to under-representation of studies with small sample sizes. We used the "trim and fill" method to calculate the effect of potential file drawer problems (Duval & Tweedie, 2000a, 2000b). We

also computed the fail-safe number (*N*fs), i.e., the number of studies with null results that have not been published but have to exist to overturn the association between Dialogic Reading and vocabulary gains to a level of non significance (Lipsey & Wilson, 2001). Furthermore, homogeneity across studies was assessed by means of the *Q*-statistic to determine whether variability among individual effect sizes was larger than should be expected based on subject level sampling error. Significant *Q*s imply heterogeneity, which indicates that the separate effect sizes do not all estimate the same population mean effect size (Lipsey & Wilson, 2001). Additionally, I-squared (I^2) is presented, which measures the degree of inconsistency between studies. I^2-values larger than 75% implying heterogeneity (Petticrew & Roberts, 2006). When heterogeneity was indicated by both measures, a random effects model was preferred, assuming that variability between studies was random instead of systematic (Rosenthal, 1995).

Effects of moderator variables such as population at school, risk status, publication year, and treatment integrity were tested by contrasting sub-samples or by applying a meta-regression model. It is important to note that each subset had to consist of at least four studies before contrasts were considered (Bakermans-Kranenburg, van IJzendoorn, & Juffer, 2003). Fixed effects models were applied in the case of homogeneous sets of outcomes, either by a significant $Q_{between}$ (heterogeneity between subsets) and a non-significant Q_{within} (homogeneity within the subsets), or an I^2 less than 75%. More conservative random effects model tests were applied when heterogeneous outcomes were present.

Results

Preliminary Analysis

No outlying values appeared to be present. That is, standardized *z* values were smaller than 3.26 and/or larger than -3.26 for all effect sizes ($p < .001$). Evidence for publication bias was absent, as the funnel plot of precision showed symmetry around the point estimate. Further, we tested whether more controlled experiments revealed stronger results than studies that hardly checked the content or frequency of book reading sessions. A meta-regression showed no significant differences ($d = .05$, $p > .05$), implying that the quality of the intervention did not affect the overall effect size. Furthermore, a moderator analysis revealed that studies conducted within the research group of Whitehurst ($k = 5$ studies, $n = 136$ children) did not significantly differ from experiments by other research groups ($Q = 3.64$; $p > .05$). The minor variation in duration of the interventions did not influence any effects either. We were not able to test for publication bias by comparing published and unpublished studies because only two unpublished studies were located.

The fixed effects model was applied in order to calculate the mean effect size for overall, expressive, and receptive vocabulary because either *Q* or I^2 indicated that the samples were homogeneous ($Q_{overall} = 34.10$, $p < .01$, $I^2 = 56.01$; $Q_{receptive} = 21.34$,

$p > .05$, $I^2 = 34.40$; $Q_{expressive} = 11.27$, $p > .05$, $I^2 = 29.00$). Standardized differences in means, 95% Confidence Intervals (CIs) per study, and outcome measure are presented in Appendix 3.1.

The Additional Effect of Dialogic Reading Intervention on Vocabulary Measures

For all included studies ($k = 16$, $N = 626$), Cohen's d equaled .42 ($SE = .06$; *95%* CI = .30, .53; $p < .001$). The fail-safe number indicated that 123 additional studies with null or non-significant results needed to be added to negatively influence this significant but small effect size. An overview of the distribution of the combined effect sizes per study is graphically displayed in Table 3.1.

Table 3.1

Stem-and-Leaf Display of the Effect Sizes at Posttesting on All Vocabulary Measures.

Stem	Overall Vocabulary [a]	Expressive Vocabulary [b]	Receptive Vocabulary [c]
1.0		3, 3	
.9	0, 1		
.8			
.7	3, 3	3	3
.6			7
.5	1, 3, 8	9, 9	8, 8
.4	4		1, 4
.3	9	2	3, 9
.2	6	3	7
.1	6	6, 8	3
.0			
-.0	5		
-.1	5, 9		5, 8, 9
-.2	6, 6		6
-.3			
-.4			
-.5			
-.6			
-.7			
-.8			8

