



Published in final edited form as:

*J Learn Disabil.* 2011 ; 44(3): 258–275. doi:10.1177/0022219410378444.

## **A Synthesis of Read-Aloud Interventions on Early Reading Outcomes Among Preschool through Third Graders at Risk for Reading Difficulties**

**Elizabeth A. Swanson,**  
The University of Texas at Austin

**Jeanne Wanzek,**  
Florida State University

**Yaacov Petscher,**  
Florida Center for Reading Research

**Sharon Vaughn,**  
The University of Texas at Austin

**Jennifer Heckert,**  
Educational Consultant

**Christie Cavanaugh,**  
The University of Florida

**Guliz Kraft,** and  
The University of Texas at Austin

**Katie Tackett**  
American Institutes for Research

### **Abstract**

A synthesis and meta-analysis of the extant research on the effects of storybook read aloud interventions for children at-risk for reading difficulties ages 3–8 is provided. A total of 29 studies met criteria for the synthesis, with 18 studies providing sufficient data for inclusion in the meta-analysis. Read aloud instruction has been examined using dialogic reading, repeated reading of stories, story reading with limited questioning before, during, and/or after reading, computer assisted story reading, and story reading with extended vocabulary activities. Significant, positive effects on children’s language, phonological awareness, print concepts, comprehension, and vocabulary outcomes were found. Despite the positive effects for read aloud interventions, only a small amount of outcome variance was accounted for by intervention type.

---

Reading aloud to young children has been a commonplace practice in homes and schools for years. Parents, educators, policymakers, and politicians have promoted read alouds with the common belief that reading to children makes a difference in children’s literacy development. Indeed, adults reading to children under varying conditions may have social-emotional benefits for the children (Bus, 2001) as well as positive impacts on children’s language and literacy development (van Kleeck, 2004; Lonigan, 1994; NRC, 1998; Wasik & Hendrickson, 2004) and motivation to read (Gambrell & Marinak, 2009).

Many children learn to read effortlessly; however, an estimated 10% of children have difficulty acquiring reading skills and need additional support or specialized instruction (Catts & Hogan, 2003). Conditions that place children at risk for reading difficulties include poverty, cultural and linguistic differences, neurologically-based problems, inadequate

instruction, limited development-enhancing opportunities, or familial history of reading disabilities (McCoach, O'Connell, Reis, & Levitt, 2006; Neuman & Dickinson, 2001). Knowledge of these early risk factors as well as the negative impact of reading difficulties on future success in school and life (National Center for Educational Statistics 2004; 2006) have led to an increased focus on the early identification and prevention of reading difficulties in the early grades, including preschool (Catts & Hogan, 2003; Lyon, 1998).

The focus on early intervention for children at-risk for reading difficulties has led early childhood and elementary educators and researchers to examine early literacy instructional practices including the commonplace practice of storybook read alouds (National Early Literacy Panel, 2009). A corpus of research investigating the effects of read aloud practices in preschool and the early elementary grades has developed over the past several decades. Previous meta-analyses (e.g. Blok, 1999; Bus IJzendoorn & Pellegrin, 1995; Mol, Bus & DeJong, 2009; National Early Literacy Panel, 2009), and syntheses (Karweit & Wasik, 1996; Scarborough & Dobrich, 1994) have provided important information to inform the use of read alouds with young children, including a) support for reading aloud as an effective practice for improving child outcomes, b) initial work to identify study features associated with stronger effects, and c) some support of read-alouds as a practice for children at risk for reading difficulties, and d) identification of methodological concern for further study.

Scarborough and Dobrich (1994) synthesized 31 studies conducted over a period of thirty years that explored the influence of parent-child read aloud experiences on language and literacy development. Effect sizes for some studies were reported, and were moderately positive. Experimental groups outperformed control groups in interventions that involved parent training for shared book reading strategies. However, the strength of the association was inconclusive based on variability of results and methodology across the studies reviewed.

While the extensively cited Scarborough and Dobrich (1994) review was replete with evidence to support reading to children, the authors also drew attention to the potential overconfidence that educators, parents, and researchers place on the contributions of parent-child reading to the promotion of literacy development. They cited an ongoing debate about the bidirectional effect between child interest in reading and being read to. In other words, does reading to young children promote interest and skill development or do inherent interest and skill elicit more reading, thus impacting further skill development? Scarborough and Dobrich concluded that reading aloud to children is a complex practice, worthy of additional research and thus, subsequent studies and reviews on the practice have since flourished.

Karweit and Wasik (1996) complemented Scarborough and Dobrich's (1994) work by examining studies that focused on the effects of storybook reading on 4- and 5-year-olds in school settings. In their narrative review of studies that included children from disadvantaged settings conducted between 1988 and 1994, the authors identified the following practices associated with positive outcomes: small group size in favor of whole group or individual story reading, repeated readings to facilitate child involvement and comprehension, vocabulary instruction using synonyms or role playing, and teacher interaction and questioning strategies to support vocabulary learning and comprehension. Although the authors identified disadvantaged children as a focus, the data were not disaggregated to describe differential effects for children with varying abilities or risk status.

Blok (1999) conducted a meta-analysis of ten studies, published between 1968 and 1994, in which read alouds took place in educational settings. The studies included aggregated data for disadvantaged (not defined) and non-disadvantaged children ranging in age from 31 to

90 months. The overall effects of a variety of read aloud interventions were medium in size for oral language ( $ES = 0.63$ ) and reading ( $ES = 0.41$ ) outcomes. Blok (1999) identified study characteristics that acted as moderators for oral language and wide reading outcomes, citing that instruction with younger children and smaller instructional group sizes were associated with larger effects. In addition, the use of untrained adult readers was associated with greater effects in wide reading. It is unclear though, whether the provision of training for the readers was a constant variable across studies in which this was reported. An important, general conclusion of the meta-analysis was that although some studies revealed positive effects, the authors also reported some poor quality research and highlighted the need for additional empirical work on the use of reading aloud in educational settings as a means to improve literacy outcomes.

An increase in high-quality research on story read alouds has since occurred. A recent meta-analysis of early literacy practices for preschool and kindergarten children examined experimental and quasi-experimental literacy literature through 2003 (National Early Literacy Panel, 2009). The panel reported that shared reading interventions demonstrated moderate effects on children's print knowledge and oral language skills. Shared reading was examined across settings (e.g. schools, home, pediatricians provided books to parents, etc.) and across adult readers (parents, teachers). While the overall effect size for children who were not at risk seemed to be higher than for children who were at risk, the difference in effects was not statistically significant. Information on features of instruction and the effects on specific literacy outcomes for the at risk population were not analyzed.

Our synthesis and meta-analysis differs from previous reviews in several ways. First, we focus on only teacher-delivered interventions, excluding parent interventions (e.g. Bus et al., 1995; NELP, 2009). Second, we limit our review to studies that focus on students at risk for reading difficulty. Third, because previous meta-analytic results indicate no differences in effect size due to age differences (NELP, 2009), we have included preschool through third graders. Finally, previous syntheses have focused on limited interventions (e.g. Bus et al., 1995 focused on frequency of book reading) or limited outcomes (e.g. Mol et al., 2009 focused on vocabulary and print knowledge). Our synthesis considers all early reading and language outcomes. We address the following questions:

1. What outcomes result from read-aloud practices in educational settings for children considered at risk for reading difficulty?
2. What features of read aloud instruction are associated with improved outcomes for children at risk for reading difficulty?

## Method

A comprehensive search of the literature was performed through a three-step process (Cooper, 1998). First, a computer search of ERIC and PsycInfo was conducted to locate studies published between 1984 and 2008. Topic-related terms or root forms of those terms (story telling, story reading, book reading, storybook reading, read aloud, oral reading, retelling, and dialogic preschool\*, prekindergarten, pre-kindergarten, kindergarten\*, young children, early childhood, daycare, school, elementary school, teaching, intervention, literacy, oral language, communication, comprehension, language, LD, at-risk, disadvantaged, low income, disabilities, learning dis\*, learning difficulties, language delays, vocabulary, struggling readers, emerging readers, reading disabilities, reading difficulties) were used in various combinations to identify the greatest number of related articles. A total of 3,752 articles were identified in the initial search.