Note. Combine the numbers under 'stem' with the numbers in the other columns to find all effect sizes for overall, expressive, and receptive vocabulary; for instance, negative effect sizes for 'overall vocabulary' were: -.05, -.15, -.19, -.26, and -.26. [a] Overall mean $d = .42$, $k = 16$ studies, $n = 626$ children ; [b] Overall mean $d = .59$, $k = 9$, $n = 322$; [c] Overall mean $d = .22$, $k = 15$, $n = 608$

When focusing on measures of expressive vocabulary in particular ($k = 9$; $n = 322$), Cohen's d was .59 ($SE = .08$; *95%* CI = .44, .75; $p < .001$), which is a moderate effect size. For the studies that reported receptive vocabulary measures ($k = 15$; $n = 608$), Cohen's d was small ($d = .22$, $SE = .09$; *95%* CI = .05, .39; $p < .01$). As the

95% CIs of these two outcome measures showed no overlap, the hypothesis that Dialogic Reading affects expressive vocabulary significantly more than receptive vocabulary was accepted. The number of missing studies that were needed to overturn these significant results were *N*fs = 92 for expressive and *N*fs = 9 for receptive vocabulary, respectively.

Explaining the Variability in Effect Sizes

A moderator analysis was conducted to test whether the intervention had a greater impact on younger children in preschool (k = 10, n = 351) than on older children in kindergarten (k = 6, n = 275). As is presented in Table 3.2, the overall vocabulary of preschool children benefited significantly more from the Dialogic Reading intervention (Q = 7.14, p < .01; d = .50, *SE* = .12; *95%* CI = .37, .64) than children in kindergarten (d = .14, *SE* = .07; *95%* CI = -.10, .37). This moderator could not be tested for expressive vocabulary outcomes, because the kindergarten subset consisted of less than four studies. With regard to receptive vocabulary, the moderator did not remain significantly different between subsets (Q = 1.30, p > .05).

Furthermore, we investigated whether children at risk benefited less from the Dialogic Reading intervention than children not at risk. Because SES was not clearly reported in all studies, risk status was based on the demographic variables of income or maternal education. Families designated as at risk (k = 7, n = 208) received governmental support (Lonigan & Whitehurst, 1998), had low incomes (Cronan, Cruz, Arriaga, & Sarkin, 1996), or had low educated mothers (Crain-Thoreson & Dale, 1999; Fielding-Barnsley & Purdie, 2002; 2003). Samples were designated as not at risk (k = 9, n = 418) if families had a modal income (Whitehurst et al., 1988), and mothers were educated at the tertiary (Arnold et al., 1994; Huebner, 2000b) or secondary level (Blom-Hoffman, O'Neill-Pirozzi, Volpe, Cutting, & Bissinger, 2006; Chow & McBride-Chang, 2003; Chow et al., 2007). The effect of Dialogic Reading did significantly differ between subsets (Q = 9.52, p < .01), with a moderate effect size for children not at risk (d = .53, *SE* = .07; *95%* CI = .40, .67) and only a small effect for children at risk (d = .13, *SE* = .11; *95%* CI = -.08, .35). When focusing solely on expressive vocabulary, the same significant differences in effect sizes were present (Q = 6.80, p < .01). When selecting receptive vocabulary outcomes, however, the moderator did not remain significant (Q = .002, p > .05).

Table 3.2
Meta-analytic Results of Dialogic Reading Intervention Studies and Moderators split for Overall vocabulary, Expressive and Receptive Vocabulary Outcome Measures.

	k	*n*	*d*	*95%* CI	*Q* [a]	*p*	I^2
Overall Vocabulary							56.01
Total set	16	626	.42	.16, .54	34.10	.00	
At risk					9.52	.00	
Yes	7	208	.13	-.08, .35	8.95	.18	32.94
No	9	418	.53	.40, .67	15.63	.05	48.83
Age group [b]					7.14	.01	
PreS	10	351	.50	.37, .64	18.60	.03	40.27
K	6	275	.14	-.10, .37	8.37	.14	51.60
Expressive Vocabulary							
Total set	9	322	.59	.44, .75	11.27	.19	29.01
At risk					6.80	.01	
Yes	4	96	.22	-.10, .54	0.13	.99	0.00
No	5	226	.71	.53, .89	4.34	.36	7.78
Age group					.66	.42	
Receptive Vocabulary							
Total set	15	608	.22	.05, .39	21.34	.09	34.40
At risk					1.14	.29	
Age group					1.30	.26	

Note. *k* = number of studies; *n* = total number of participants; *95%* CI = Confidence Interval; [a] *Q* statistic for moderator stands for effects of contrasts (*df* = number of subsets - 1); *Q* statistic for subset stands for homogeneity (*df* = k - 1); [b] *PreS* = Preschool; *K* = Kindergarten

Discussion

This meta-analysis tested the feasibility of an intervention designed to increase the quality of shared book reading among parents and their two- to six-year-old children. The meta-analysis demonstrates that enhancing the dialogue between parent and child during reading sessions strengthens the effects of book reading. However, the correlation between the intervention and a compound of linguistic skills was moderate (r = .20), explaining about 4% of the outcome measures in a set of 16 studies that included 626 children. When we restricted the analyses to studies that assessed expressive vocabulary (k = 9 studies, n = 322 children), the relation became stronger (r = .29), explaining about 8% of the variance. Apparently, not only does the exposure to a story promote language development, but it is also important that parents stimulate active involvement by eliciting verbal responses to the story with the help of open-ended questions. This outcome means it is likely that the quality of book reading is as important for language development as its frequency. The large effect size reported in Whitehurst et al.'s (1988) study was never replicated, but an authorial bias was not present as appears from the comparison of studies executed by Whitehurst and colleagues and studies carried out by other

research groups as well as from a non-significant publication bias. Besides that, a cumulative analysis did not show a decreasing effect of the intervention with an increase in publication year; a common result of other meta-analyses in which pioneering studies often appear as outliers that are unlikely to be replicated (Bus & van IJzendoorn, 2004; Petticrew & Roberts, 2006).

Not all children need Dialogic Reading to profit from parent-child book sharing, however. We also found evidence supporting the hypothesis that Dialogic Reading with older children does not have as great an impact as Dialogic Reading with younger age groups. The older children in this set of studies, the five- to six-year-olds, scarcely benefited from Dialogic Reading (d = .14), explaining less than 1% of the variance. Insofar as Dialogic Reading causes additional effects, these effects were only manifested in the younger, two- to four-year old age group (d = .50). In the latter group, about 4% to 5% of the differences in outcome measures were explained by stimulating active child participation through Dialogic Reading. It is not plausible that a lack of challenge causes the low effect sizes in kindergarten children, but it is conceivable that parents fail to adapt the technique to older children. That is to say, a set of specific techniques has been developed for reading with children aged four to five years alongside a set for reading with children aged two to three years (Zevenbergen & Whitehurst, 2003). The set for older children targets more advanced skills by asking specific types of questions, evaluating and expanding on the child's responses, and having the child repeat the expanded phrases. However, the information provided in the articles was mostly not sufficient to decide whether researchers indeed adapted Dialogic Reading in order to implement more challenging interventions for the older children.

The finding that older and more experienced children may depend less on external support in understanding and enjoying the story is in line with the assumption that children internalize previous experiences and generalize those to new situations. As a result, kindergarten children need less help and support to remain attentive and to discover exciting parts of the stories, even when stories are new. It is possible that the same amount of dialogue takes place during parent-child interaction, but that older children depend less on adults because they are more inclined to initiate dialogue when they do not grasp the story content or are deeply affected by events. As is also suggested by the outcomes of studies comparing the amount of talk in older and younger age groups (Bus & van IJzendoorn, 1988; Martin, 1998), another option is that sessions become less dialogic with age. Older children may even prefer hearing the story without interruptions because they have sufficient linguistic skills and knowledge to sustain interest in the story without a parent focusing their attention by asking questions and providing explanations. In fact, they may not experience interruptions for questions as stimulating, but rather as annoying and interfering. This interpretation is further supported by the number of studies involving kindergarten children that reported negative results for Dialogic Reading, namely three out of six experiments, in contrast to two out

of ten studies with negative results when participants were preschool children. Unfortunately, it is not possible to check this explanation because there are no studies demonstrating that older children indeed respond differently to Dialogic Reading compared with younger children.