Second, to ensure the most recently published articles were included, a hand search of the following nine journals from 2004 to 2008 was conducted: *Journal of Educational Psychology*; *Reading Research Quarterly*; *Language or Speech or and Hearing Services in Schools*; *Early Childhood Research Quarterly*; *Topics in Early Childhood*; *American Journal of Speech Language Pathology*; *Exceptional Children*; *American Educational Research Journal*; and *Journal of Communication Disorders*. Third, we searched the reference list of each qualifying study to ensure that all studies meeting our criteria were identified.

Studies were selected if they met the following criteria:

1. Participants were in preschool – 3<sup>rd</sup> grade or ages 3 to 8. When a sample also included older participants, the study was retained if at least 50% of the sample met the age/grade criteria or outcome data for the targeted age/grade range (3–8 years; preschool-grade 3) could be disaggregated from the larger sample.
2. Children were at-risk for reading difficulties based on at least one of the following categories: low achievement in phonemic awareness, vocabulary, or letter identification, few preschool or home literacy experiences, low socioeconomic status, family history of reading disability, or attended a school with historically low reading achievement. When a study included average achieving children as well as those at risk, the study was retained if at least 50% of the sample was at-risk for reading difficulties or the reading outcomes could be disaggregated for at-risk children.
3. A treatment-comparison, multiple treatment, single-group, or single subject research design was utilized. Case studies and perception studies (e.g. Collins-Stanley & Gan, 1996) were excluded from the synthesis.
4. The study took place in a preschool, day care, or school setting. If an intervention took place in one of the selected settings and children also participated in a home-based or parental intervention, the study was included. Interventions conducted solely by parents were excluded.
5. The intervention consisted of a read aloud of a storybook in an alphabetic language by an adult or using audio-tape or computer assisted applications. Multi-component studies for which read aloud was not the primary focus (e.g. Hindson, Byrne, Fielding-Barnsley, Newman, Hine, & Shankweiler, 2005; Ukrainetz, Cooney & Dyer, 2000), interventions that included read alouds of expository text (e.g. Daly & Martens, 1994), and observation studies of existing practices (e.g. Purcell-Gates, McIntyre & Freppon, 1995) were excluded.
6. At least one dependent measure assessed reading or language-related outcomes.
7. The study was published in a refereed journal.

Twenty-seven articles (29 studies) met these criteria. Two articles included more than one study (Beck & McKeown, 2007; Coyne, McCoach & Kapp, 2007).

## Data Analysis

**Coding procedures**—We coded each study to identify and organize pertinent information using a code sheet employed in earlier syntheses (Kim, Vaughn, Wanzek, & Wei, 2004; Vaughn, Kim, Sloan, Hughes, Elbaum, & Sridhar 2003) and aligned with the What Works Clearinghouse Design and Implementation Assessment Device (Institute for Educational Sciences, 2003), which is used to evaluate the quality of studies. The code sheet

included information on participants, design, description of conditions, clarity of causal inference, possible sources of intervention contamination, and reported findings.

The participant section of the code sheet included four forced choice items (socioeconomic status, risk type, exceptionality, and gender) and two open-ended items (age as described in the text and risk type as described in the text). Design information was recorded using four forced choice items (research design, assignment of participants to study groups, use of fidelity checks, and report of pretest scores) and one open ended item (criteria for participant selection). Detailed information on both treatment and comparison groups was coded through open-ended items (age and grade of participants, site of intervention, role of person implementing intervention, length of each session, duration of intervention, total number of sessions, and frequency of sessions).

Clarity of causal inference was recorded using six items for experimental designs (e.g. presence/absence of differential attrition between intervention and comparison group, sample size of the intervention group at the beginning and end of the study) and nine items for quasi-experimental designs (e.g. procedures used to equate groups, evidence of differential attrition between groups). Additional items allowed coders to describe assessment measures. Finally, precision of outcomes for both effect size estimation and statistical reporting was coded using a series of eight forced-choice or yes/no questions (e.g. evidence of assumptions of independence, normality, and equal variance). Effect sizes were calculated using reported performance on outcome measures and sample sizes for each treatment and comparison condition.

Six researchers participated in a series of training sessions (approximately 8 hours total) on interpreting the code sheet. Following training, researchers coded one article to determine inter-rater agreement (agreements divided by agreements plus disagreements). Inter-rater reliability of .95 for study characteristic coding and .91 for effect size coding was achieved. Teams of two researchers coded each article, resolved any differences in coding, and reached final decisions by consensus. After coding was completed, information was compiled in table format. Table 1 includes information on study design, number and age of participants, type and duration of intervention, and person who implemented the intervention. In Table 2, interventions are described and effect sizes are reported when appropriate data were available.

**Article categorization**—A team of two researchers independently categorized articles by intervention type. A total of six read-aloud intervention types were identified. Descriptors for each intervention type were developed (see Table 3) and used by a third researcher to verify the categorization.

**Effect size calculation**—The primary statistic used to analyze the relationship between treatment and outcomes across the outcome measures was Cohen's  $d$  (Cohen, 1988). This statistic indicates, in standard deviation units, the extent to which the treatment groups outperformed the control groups. For studies in this synthesis that employed a treatment-comparison design, effect sizes can be interpreted as  $d = 0.2$  as small,  $d = 0.5$  as medium, and  $d = 0.8$  as a large effect (Cohen, 1988). The authors calculated Cohen's  $d$  by taking the difference of the treatment and control group means and dividing by the pooled standard deviation. Instances where  $F$ -test statistics were provided instead of means and standard deviations, the effect size was calculated using the  $F$  statistic and degrees of freedom. Nine studies did not report sufficient data to allow effect size calculation. Moreover, it is known that effect sizes from small studies tend to be biased, thus a sample weighted estimate of the standardized difference was computed.

**Meta-analysis plan**—We conducted a meta-analysis of a subset of 18 studies that employed a treatment-comparison design and included information about outcomes and sample sizes that allowed for meta-analytic analysis (studies are denoted in tables). Many of the studies included multiple treatment or control groups and administered more than one measure in order to assess performance on various outcomes. As such, in order to account for treatment and control differences specific to the different types of outcomes, individual meta-analyses were run for: language, phonological awareness, word recognition, print concepts, vocabulary, and reading comprehension. Estimation of the mean effect size, variability across studies, and moderator effects were assessed using a multilevel model approach using HLM 6.0 software (Raudenbush & Bryk, 2002). Multilevel models are aptly applied to meta-analytic data for a number of reasons: a) multi-level modeling can account for subjects nested within studies by calculating the variance at both the subject and study level; b) multilevel modeling provides flexibility in mixed model estimation; and c) the researcher can simultaneously estimate the average effect size (fixed effect), variance of the effect (random effect), and residual variance modeled by moderator variables.

Multilevel modeling has the benefit of estimating results based on several methods, including restricted maximum likelihood (REML; Raudenbush & Bryk, 1986). Traditionally, ordinary least squares or weighted least squares are used in a regression analysis, with both methods designed to reduce the squared differences between predicted and observed values of the dependent variables. In the context of a multilevel design however, both methods are less efficient and require more participants to achieve results comparable to those estimated by REML (Kreft & Leeuw, 1998). Although multilevel analyses are typically performed with the presumption that the variances are previously unknown, the meta-analysis must be performed under the known variance models. In these models, the variances must be calculated and entered into the multivariate data matrix along with the study ID, effect size, and moderator variables. The estimation of multilevel meta-analytic models is performed in a two-stage process. Initially, the unconditional model is run to estimate the grand mean effect size and the variance of the true effects. The level-1 model (within-studies) is summarized as:

$$d_j = \delta_j + e_j \quad (1)$$

where

$d_j$  = the effect size from a study

$\delta_j$  = the true effect size of  $d_j$  across all studies, or the corresponding population parameter

$e_j$  = the sampling error associated with  $d_j$  as the estimate of  $\delta_j$ .

Should statistically significant variability be observed, the analyst may perform a conditional analysis, which includes the level-2 moderator variables. The purpose of this second model is to generate an effect size for prediction, and estimate remaining residual variance after controlling for the entered level-2 moderators, summarized as:

$$\delta_j = \gamma_0 + \gamma_1 W_{1j} + \dots + \gamma_s W_{sj} + u_j \quad (2)$$

where

$\gamma_0$  = grand mean effect size

$\gamma_s$  = regression coefficients

$W_{1j} \dots W_{sj}$  = moderator variables predicting  $\delta_j$



$u_j$  = level-2 random error.