Is Dialogic Reading a helpful technique to promote school readiness in children who are most in need of effective language promotion and pre-literacy experiences? In this context, any implementation of Dialogic Reading that includes parents with greater educational diversity should be considered (Huebner & Meltzoff, 2005). Parents with relatively low levels of education in particular could benefit from evocative, literacy-stimulating techniques, as dialogue during book sharing is not a self-evident phenomenon in low-educated families. A low-educated mother, in contrast to a better educated mother, often just explains details of the picture without any attempt to involve the child in thinking about the event (Arnold et al., 1994; Bus & van IJzendoorn, 1995; Cronan et al., 1996; Huebner & Meltzoff, 2005). It may therefore be hard for these parents to incorporate such behavior into book reading routines (e.g., Bus & van IJzendoorn, 1995; Heath, 1982; Ninio, 1980). For that reason, we tested whether effect sizes variedy as a function of risk status, a variable that, fortunately, was not confounded with age group in the present set of studies.

A unique result of this meta-analysis is that groups at risk for language and literacy impairments benefit less from Dialogic Reading than groups not at risk. Specifically, in the group at risk, we found a minimum effect size (explaining 1% of the variance), whereas in groups not at risk the effect size was substantial (7%). Two explanations for this disappointing result are germane. First, one could argue, in line with the so-called Matthew effect, that parents are required to have a strong educational background in order to use Dialogic Reading effectively (Fung et al., 2005). A critical test of this explanation is missing because there are no studies testing to what extent Dialogic Reading is actually realized in lower-and higher-educated families. Second, it may be that children at risk do not benefit from Dialogic Reading because making inferences (and similar requests) goes beyond their present abilities. Attempts to expand children's behavior may not be effective as long as the new behavior is not part of children's developing repertory of responses (Bus & de Jong, 2006). It is therefore imaginable that older children at risk benefit more from Dialogic Reading than younger children at risk but in view of the number of studies we were unable to test this hypothesis.

Cautions and Limitations

Because of a number of drawbacks, it seems important to replicate this meta-analysis a few years from now when the number of studies has grown. The present study included only a small set of studies and a moderate number of participants. Furthermore, to separate benefits of book reading and encouragement from the effects unique to the Dialogic Reading method, we need to be sure that the

frequency of book reading to children is similar in the control and experimental groups and that the sessions in the experimental group are more interactive as a result of training parents in Dialogic Reading. Many studies lack control of what actually happens in the control and experimental groups. Data describing the behavior of the control group and/or intervention group are often missing or scant. Providing more descriptive data is advisable, even though a quality variable coding the amount of information relating to the content and frequency of book reading did not correlate significantly with outcome measures. Furthermore, we cannot exclude bias due to the fact that the program was voluntary in many studies. It is possible that parents attracted to the intervention were those who were more likely to carry out program requirements. Lastly, it was not always possible to be certain of the origin of the intervention effects. In some studies (e.g., Crain-Thoreson & Dale, 1999; Lonigan & Whitehurst, 1998) the home interventions were part of a larger study that also targeted teacher-child reading. Although we solely selected children in the separate home condition, it cannot be ruled out that the interventions in the school environment affected their outcomes.

Practical Implications

The literature suggests that Dialogic Reading has potential for enhancing the language development of very young children, thus increasing the readiness with which they enter school (e.g., Cutspec, 2004). The behavioral change in parents that occured while sharing books with children, coupled with the expressive language gains demonstrated by the children who participated in the studies, provides an early childhood intervention that is worth implementing in families. However, such a conclusion with far-reaching consequences for intervention programs is only partly supported by the results of this quantitative meta-analysis based on a review of 16 experiments. Our findings indicate that Dialogic Reading does not form a scaffolding of parent-child opportunities for early literacy development for all parents. For reasons to be addressed and specified in further research, this meta-analysis indicates that a book-reading intervention standardized on middle-class White or suburban samples may not be appropriate for lower class families, analogous to outcomes of experiments in libraries and other domains (e.g., Neuman & Celano, 2006). In particular, the present set of studies suggests that Dialogic Reading can change the home literacy activities of families with two- to four-year-old children but not those of families with children at greatest risk of school failure.

Appendix

Appendix 3.1

Characteristics of Studies Included in the Meta-Analysis.