When combining the two equations, two sources of error are contained; one for the unconditional model ( $u_j$ ) and one for the conditional level-2 model ( $e_j$ ). Determination of whether the unconditional or conditional model is selected is based on the estimation of  $\tau$  (i.e. observed score variance), whereby a value of zero indicates that no variance in study effect sizes remains after partitioning sampling errors, thus the level-2 model would not be needed. Should  $\tau$  be significantly different from zero, the conditional model would be employed to test if the residual variance could be explained by the selected moderator variables.

Furthermore, at both the unconditional and conditional levels, empirical bayes (EB) estimates may be derived to compare with the maximum likelihood estimates (REML). This process allows the researcher to supply the grand mean effect size and create a normal distribution of effect sizes, shrinking the extreme estimates closer towards the mean. An advantage of this procedure is that the analyst may observe the degree of unreliability that exists within a particular study effect size. Whereas the REML approach considers all values as likely, the EB estimates generate a normal distribution around the grand mean for the individual cases, with those showing the most shrinkage as evidence that the units have little data, or may be considered to be outliers (Raudenbush & Bryk, 2002). Moreover, from both models, the variance components for the grand means can be used to estimate the proportion of reduction in variance accounted for when using the selected moderators in the analyses. In this equation  $\hat{\tau}_{qq}(UC)$  represents the variance component for the unconditional model, and  $\hat{\tau}_{qq}(C)$  is the variance component for the conditional model (Raudenbush & Bryk, 2002).

Prior to meta-analysis, estimation of the file-drawer problem was conducted to determine the extent to which unpublished, unsuccessful studies would need to be incorporated to make results non-significant. Since most published studies report statistically significant findings, it is supposed that greater non-published research exists for the topic. Results indicated that the file drawer problem did not pose a threat to the current study ( $1.72 < 1.96$ ).

## Results

A total of 27 articles, comprising 29 studies (Beck & McKeown, 2007 and Coyne et al., 2007 each contain two studies) are included in this synthesis, and reflect a variety of study designs. Therefore, we conducted several types of analyses to fully explain the results of these studies. First, we analyzed study features (e.g. study design and use of random assignment). Second, we conducted a meta-analysis of all treatment-comparison design studies to determine the effect of read aloud interventions on several early reading outcomes, along with a follow-up moderator analysis to examine differences between studies using criterion measures (“hard” criteria; i.e. below average achievement) or background characteristics (“soft” criteria; i.e. low SES, low school quality, history of family risk) and differences using standardized or researcher developed measures. Finally, we synthesized all single-group, single subject, multiple treatment, and three treatment-comparison studies not included in the meta-analysis by outcome and intervention type.

### Study Features

In the following section, we summarize information about study characteristics and design elements. Detailed information about each study is included in Tables 1 and 2.

**Sample characteristics**—Sample sizes ranged from one to 254 children, with a mean of 56.6, and a median of 58 participants. The majority of studies targeted preschool through

kindergarten children ( $n = 22$ ). However, three studies included first graders and two studies included second or third graders.

**Study design**—This corpus of studies included 21 treatment-comparison, five multiple treatment, two single group, and one single subject study. Several design elements strengthen the reliability and lend credibility to findings from treatment-comparison studies. These include use of random assignment, fidelity of treatment procedures, and the use of standardized dependent measures (U.S. Department of Education, 2003; Raudenbush, 2005; Shadish, 2002). The number of treatment-comparison studies that possess these design elements are reported in Table 4. Of note, two treatment-comparison studies (Lonigan, Anthony, Bloomfield, Dyer & Samwel, 1999; Whitehurst, Arnold, Epstein, Angell, Smith & Fischel, 1994) reported all three elements.

**Interventions**—Studies were placed into one of six categories (computer assisted, dialogic reading, limited questioning, repeated reading, vocabulary, and other) based on the authors' description of the intervention. The distribution of intervention type by design is reported in Table 5. In addition, Table 2 includes a description of each study's treatment and comparison conditions.

Seventeen studies reported number of intervention sessions with a range from 3 to 155, and a mean of 30 sessions. The number of sessions was not reported in 12 studies. Seven studies reported the length of sessions, with a range of 6 to 40 minutes and a mean of 24 minutes. Among treatment-comparison studies, the number of intervention sessions averaged 29, while the length of sessions averaged 20 minutes. Available information about frequency and length of sessions is reported in Table 1.

## Meta-Analysis

In this section, we report results of the meta-analysis by outcome (see Table 6) coupled with report of our investigation of moderating effects of intervention type on each outcome (see Table 7).

### Language Outcomes

The range of the 23 weighted effect sizes for language outcomes was  $-0.48$  to  $1.79$ , with an overall weighted mean from the unconditional model of  $0.29$  ( $t(23) = 3.16$ ,  $p = .005$ ). This suggests that children who received read aloud interventions significantly outperformed children in the comparison group on measures of language. The variance component from the multilevel model of  $0.003$  was not statistically significant ( $p = >.500$ .; Table 6), indicating that while intervention children outperformed comparison children, there was not significant variability among studies. Although the range suggested larger deflections from the mean in terms of magnitude, the amount of error in effect sizes precluded a rejection of the null hypothesis.

### Phonological Awareness Outcomes

A large grand mean difference was observed ( $d = 0.78$ ) for the phonological awareness effect sizes ( $t(21) = 5.23$ ,  $p < .001$ ). The range of effect sizes was  $-0.22$  to  $3.01$ , and the error around the estimates was relatively small. This resulted in a significant variance component ( $0.42$ ,  $p < .001$ ), suggesting that studies varied considerably in the magnitude of difference in performance on phonological awareness outcomes between treatment and control children. The second stage of analysis included the modeling of variance with selected moderators, namely, types of interventions used. Within the conditional model analysis, four types of treatment were identified, including: dialogic reading, computer



assisted, limited questioning, and repeated reading. Across the 14 effect sizes, dialogic reading was the most frequently observed intervention, and was set as the referent condition. The mean effect size from the conditional model was 0.84, indicating that children who were exposed to dialogic reading outperformed control children. Computer assisted interventions were weaker than dialogic reading by  $-0.39$  units, with a fitted mean effect size of 0.45. Limited questioning interventions were also weaker than dialogic reading by  $-1.07$ , producing a fitted mean effect size of  $-0.23$ . Neither computer assisted nor limited questioning mean effect sizes were statistically different from dialogic reading, likely due to the relatively few number of effect sizes represented by the two conditions. Conversely, children in the repeated reading condition demonstrated a significantly larger mean effect size (i.e., 2.59) compared to dialogic reading ( $t(19) = 2.27, p < .05$ ). A follow-up moderator analysis tested the extent to which differences in the definition of risk criteria used in the study related to residual unexplained variance in the effect sizes. Testing differences between studies using standardized or researcher developed measures was not warranted as all studies using phonological awareness used researcher developed assessments. Results from this test indicated that no significant differences were observed between studies using soft or hard criteria for risk ( $p = 0.64$ ).

By modeling treatment differences, the amount of variance was reduced from 0.42 in the unconditional model to 0.41. Using the estimates from both models, the variance reduction was 0.20, indicating that 2% of the variance in effect sizes was explained by the type of intervention that was administered.

### Print Concepts Outcomes

A total of ten effect sizes were calculated for print concepts outcomes, with a range of  $-0.77$  to 2.35. The grand mean from the unconditional model was 0.86 and was significantly different from a mean of 0 ( $t(10) = 3.17, p = .01$ ), with significant variance existing among these ten effect sizes ( $p < .001$ ). Across studies with print concept outcomes, three different interventions were coded—dialogic reading, repeated reading, and computer assisted. Similar to the phonological awareness analysis, dialogic reading was set as the referent condition, with a mean effect size of 0.70 observed. Given the large standard error associated with this mean (0.43), a rejection of the null hypothesis was not warranted. Similarly, the fitted means of both repeated reading (1.30) and computer assisted (0.18) were associated with large standard errors, and were not significantly different from dialogic reading. The secondary moderator analysis tested only the differences between soft and hard risk criteria, as all print concepts measures were researcher developed. Results indicated that the mean effect size for studies using soft criteria was significantly less (i.e.,  $-2.11$ ) than studies using hard criteria ( $t(7) = -4.84, p = .001$ ). The variance components between the unconditional and conditional models did decrease as a function of including intervention type; however, this only resulted in a 6% reduction. When accounting for differences due to risk criteria, the variance components decreased by a total of 82%. Residual variance after covarying interventions suggested that significant variability still existed that could be explained by other non-coded predictors.