First Author	*Year*	*Continent*	*At Risk*	*School*[a]	M_{age} *int. gr.*	*Design*[b]	*Training DR*[c]	*Duration (weeks)*	*Control gr.*[d]	*N* ($n_{exp.}$ & $n_{cont.}$)[e]	*Outcome*[f]	*ES (d)* [g]	*95% CI*
Arnold (*st 1*: DR vs C)[1]	1994	U.S.	No	PreS	28.20	E	Group	4	No int.	37 (23 & 14)	*RV*	.41	-.26, 1.08
											EV	.59	.11, 1.07
											OL	.53	.14, .92
Arnold (*st 2*: DR_Video vs C)[1]	1994	U.S.	No	PreS	29.70	E	Video	4	No int.	27 (14 & 13)	*RV*	.67	-.11, 1.45
											EV	1.03	.46, 1.59
											OL	.90	.44, 1.36
Blom-Hoffman	Unpub.	U.S.	No	PreS	44.50	E	Video	6	Books	18 (8 & 10)	*EV*	.73	-.23, 1.69
Chow (*st 1*: DR vs TR)[2]	2003	Asia	No	K	63.72	E	Indiv.	8	Books	41 (13 & 28)	*RV*	-.15	-.81, .51
Chow (*st 2*: DR vs C)[2]	2003	Asia	No	K	63.72	E	Indiv.	8	No int.	41 (14 & 27)	*RV*	-.19	-.84, .46
Chow (*st 1*: DR vs TR)[2]	Unpub.	Asia	No	K	63.11	E	Group	12	Books	55 (18 & 37)	*RV*	.44	-.13, 1.01
Chow (*st 2*: DR vs C)[2]	Unpub.	Asia	No	K	63.11	E	Group	12	No int.	55 (19 & 36)	*RV*	.39	-.17, .96
Crain-Thoreson	1999	U.S.	Yes	PreS	49.90	E	Group	8	No int.	19 (10 & 9)	*RV*	.13	-.78, 1.03
											EV	.18	-.46, .82
											OL	.16	-.36, .69
Cronan (*st 1*: 18 Instr. visits vs C)[3]	1996	U.S.	Yes	PreS	27.80	E	Indiv.	28	No int.	32 (21 & 11)	*RV*	.73	-.02, 1.48
Cronan (*st 2*: 3 Instr. visits vs C)[3]	1996	U.S.	Yes	PreS	28.00	E	Indiv.	28	No int.	31 (21 & 10)	*RV*	-.26	-1.01, .50

Fielding-Barnsley	2002	Australia	Yes	K	63.00	QE	Indiv.	8	No int.	34 (17 & 17)	*RV*	-.88	-1.59, -.18
											EV	.32	-.36, .10
											OL	-.26	-.74, .23
Fielding-Barnsley	2003	Australia	Yes	G1	70.20	QE	Indiv.	8	No int.	49 (26 & 23)	*RV*	.58	.00, 1.15
Huebner	2000b	U.S.	No	PreS	28.61	E	Group session	6	Books	115 (79 & 36)	*RV*	.27	-.13, .67
											EV	.59	.36, .83
											OL	.51	.31, .71
Lonigan (*st 1*: low compliance)	1998	U.S.	Yes	PreS	41.90	E	Indiv.	6	No int.	20 (9 & 11)	*RV*	.33	-.56, 1.22
											EV	.23	-.40, .85
											OL	.26	-.25, .77
Lonigan (*st 2*: high compliance)	1998	U.S.	Yes	PreS	46.90	E	Indiv.	6	No int.	23 (7 & 16)	*RV*	-.18	-1.07, .71
											EV	.16	-.46, .79
											OL	-.05	-.56, .47
Whitehurst	1988	U.S.	No	PreS	29.40	E	Indiv.	4	No int.	29 (14 & 15)	*RV*	.58	-.17, 1.32
											EV	1.03	.58, 1.48
											OL	.91	.53, 1.29

Note. [a] School type at the start of the study: preschool (PreS), kindergarten (K), first grade (G1). [b] Design: experimental (E), quasi-experimental (QE). [c] Type of Dialogic Reading parent training (video instruction, group session, individual training). [d] Control group: got no intervention (No int.) or received books but no dialogic reading instruction (Books). [e] *n*'s were adapted when studies included more than one intervention or control group and all groups met the inclusion criteria. [f] Outcome measures were indicators of *receptive vocabulary* (RV): Peabody Picture Vocabulary Test (PPVT, e.g. Dunn & Dunn, 1997), Bracken Basic Concept Scale (BBCS, Bracken, 1984); and indicators of *expressive vocabulary* (EV) that were combined into an overall EV-*d*: Expressive One-Word Vocabulary Test (EOWPVT, Gardner, 1981), expressive vocabulary subtest of the Illinois Test of Psycholinguistic Abilities (ITPA-VE, Kirk, McCarthy & Kirk, 1968), mean and/or total length of utterances by the child during reading sessions (MLU); *Oral Language (OL)*: effect sizes of indicators of receptive and expressive vocabulary were combined into an overall OL-*d*. [g] Based on posttest data: means and standard deviations. When posttest means and standard deviations were not reported, Cohen's *d* was calculated using: t-test (Cronan et al., 1996), covariate adjusted F-test and *p*-values (Arnold et al., 1994), and F-test statistic based on pre-post gain scores (Fielding et al., 2002). [1] The size of the control group was split into two groups in order to include results of both Dialogic Reading-interventions. [2] The size of the Dialogic Reading group was split into two groups in order to include both control groups. Means and standard deviations remained unchanged. [3] We excluded the one- to two-year-olds because outcome measures were based on parent reports. Test results of the three-year-olds were based on objective measures of receptive vocabulary and included in the meta analysis. The n's of the intervention and control group were estimated by adding two to the number of degrees of freedom of the reported t-test statistic before dividing the adjusted degrees of freedom by two. Then, the size of the control group was split into two groups in order to include results of both dialogic reading interventions.

References

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

* Arnold, D. H., Lonigan, C. J., Whitehurst, G. J., & Epstein, J. N. (1994). Accelerating language development through picture book reading: Replication and extension to a videotape training format. *Journal of Educational Psychology, 86,* 235-243.

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.

* Blom-Hoffman, J., O'Neill-Pirozzi, T., Volpe, R., Cutting, J., & Bissinger, E. (2006). Instructing parents to use dialogic reading strategies with preschool children: Impact of a video-based training program on caregiver reading behaviors and children's related verbalizations. *Journal of Applied School Psychology,* 23, 117-131.

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

Bracken, B.A. (1984). *The Bracken Basic Concept Scale: Examiner's manual.* Columbus, OH: Bell & Howard.

Britto, P. R., Brooks-Gunn, J., & Griffin, T. M. (2006). Maternal reading and teaching patterns: Associations with school readiness in low-income African American families. *Reading Research Quarterly, 41,* 68-89.

Bus, A. G., & de Jong, M. T. (2006). Relationships in book sharing: A developmentally appropriate way to foster pre-academic growth. In S. Rosenkoetter & J. Knapp-Philo (Eds.), *To read the world: Language and literacy in the first 3 years* (pp. 123-144). Washington, DC: Zero-to-Three Press.

Bus, A. G., Leseman, P. P., & Keultjes, P. (2000). Joint book reading across cultures: A comparison of Surinamese-Dutch, Turkish-Dutch, and Dutch parent-child dyads. *Journal of Literacy Research, 32,* 53-76.

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., 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.

* Chow, B. W.-Y., & McBride-Chang, C. (2003). Promoting language and literacy development through parent-child reading in Hong Kong preschoolers. *Early Education and Development, 14,* 233-248.

* Chow, B. W.-Y., McBride-Chang, C., Cheung, H., & Chow, C. S.-L. (2007). Dialogic reading and morphology training in Chinese children: Effects on language and literacy. Manuscript submitted for publication.

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

Cooper, H., & Hedges, L. V. (Eds) (1994). *The handbook of research synthesis.* New York: Russell Sage Foundation.

* 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.

* Cronan, T. A., Cruz, S. G., Arriaga, R. I., & Sarkin, A. J. (1996). The effects of a community-based literacy program on young children's language and conceptual development. *American Journal of Community Psychology, 24,* 251-272.

Cutspec, P. A. (2004). Influences of dialogic reading on the language development of toddlers. *Bridges; Practice-Based Research Syntheses, 2,* 1-12. Retrieved February 21, 2007.

Dale, P. S., Crain-Thoreson, C., Notari-Syverson, A., & Cole, K. (1996). Parent-child book reading as an intervention technique for young children with language delays. *Topics in Early Childhood Special Education, 16,* 213-235.

DeTemple, J. & Snow, C. E. (2003). Learning words from books. In A. van Kleeck, S. A. Stahl, & E. B. Bauer (Eds.), *On reading books to children: Parents and teachers* (pp. 16-36). Mahwah, NJ: Erlbaum.

Dunn, L. & Dunn, L. (1997). *Peabody Picture Vocabulary Test* (PPVT-III, 3rd ed.). Circle Pines, MN: American Guidance Service.