### Comprehension Outcomes

Comprehension outcomes were represented by 22 effect sizes, with a range of  $-0.52$  to 1.47. The grand mean effect size for this outcome in the unconditional model was 0.70, and was statistically different from 0 ( $t(20) = 4.94, p < .001$ ). Significant variability existed in the mean, and was modeled by dialogic reading, computer assisted interventions, and a group of interventions labeled here as “other”. The mean effect size for dialogic reading was 0.60 ( $t(18) = 3.81, p < .01$ ). Computer assisted interventions showed stronger differentiation between treatment and control groups, with a fitted mean of 1.27. This value was more than

0.50 standard deviation units larger than the effect size for dialogic reading, and was significantly different. Conversely, the group of effect sizes “other” demonstrated a small negative deflection from dialogic reading ( $-0.18$ ); however, this was not considered to be a significant deviation. The secondary moderator analysis tested only the differences between researcher and standardized measures, as all study samples met soft criteria for risk status. The test of differences between researcher and standardized measures indicated that no statistically significant effects were observed ( $p=.07$ ). While the amount of variance in effect sizes decreased as a function of intervention types, only 4% was explained using these covariates, and a significant portion remained that could be explained by other variables.

### Vocabulary Outcomes

Vocabulary outcomes comprised the largest number of effect sizes ( $n = 51$ ), with a range across the weighted estimates of  $-1.19$  to  $4.73$ . From the unconditional model, a large mean coefficient of  $1.02$  was observed, which was significantly different from 0 ( $t(50) = 5.83, p < .001$ ), with a variance component that was relatively large ( $1.68$ ) when compared to values from the other outcomes. A total of five intervention categories measured vocabulary outcomes, including computer assisted, limited questioning, vocabulary, other interventions, and dialogic reading again set as the referent. The results from the conditional model indicated that the mean effect size for the referent was moderate in nature ( $0.57$ ), and was also significantly different from 0 ( $t(46) = 2.97, p < .01$ ). Vocabulary interventions showed to be weaker in magnitude than dialogic reading by  $-0.42$  units, but due to the large standard error, this was not a statistically meaningful difference. Similarly, while the mean effect size for limited questioning was larger in magnitude than the referent (i.e.,  $1.43$ ), a large standard error precluded a meaningful difference from dialogic reading. Interventions identified as “other” produced a substantially larger effect size compared to the referent, with a fitted value of  $2.18$ , and limited questioning interventions were significantly stronger as well with a fitted mean of  $2.00$ . These fitted means are representative of the studies based on inclusion criteria, however, the number of studies are relatively few. The supplementary test of both risk criteria and measures differences indicated no significant relationships existed between these moderators ( $p=.51$  and  $.16$ , respectively) and effect sizes. After accounting for the effects of intervention type in the model, the variability in the grand mean reduced from  $1.48$  to  $1.06$ , which translated into a 28% reduction of the variance in effect size differences.

### Findings from Additional Studies

The meta-analysis of findings from treatment-comparison studies provides confident conclusions about causal inferences. However, results from single-group, multiple treatment, and single-subject studies can be used to support or refute findings from the meta-analysis. Findings are summarized by comprehension, vocabulary, language, and word recognition outcome measures. Within each outcome measure, findings from single group, multiple treatment, and single-subject studies are reported. Because all studies with phonological awareness or print concept outcomes were included in the meta-analysis, no additional information for these outcome variables is reported below.

### Language Outcomes

In addition to the three studies included in the meta-analysis, one computer assisted (Verhallen, Bus & deJong, 2006), one limited questioning (Bygrave, 1994), one “other” (Bellon-Harn, Hoffman & Harn, 2004), and one repeated reading study (Pemberton & Watkins, 1987) included at least one language outcome. Verhallen and colleagues (2006) assessed understanding of syntax by having children repeat sentences from a story, with the idea that as children better understand the grammar of sentences, they make fewer mistakes

in repeating the sentences. Effect sizes were larger for kindergarteners who listened to a story read four times using an interactive computer-based, story reading program ( $ES = 1.04$ ) than children who listened to the story one time on the computer based program ( $ES = 0.27$ ) or listened to an adult read the story aloud four times ( $ES = -0.04$ ). Bygrave (1994) reported no differences among seven year olds between limited questioning and typical practice groups on a measure of receptive language concepts. In Bellon-Harn and colleague's (2004) single subject study, five and six year olds with phonological and language impairments increased the complexity of utterances after several types of scaffolds during read-aloud sessions, including cloze procedures (e.g. "There is some paint. They made..." and student completes the sentence), or a contrast word procedure (e.g. student says "weg" and teacher says, "weg or leg?") paired with asking a series of "wh" questions. Pemberton and Watkins (1987) reported no difference in language outcomes for pre-kindergarteners who were repeatedly read a story (6–7 times) written with base sentence-recast pairings (e.g. This is a frog. He is a big green frog) versus children who were read a book without the recast pairings ( $ES = 0.06$ ).

### Comprehension

In addition to the five treatment-comparison studies included in the meta-analysis, three dialogic reading (Combs, 1987; Davies, Shanks & Davies, 2004; Reutzel, Hollingsworth & Eldredge, 1994), two limited questioning (Bygrave, 1994; Morrow, 1984), and one study categorized as "other" (Boykin & Cunninham, 2001) included at least one comprehension outcome. The goal of dialogic reading treatments in three studies (Combs, 1987; Davies et al., 2004; Reutzel et al., 1994) was for children to develop a deep understanding of the story through the use of several strategies, including retelling. On tests of story recall, Combs (1987) reported that below average pre-kindergarteners doubled their recall of stories at posttest. Davies and colleagues (2004) reported statistically significant growth in the quantity of information children recalled, but no statistically significant growth from pre to post test in the quality of story retellings. In Reutzel et al.'s (1994) study, second graders were assigned to one of two different dialogic reading interventions (oral recitation lessons or shared book experience), whereby each group was read stories aloud and provided opportunities to practice retelling the stories. Additionally, children in the oral recitation group each had a copy of the text and spent time focused on developing reading fluency. There were little detectable differences between the two groups on measures of oral retelling ( $ES = 0.09$ ) and the comprehension subtest of the TTBS ( $ES = 0.13$ ). Small effects sizes were detected on a measure of answering explicit questions ( $ES = 0.25$ ). However, children in the shared book experiences group out-performed children in the oral recitation group when answering implicit questions ( $ES = 0.85$ ).

Bygrave (1994) and Morrow (1984) each implemented a limited questioning read-aloud intervention. In Bygrave's study, a story reading condition whereby seven-year-olds were read one short story per day and asked questions aimed at increasing comprehension and memory skills over a 23-week period was implemented. Morrow (1984) provided teacher-identified, low achieving kindergarteners with one of three types of questions before and after read aloud sessions conducted once per week over an 8-week period: (a) questions focused on story structure (e.g. setting, theme), (b) "traditional discussion" of the story, or (c) combined structural and traditional discussion. Bygrave (1994) and Morrow (1984) both reported comprehension outcomes for children at risk for reading difficulties assigned to the treatment conditions did not exceed those in the control conditions.

### Vocabulary

Fifteen treatment/comparison studies with vocabulary outcomes were included in the meta-analysis. One additional dialogic (Reutzel et al., 1994), one limited questioning (Bygrave,

1994), one “other” (Pemberton & Watkins, 1987) and two vocabulary studies (Beck & McKeown, 2007 study 2; Coyne et al., 2007 study 2) reported at least one vocabulary outcome measure. Second graders who received dialogic reading intervention (Reutzel et al., 1994) slightly outperformed children in a more fluency focused intervention on a standardized measure of vocabulary (ES = 0.19). However, Bygrave (1994) reported no significant differences between groups at post test. Pemberton and Watkins (1987) examined the effect of language recastings on vocabulary outcomes. They provided one group of pre-kindergarteners with multiple readings of a text written with language recastings, and the comparison group with repeated readings of a story with the same content, but no language recastings. Teachers read the book aloud to children six to seven times. Authors reported no statistically significant differences between the treatment and comparison groups on a standardized test of vocabulary (ES = 0.05). However both groups demonstrated statistically significant vocabulary gains from pre to post test.

Two authors designed vocabulary based interventions delivered during read aloud sessions (Beck & McKeown, 2007 study 2; Coyne et al., 2007 study 2). Both studies provided one treatment group with rich, extended vocabulary instruction that included a variety of interactive activities before, during, and following read alouds that focused on recognizing, discussing, and answering questions about target vocabulary words. However, the second treatment condition differed between these two studies. When Coyne and colleagues (2007, study 2) compared an extended vocabulary treatment with read alouds containing simple definitions of target words provided during reading, kindergarteners in the extended vocabulary treatment scored significantly higher on researcher developed measures of expressive (ES = 1.70), receptive (ES = 0.99), and context vocabulary (ES = 1.12). In addition, children with higher PPVT scores prior to treatment were more likely to learn word meanings through extended instruction than children with lower initial PPVT scores.

Beck and McKeown (2007, study 2) compared an extended vocabulary treatment (“Rich Instruction”) with a more intensive intervention where children received more frequent extended vocabulary instruction for a longer duration (“More Rich Instruction”). While effect sizes could not be calculated, authors reported that among both kindergarten and first graders, the pre to post test gain in words learned among children who received “More Rich Instruction” was higher than children in the “Rich Instruction” condition.

### Word Recognition

One dialogic reading study (Reutzel, et al., 1994) reported a word recognition outcome. Reutzel and colleagues (1994) reported that second graders who received the dialogic intervention plus fluency instruction outperformed children who received the dialogic intervention alone on the Iowa Test of Basic Skills (Hoover, Hieronymus, Frisbie & Dunbar, 1993) Word Analysis subtest (ES = -1.36) and oral reading error rate (ES = -1.01). However, the dialogic intervention only group read a greater number of words per minute at post test (ES = 0.63)

### Discussion

In this synthesis we sought to examine the effects of read-aloud practices on the literacy outcomes of children at-risk for reading difficulties. Read aloud instruction has been examined in several instructional formats including dialogic reading, repeated reading of stories, story reading with limited questioning before, during, and/or after reading, computer assisted story reading, and story reading with extended vocabulary activities. Dialogic reading has received the most examination in the literature. A few additional read aloud instructional techniques have been examined in single studies, reviewed in this synthesis as “other” read aloud interventions.

Previous syntheses of storybook reading have indicated a lack of high quality research, qualifying the findings and decreasing the ability to make robust statements regarding the effects of read alouds on literacy outcomes for children (Blok, 1999; Scarborough & Dobrich, 1994). Our synthesis of the current literature indicates the amount of high quality of research has increased. We found 20 treatment-comparison studies of read aloud interventions specifically for children with risk factors for reading difficulties. Fourteen studies included standardized assessments of children's literacy outcomes, allowing for reliable and valid examination of broad literacy outcomes following interventions. The larger and higher quality corpus of studies allowed us to examine more differentiated categories of independent and dependent variables than has been previously reported.

The meta-analysis revealed significant, positive effects for read aloud interventions on children's language, phonological awareness, print concepts, comprehension, and vocabulary outcomes. These results suggest that read aloud interventions provide children at-risk for reading difficulties with higher literacy outcomes than children who do not participate in these interventions. Although the effect on language outcomes was small and not moderated by intervention type, all other mean effects were large with intervention type explaining some of the variance in outcomes. These results are line with the recent National Early Literacy Panel (NELP, 2008) review of the effects of shared-reading interventions. The Panel reported significant effects for oral language (including listening comprehension and vocabulary) and print knowledge, but did not disaggregate these individual outcomes for children at-risk for reading difficulties. Encouragingly, our findings suggest children at-risk for reading difficulties do benefit from read aloud interventions in these areas. In contrast to the NELP (2008) results, we also found significant effects for phonological awareness outcomes suggesting a possible additional benefit for children who are at-risk for reading difficulties when read aloud interventions are implemented in educational settings.

Dialogic reading has the most causal evidence to support its effects on children's literacy outcomes with 8 experimental studies. The meta-analysis indicated moderate to large mean effect sizes for dialogic reading interventions on child outcomes of phonological awareness, print concepts, reading comprehension, and vocabulary. Thus, the evidence suggests that extended dialogue around the read aloud with transfer of storytelling to the child can improve literacy outcomes for children at-risk for reading difficulties.

The meta-analysis also revealed that computer-assisted interventions demonstrated small to large mean effects on literacy outcomes while limited questioning interventions revealed a small, negative mean effect size on phonological awareness yet a large effect size related to vocabulary outcomes. In both cases, there were few effect sizes contributing to these mean effects, and the mean effects for limited questioning were not significantly different than the mean effect for dialogic reading. Computer-assisted interventions demonstrated significantly higher effects on reading comprehension and vocabulary outcomes for children than the dialogic reading interventions. The higher effects on vocabulary may be the result of the outcome measures. The computer-assisted studies used researcher-developed measures assessing student knowledge of the meaning of the words in the book read aloud during the intervention. In contrast, the dialogic reading studies largely measured broad vocabulary outcomes (beyond words in the intervention storybooks) on standardized measures of vocabulary.

The practice of repeatedly reading stories to children at-risk for reading difficulties has received limited attention, with only two studies examining the intervention effects on only two literacy outcomes, language and print concepts. A large mean effect was noted for print concepts; however these effects are reported from only one study and were not significantly different from the print concepts outcomes for dialogic reading. A second study found no



differences in language outcomes between two versions of a repeated reading treatment, but reported significant differences between pretest and posttest scores for children involved in the treatments. The evidence suggests that repeated reading interventions may have potential for positively effecting student outcomes, but additional research is needed.

Intervention type accounted for 28% of the variance in vocabulary effect sizes. Although seemingly counterintuitive, the meta-analysis indicates that read alouds with extended vocabulary had the lowest mean effect size for vocabulary outcomes. However, this small effect was not significantly different than the moderate effect for dialogic reading. A closer examination of the vocabulary studies indicates children are making gains in the words that are taught in the intervention with the storybook, but demonstrating less gain in uninstructed vocabulary words. Two studies in this synthesis also noted differential effects on vocabulary outcomes, reporting children with the lowest vocabulary at pretest made the most gains in vocabulary (Coyne et al., 2004; Justice et al., 2005). Importantly, our synthesis reveals that participation in the different types of read aloud interventions seems to stimulate vocabulary development for children at-risk for reading difficulties, even on standardized measures of vocabulary.

Despite the positive effects for read aloud interventions, only a small amount of variance was accounted for by intervention type in the outcomes. The meta-analysis indicates that other unknown factors beyond the provided intervention explain significant amounts of variance in child outcomes on each of the measures. Similarly, several studies examining multiple read aloud treatments reported no significant differences in outcomes for the different treatment conditions, though significant pre to posttest gains were realized in each treatment. These findings may suggest that while read aloud interventions are valuable for children's literacy outcomes, some of the specific features related to improvement have not been fully realized in the literature. Currently, the strongest evidence comes from dialogic reading interventions suggesting that incorporating extended child-adult dialogue and questioning around storybooks is a valuable practice in educational settings. Computer-assisted read alouds demonstrate promise for improving children's literacy outcomes as well. The studies examining computer-assisted read alouds generally included interactive features allowing the child to engage in key aspects of the story line or character manipulation.

The literature provides little information regarding the long-term effects of read aloud interventions. Only 3 studies examined delayed or long-term outcomes of the interventions (Coyne et al., 2007 Study 1 and Study 2; Hindson et al., 2005). In both of their studies, Coyne et al. (2007) reported children in the read aloud intervention continued to outperform children in the control group 6 weeks after intervention on words taught in the intervention. Hindson et al. (2005) reported 1 year follow-up results indicating children in the read aloud intervention continued to perform at grade level expectations, but children not at-risk for difficulties made more gains in the year thus widening the gap. Future research on read aloud interventions should provide additional information on the long-term outcomes of participation in the interventions to better inform practice.