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.

* Fielding-Barnsley, R., & Purdie, N. (2002). Developing pre-literacy skills via shared book reading: The assessment of a family intervention program for preschool children at risk of becoming reading disabled. *Australian Journal of Learning Disabilities, 7,* 13-19.

* Fielding-Barnsley, R., & Purdie, N. (2003). Early intervention in the home for children at risk of reading failure. *Support for Learning, 18,* 77-82.

Fung, P. C., Chow, B. W. Y., & McBride-Chang, C. (2005). The impact of a dialogic reading program on deaf and hard-of-hearing kindergarten and early primary school-aged students in Hong Kong. *Journal of Deaf Studies and Deaf Education, 10,* 82-95.

Gardner, M. F. (1981). *Expressive One-Word Picture Vocabulary Test*. Novata, CA: Academic Therapy Publications.

Gormley, S., & Ruhl, K. L. (2005). Dialogic shared storybook reading: An instructional technique for use with young students in inclusive settings. *Reading and Writing Quarterly: Overcoming Learning Difficulties, 21,* 307-313.

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.

Heath, S. B. (1982). What no bedtime story means: Narrative skills at home and school. *Language in Society, 11,* 49-76.

Huebner, C. E. (2000a). Community-based support for preschool readiness among children in poverty. *Journal of Education for Students Placed at Risk, 5*, 291-314.

* Huebner, C. E. (2000b). Promoting toddlers' language development through community-based intervention. *Journal of Applied Developmental Psychology, 21*, 513-535.

Huebner, C. E., & Meltzoff, A. N. (2005). Intervention to change parent-child reading style: A comparison of instructional methods. *Journal of Applied Developmental Psychology, 26,* 296-313.

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: Guilford Press.

Kirk, S. A., McCarthy, J. J., & Kirk, W. D. (1968). *Illinois Test of Psycholinguistic Abilities.* Urbana: University of Illinois Press.

Laakso, M. L., Poikkeus, A. M., &Lyytinen, H. (1999). Shared reading interaction in families with and without genetic risk for dyslexia: Implications for toddlers' language development. *Infant and Child Development, 8,* 179-195.

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

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 of preschool children from low-income backgrounds. *Early Childhood Research Quarterly, 13,* 263-290.

Martin, L. E. (1998). Early book reading: How mothers deviate from printed text for young children. *Reading Research and Instruction, 37,* 137-160.

Neuman, S. B., & Celano, D. (2006). The knowledge gap: Implications of leveling the playing field for low-income and middle-income children. *Reading Research Quarterly, 41,* 176-201.

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

Ninio, A., & Bruner, J. (1978). The achievement and antecedents of labelling. *Journal of Child Language, 5,* 1-15.

Park, J. Y. K. (2006). *Facilitating young children's second language development through picture book sharing.* Unpublished doctoral dissertation, University of Kansas, USA.

Petticrew, M., & Roberts, H. (2006). *Systematic reviews in the social sciences; A practical guide.* Oxford, England: Blackwell Publishing.

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

Rosenthal, R. (1995). Writing meta-analytic reviews. *Psychological Bulletin, 118,* 183-192.

Silvén, M., Ahtola, A., & Niemi, P. (2002). Early words, multiword utterances and maternal reading strategies as predictors of mastering word inflections in Finnish. *Journal of Child Language, 30,* 253-279.

Sulzby, E. (1985). Children's emergent reading of favorite storybooks: A developmental study. *Reading Research Quarterly, 20,* 458-481.

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.

Verhallen, M. J. A. J., Bus, A. G., & De Jong, M. T. (2006). The promise of multimedia stories for kindergarten children at risk. *Journal of Educational Psychology, 98,* 410-419.

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*(2), 243–250.

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., DeBaryshe, F. D., Valdez-Menchaca, M. C., et al. (1988). Accelerating language development through picture book reading. *Developmental Psychology, 24,* 552-559.

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

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.

Wilson, D.B. (2001). Effect size determination program. Retrieved February 28, 2007.

Zevenbergen, A. A., & Whitehurst, G. J. (2003). Dialogic reading: A shared picture book reading intervention for preschoolers. In A. van Kleeck, S. A. Stahl, & E. B. Bauer (Eds.), *On reading books to children: Parents and teachers* (pp. 177-200). Mahwah, NJ: Erlbaum.