## Limitations

We have the most confidence in reporting findings related to dialogic reading given the relatively large number of studies examining this type of intervention. Only 2–3 studies each have examined computer-assisted read alouds, repeated reading of read alouds, or limited questioning. Thus, often there was a lot of error around the estimates. The low number of studies with common intervention features precluded us from examining more specific features of intervention implementation such as group size and amount of time in intervention that may moderate student outcomes. These factors may explain some of the



additional variance in child literacy outcomes and warrant continued research to inform early intervention efforts for children at-risk for reading difficulties. Finally, one should interpret the findings of the supplementary tests of differing risk criteria and standardized or researcher based assessments with caution. The very small number of effect sizes contributing to some of these supplementary tests was underpowered to accurately detect statistically significant differences, meaning that we cannot completely rule out the possibility of Type I error.

## References

- \*. Aram D, Biron S. Joint storybook reading and joint writing interventions among low SES preschoolers: Differential contributions to early literacy. *Early Childhood Research Quarterly*. 2004; 19:288–610.
- \*. Beck, IL.; McKeown, MG. Different ways for different goals, but keep your eye on the higher verbal goals. In: Wagner, RK.; Muse, A.; Tannenbaum, K., editors. *Vocabulary Acquisition: Implications for Reading Comprehension*. New York: Guilford Publications; 2007. p. 182-204.
- \*. Bellon-Harn ML, Hoffman PR, Harn WE. Use of cloze and contrast word procedures in repeated storybook reading: targeting multiple domains. *Journal of Communication Disorders*. 2004; 37:53–75. [PubMed: 15013379]
- Blok H. Reading to young children in educational settings: A meta-analysis of recent research. *Language Learning*. 1999; 49(2):343–371.
- \*. Boykin AW, Cunningham R. The effects of movement expressiveness in story content and learning context on the analogical reasoning performance of African American children. *The Journal of Negro Education*. 2001; 70(1/2):72–83.
- Bus, AG. Joint caregiver-child storybook reading: A route to literacy development. In: Neuman, SB.; Dickinson, DK., editors. *Handbook of early literacy research*. Vol. 1. New York: Guilford Press; 2001. p. 179-191.
- Bus AG, van IJzendoorn JH, Pellegrini AD. Joint book reading makes for success in learning to read: A meta-analysis on intergenerational transmission of literacy. *Review of Educational Research*. 1995; 65:1–21.
- \*. Bygrave PL. Development of listening skills in students in special education settings. *International Journal of Disability, Development, and Education*. 1994; 41(1):51–60.
- Catts H, Hogan T. Language basis of reading disabilities and implications for early identification and remediation. *Reading Psychology*. 2003; 24:223–246.
- Cohen, J. *Statistical power analysis for the behavioral sciences*. 2. Hillsdale, NJ: Lawrence Earlbaum Associates; 1988.
- Collins-Stanley T, Gan S. Choice of romantic or violent or and scary fairy-tale books by preschool girls and boys. *Child Study Journal*. 1996; 26:279–303.
- \*. Combs M. Modeling the reading process with enlarged texts. *The Reading Teacher*. 1987; 40(4): 422–426.
- Cooper, H. *Synthesizing research: A guide for literature reviews*. Thousand Oaks, CA: Sage Publications; 1998.
- \*. Coyne MD, McCoach DB, Kapp S. Vocabulary intervention for kindergarten students: Comparing extended instruction with embedded instruction and incidental exposure. *Learning Disabilities Quarterly*. 2007; 30:74–88.
- \*. Coyne MD, Simmons DC, Dame'enui EJ, Stoolmiller M. Teaching vocabulary during shared storybook readings: An examination of differential effects. *Exceptionality*. 2004; 12(3):145–162.
- Daly E, Martens B. A comparison of three interventions for increasing oral reading performance: Application of the instructional hierarchy. *Journal of Applied Behavior Analysis*. 1994; 27:459–469. [PubMed: 16795833]
- \*. Davies P, Shanks B, Davies K. Improving narrative skills in young children with delayed language development. *Educational Review*. 2004; 56(3):271–286.

- \*. Dickinson DK, Smith MW. Long-term effects of preschool teachers' book readings on low-income children's vocabulary and story comprehension. *Reading Research Quarterly*. 1994; 29(2):105–122.
- Ezell, HK.; Justice, LM. *Shared storybook reading: Building young children's language and emergent literacy skills*. Baltimore, MD: Brookes; 2005.
- Gambrell, L.; Marinak, B. Reading motivation: What the research says. 2009. Retrieved September 3, 2009 from <http://www.readingrockets.org/article/29624>
- \*. Hargrave AC, Sénéchal M. A book reading intervention with preschool children who have limited vocabularies: The benefits of regular reading and dialogic reading. *Early Childhood Research Quarterly*. 2000; 15(1):75–90.
- Hindson B, Byrne B, Fielding-Barnsley R, Newman C, Hine DW, Shankweiler D. Assessment and Early Instruction of Preschool Children at Risk for Reading Disability. *Journal of Educational Psychology*. 2005; 97(4):687–704.
- Hoover, H.; Hieronymus, A.; Frisbie, D.; Dunbar, S. *Iowa tests of basic skills*. Chicago: Riverside; 1993.
- Institute of Education Sciences. What Works Clearinghouse Study Review Standards. 2003. Retrieved September 24, 2005 from What Works Clearinghouse Web site: [http://www.whatworks.ed.gov/reviewprocess/study\\_standards\\_final.pdf](http://www.whatworks.ed.gov/reviewprocess/study_standards_final.pdf)
- \*. Justice LM, Ezell HK. Use of storybook reading to increase print awareness in at-risk children. *American Journal of Speech-Language Pathology*. 2002; 11:17–29.
- \*Justice LM, Meier J, Walpole S. Learning new words from storybooks: Findings from an intervention with at-risk kindergarteners. *Language, Speech, and Hearing Services in Schools*. 2005; 36:17–32.
- Karweit N, Wasik BA. The effects of story reading programs on literacy and language development of disadvantaged preschoolers. *Journal of Education for Students Placed at Risk*. 1996; 1(4):319–348.
- Kim A, Vaughn S, Wanzek J, Wei S. Graphic organizers and their effects on the reading comprehension of students with learning disabilities. *Journal of Learning Disabilities*. 2004; 37:105–118. [PubMed: 15493233]
- \*. Korat O, Shamir A. Electronic books versus adult readers: Effects on children emergent literacy as a function of social class. *Journal of Computer Assisted Learning*. 2007; 23:248–259.
- Kreft, I.; De Leeuw, J. *Introducing multilevel modeling*. Thousand Oaks, CA: Sage; 1998.
- Lonigan CJ. Reading to preschoolers exposed: Is the emperor really naked? *Developmental Review*. 1994; 14:303–323.
- \*. Lonigan CJ, Anthony JL, Bloomfield B, Dyer SM, Samwel C. Effects of two preschool shared reading interventions on the emergent literacy skills of children from low-income families. *Journal of Early Intervention*. 1999; 22:306–322.
- \*. Lonigan CJ, Whitehurst GJ. Relative efficacy of parent and teacher involvement in a shared-reading intervention for preschool children from low-income backgrounds. *Early Childhood Research Quarterly*. 1998; 17:265–292.
- Lyon GR. Why reading is not a natural process. *Educational Leadership*. 1998; 55(6):14–18.
- McCoach DB, O'Connell AA, Reis SM, Levitt H. Growing readers: A hierarchical linear model of children's reading growth over the first two years of school. *Journal of Educational Psychology*. 2006; 98:14–28.
- Mol SE, Bus AG, deJong MT. Interactive book reading in early education: A tool to stimulate print knowledge as well as oral language. *Review of Educational Research*. 79:979–1007.
- \*. Morrow LM. Reading stories to young children: Effects of story structure and traditional questioning strategies on comprehension. *Journal of Reading Behavior*. 1984; 16(4):273–288.
- \*. Morrow LM. Young children's responses to one-to-one story readings in school settings. *Reading Research Quarterly*. 1988; 23(1):89–107.
- National Center of Educational Statistics. *The condition of education 2004*. Washington, DC: U. S Department of Education; 2004.
- National Center for Educational Statistics. *The condition of education 2006*. Washington, DC: U.S Department of Education; 2006.

- National Early Literacy Panel. Developing early literacy: Report of the National Early Literacy Panel. Washington, DC: National Institute for Literacy; 2009.
- National Research Council. Preventing reading difficulties in young children. Washington, DC: National Academy Press; 1998.
- Neuman, SB.; Dickinson, DK., editors. Handbook of early literacy research. New York: Guilford Publications; 2001.
- \*. Pemberton EF, Watkins RV. Language facilitation through stories: Recasting and modeling. *First Language*. 1987; 7:79–89.
  - \*. Peskin J, Astington JW. The effects of adding metacognitive language to story texts. *Cognitive Development*. 2004; 19:253–273.
- Purcell-Gates VP, McIntyre E, Freppon PA. Learning written storybook language in school: A comparison of low-SES children in skills-based and whole language classrooms. *American Education Research Journal*. 1995; 32(3):659–685.
- Raudenbush SW. Learning from attempts to improve schooling: The contribution of methodological diversity. *Educational Researcher*. 2005; 34 (5):25–31.
- Raudenbush SW, Bryk AS. A hierarchical model for studying school effects. *Sociology of Education*. 1986; 59:1–17.
- Raudenbush, SW.; Bryk, AS. *Hierarchical Linear Models: Applications and Data Analysis Methods*. Newbury Park, California: Sage Publications; 2002.
- \*. Reutzel DR, Hollingsworth PM, Eldredge JL. Oral reading instruction: The impact on student reading development. *Reading Research Quarterly*. 1994; 29:40–62.
- Scarborough HS, Dobrich W. On the efficacy of reading to preschoolers. *Developmental Review*. 1994; 14:245–302.
- Shadish WR. Revisiting field experimentation: Field notes for the future. *Psychological Methods*. 2002; 7(1):3–18. [PubMed: 11928889]
- \*. Talley S, Lancy DF, Lee TR. Children, storybooks, and computers. *Reading Horizons*. 1997; 38(2): 116–128.
- Ukrainetz TA, Cooney MH, Dyer SK. An investigation into teaching phonemic awareness through shared reading and writing. *Early Childhood Research Quarterly*. 2000; 15:331–355.
- \*. Ulanoff SH, Pucci SL. Learning words from books: The effects of read aloud on second language vocabulary acquisition. *Bilingual Research Journal*. 1999; 23(4):319–332.
- U.S. Department of Education. Random Assignment in Program Evaluation and Intervention Research: Questions and Answers. 2003. Retrieved September 24, 2005 from: <http://www.ed.gov/rschstat/eval/resources/randomqa.pdf>
- Van Kleeck, A. Joint caregiver-child storybook reading: A route to literacy development. In: Stone, CA.; Silliman, ER.; Ehren, BJ.; Apel, K., editors. *Handbook of language and literacy: Development and disorders*. New York: Guilford Press; 2004. p. 175-208.
- Vaughn S, Kim A, Sloan CVM, Hughes MT, Elbaum B, Sridhar D. Social skills interventions for young children with disabilities: A synthesis of group design studies. *Remedial and Special Education*. 2003; 24:2–15.
- \*. Verhallen MJAJ, Bus AG, de Jong MT. The promise of multimedia stories for kindergarten children at-risk. *Journal of Educational Psychology*. 2006; 98:410–419.
  - \*. Wasik BA, Bond MA. Beyond the pages of a book: Interactive book reading and language development in preschool classrooms. *Journal of Educational Psychology*. 2001; 93(2):243–250.
  - \*. Wasik B, Bond M, Hindman A. The effects of a language and literacy intervention on Head Start children and teachers. *Journal of Educational Psychology*. 2006; 98(1):63–74.
- Wasik, BH.; Hendrickson, JS. Family literacy practices. In: Stone, CA.; Silliman, ER.; Ehren, BJ.; Apel, K., editors. *Handbook of language and literacy development and disorders*. New York: Guilford; 2004. p. 154-174.
- \*. Whitehurst GJ, Arnold DS, Epstein JN, Angell AL, Smith M, Fischel JE. A picture book reading intervention in day care and home for children from low-income families. *Developmental Psychology*. 1994; 30(5):679–689.

Table 1

Features of Intervention Studies

Study	Study Design	Number of Participants	Age (Grade)	# of Sessions (frequency and/or duration)	Person Implementing	Type of Intervention	Type of Outcome
Boykin & Cunningham (2001) <i>Fidelity reported: No</i>	Multiple Treatment	64	84–96 months (NR)	NR	Researcher	Other	Comprehension
Coyne, Lazarus, McCaugh & Kapp (2007) <i>Fidelity reported: Yes</i>	Multiple Treatment	32	64–84 months (K)	3 sessions (NR)	Researcher	Vocabulary	Vocabulary
Dickinson & Smith (1994) <i>Fidelity reported: No</i>	Multiple Treatment	25	48 – 60 months (NR)	NR	Teacher	Dialogic	Vocabulary;
Pemberton & Watkins (1987) <i>Fidelity reported: No</i>	Multiple Treatment	20	38 – 57 months (Pre-K)	8 sessions (2/ week × 4 weeks)	Researcher	Repeated Reading	Language Development; Vocabulary
Reutzel, Hollingsworth, & Elledge (1994) <i>Fidelity reported: Yes</i>	Multiple Treatment	79	NR (2)	80 sessions (daily × 16 weeks; 30-min)	Teacher	Dialogic	Word Recognition; Vocabulary; Comprehension
Beck & McKeown (2007) <i>Fidelity reported: Yes</i>	Single Group	76	NR (K – 1st)	45 sessions (daily × 9 weeks; 20 min)	Teacher	Vocabulary	Vocabulary
Davies, Shanks, & Davies (2004) <i>Fidelity reported: No</i>	Single Group	31	59 months (K/1)	24 sessions (3/ week × 8 weeks; 40 min)	Speech Teacher Paraprofessional	Dialogic	Comprehension

Study	Study Design	Number of Participants	Age (Grade)	# of Sessions (frequency and/or duration)	Person Implementing	Type of Intervention	Type of Outcome
Beilon-Ham, Hoffman, & Ham (2004) <i>Fidelity reported: No</i>	Single Subject	3	66 – 72 months (NR)	# sessions NR (4 weeks)	Researcher	Other	Language Development
Aram & Biron (2004) <i>Fidelity reported: No</i>	Treatment/Comparison	95	36–60 months (Pre-K)	66 sessions (2/ week; 33 weeks)	Researcher	Dialogic	Phonological Awareness; Word Recognition; Comprehension
Beck & McKeown (2007) <i>Fidelity reported: Yes</i>	Treatment/Comparison	98	NR (K – 1st)	50 sessions (daily × 10 weeks)	Teacher	Vocabulary	Vocabulary
Byrge (1994) <i>Fidelity reported: No</i>	Treatment/Comparison	29	91 months (NR)	115 sessions (daily × 23 weeks)	Teacher	Limited Questioning	Language Development; Vocabulary; Comprehension
Combs (1987) <i>Fidelity reported: No</i>	Treatment/Comparison	24	NR (K)	# sessions NR (3 weeks)	Teacher	Dialogic	Comprehension
Coyne & Kapp (2007) <i>Fidelity reported: Yes</i>	Treatment/Comparison	31	64–78 months (K)	3 sessions (20–30 min)	Researcher	Vocabulary	Vocabulary
Coyne & Simms (2004) <i>Fidelity reported: No</i>	Treatment/Comparison	64	NR (K)	NR (21 weeks)	Teacher	Dialogic	Phonological Awareness; Vocabulary
Kamekuni, & Stoolmiller (2004) <i>Random assignment* Fidelity reported: No</i>	Treatment/Comparison	36	36–60 months (Pre-K)	20 Sessions (Daily × 4 weeks)	Teacher	Dialogic	Vocabulary
Justice & Ezell (2002) T: 15	Treatment/Comparison	30	41–62 months (K)	32 sessions (4/ week; 8 weeks)	Researcher	Repeated Reading	Phonological Awareness; Word Recognition

Study	Study Design	Number of Participants	Age (Grade)	# of Sessions (frequency and/or duration)	Person Implementing	Type of Intervention	Type of Outcome
Random assignment* <i>Fidelity reported: Yes</i>							
Justice, Meier & Walpole (2005)	Treatment/Comparison	57	60–77 months (K)	20 sessions (1–3 days/week × 10 weeks; 20 min)	Researcher (Graduate Students) Teacher	Vocabulary	Vocabulary
Random assignment* <i>Fidelity reported: Yes</i>							
Korat & Shamir (2007)	Treatment/Comparison	128	60–70 months (K)	T1: 3 sessions (2 weeks; 30 min) T2: 3 sessions (NR)	T1: Child/e-book T2: Researcher (undergraduate student)	T1: Computer assisted T2: Limited Questioning	Vocabulary; Phonological Awareness; Recognition Word
Random assignment* <i>Fidelity reported: No</i>							
Lonigan, Anthony, Bloomfield, Dyer, & Samwel (1999)	Treatment/Comparison	95	25–64 months (NR)	# sessions (6 weeks)	Researcher (undergraduate volunteer)	Dialogic	Phonological Awareness; Vocabulary; Comprehension
Random assignment* <i>Fidelity reported: Yes</i>							
Lonigan & Whitehurst (1998)	Treatment/Comparison	58	33–60 months (NR)	# sessions NR (6 weeks)	Teacher or Paraprofessional	Dialogic	Vocabulary; Language Development
Random assignment* <i>Fidelity reported: No</i>							
Morrow (1984)	Treatment/Comparison	254	NR (K)	8 sessions (1/week; 8 weeks)	Student Teacher	Limited Questioning	Comprehension
Random assignment* <i>Fidelity reported: No</i>							
Morrow (1988)	Treatment/Comparison	79	48 months (NR)	10 sessions (1/week; 10 weeks)	Researcher	Dialogic	Comprehension
Random assignment* <i>Fidelity reported: No</i>							



Study	Study Design	Number of Participants	Age (Grade)	# of Sessions (frequency and/or duration)	Person Implementing	Type of Intervention	Type of Outcome
Peskin & Astington (2004) <i>Fidelity reported: No</i>	Treatment/Comparison	48	49–59 months (K)	20 sessions (daily × 4 weeks)	Teacher Researcher	Other	Language Development; Vocabulary; Comprehension
Talley, Lancy, & Lee (1997) <i>Fidelity reported: Yes</i>	Treatment/Comparison	73	48 months (Pre-K)	NR	Computer	Computer Assisted	Print Concepts
Ulanoff & Pucci (1999) <i>Fidelity reported: No</i>	Treatment/Comparison	60	NR (3)	NR	Teacher	Other	Vocabulary
Verhaeghen, Bus & de Jong (2006) Random assignment* <i>Fidelity reported: No</i>	Treatment/Comparison	60	66.3–67.6 months (K)	4 sessions (2 weeks; 6 min)	Researcher	Computer Assisted	Vocabulary; Language; Comprehension
Wasik & Bond (2001) <i>Fidelity reported: No</i>	Treatment/Comparison	121	47–55 months (Pre-K)	# sessions NR (15 weeks)	Teacher	Dialogic	Vocabulary
Wasik, Bond & Himmelman (2006) <i>Fidelity reported: Yes</i>	Treatment/Comparison	207	32–58 months (Pre-K)	# sessions NR (daily × 9 months)	Teacher or paraprofessional	Dialogic	Print Concepts; Vocabulary
Whitehurst, Arnold, Epstein, Angell, Smith, & Fischel (1994) Random assignment* <i>Fidelity reported: Yes</i>	Treatment/Comparison	48	36 months (NR)	# sessions NR (6 weeks)	Teacher or Paraprofessional	Dialogic	Vocabulary

Table 2

Study Findings and Effect Sizes

Design/Intervention	Measures	Results/Findings
<p>Boykin &amp; Cunningham (2001)</p> <ul style="list-style-type: none"> <li>T1 (<i>Low Movement Theme</i>): The behaviors and activities of the characters reflected few expressive movement themes. In the first context, children were told to stand or sit while listening to the story. They then answered questions on that story. In the second context, children were told that they could dance and clap if they wanted to while listening to the story. They then answered questions on that story.</li> <li>T2 (<i>High Movement Theme</i>): The behaviors and activities of the characters reflected expressive movement themes, such as running, dancing, jumping. Same contexts as T1</li> </ul>	Comprehension	Children exposed to the high movement theme stories performed better than those exposed to low movement theme stories.
<p>Coyne, McCoach &amp; Kapp (2007) Study 2</p> <ul style="list-style-type: none"> <li>T1 (<i>Extended Word Instruction</i>): Variety of interactive activities during and following read aloud focused on recognizing, discussing, and answering questions related to target words encountered during read alouds.</li> <li>T2 (<i>Embedded Word Instruction</i>): During read aloud, simple definition of target words were provided and then sentences were reread with definition replacing the target word</li> </ul>	Vocabulary: Expressive Vocabulary: Receptive Vocabulary: Context	T1>T2 ES = 1.70 T1>T2 ES = 0.99 T1>T2 ES = 1.12
<p>Dickinson &amp; Smith (1994)</p> <ul style="list-style-type: none"> <li>T1 (<i>Co-constructive Approach</i>): Read-aloud characterized by high amounts of talk by both children and teachers during read-aloud.</li> <li>T2 (<i>Didactic – Interactional Approach</i>): Engaged in limited amounts of talk before, during, and after read-aloud.</li> <li>T3 (<i>Performance Oriented Approach</i>): Read-aloud characterized by talk before and after read-aloud; little talk during read-aloud.</li> </ul>	PPVT-R*	Children in T3 outperformed children in T2
<p>Pemberton &amp; Watkins (1987)</p> <ul style="list-style-type: none"> <li>T1 (Modeling Reading): Read-aloud story with same content as T2, but no recasting.</li> <li>T2 (Recast Reading): Read-aloud story written with base sentence-recast pairings (e.g. This is a frog. He is a big, green frog); followed the same plot as T1 story.</li> </ul>	Vocabulary, Syntax	T1>T2 ES = 0.05 T1>T2 ES = 0.06
<p>Reutzel, Hollingsworth, &amp; Eldrege (1994)</p> <ul style="list-style-type: none"> <li>T1 (<i>Shared Book Experience</i>): Texts introduced by big book read alouds, children read text with partners, echo, and choral reading</li> <li>T2 (<i>Oral Recitation Lesson</i>): Texts read aloud, children practice, rehearse, and recite texts</li> </ul>	Word analysis Oral reading words per minute Oral reading error rate Oral reading self correction Oral retelling Answer explicit questions Answer implicit questions ITBS Comprehension* ITBS Vocabulary*	T1<T2 ES = -1.36 T1>T2 ES = 0.63 T1<T2 ES = -1.01 T1>T2 ES = 0.20 T1>T2 ES = 0.09 T1>T2

Design/Intervention	Measures	Results/Findings
<p>Beck &amp; McKeown (2007) Study 2</p> <ul style="list-style-type: none"> <li>T1 (<i>Text Talk Rich Instruction</i>): Read alouds followed with rich vocabulary instruction and discussion</li> <li>T2 (<i>Text Talk More Rich Instruction</i>): Read alouds followed with more rich vocabulary instruction and discussion (i.e. instruction on words was more frequent and for a longer duration)</li> </ul>	<p>Vocabulary: # "rich" words Vocabulary: # "more rich" words</p>	<p>ES = 0.25 T1&gt;T2 ES = 0.85 T1&gt;T2 ES = 0.13 T1&gt;T2 ES = 0.19</p> <p>Pre to post test gains in knowledge of "rich" words and "more rich" words statistically significant in kindergarten and first grades.</p>
<p>Davies, Shanks, &amp; Davies (2004)</p> <ul style="list-style-type: none"> <li>T (<i>Narrative skills training</i>): Children encouraged to derive own endings for read aloud stories, retell stories, and make up own stories using puppets and role plays; structures of narratives emphasized.</li> </ul>	<p>RAPT – I* RAPT – G* Bus story information* Bus story communication* Bus story episode quality* Bus story number of connectives* Bus story connective type* Story type</p>	<p>Significant improvement in the quantity of information children presented in their stories and in the structure of children' stories</p>
<p>Bellon-Harn, Hoffman, &amp; Ham (2004)</p> <ul style="list-style-type: none"> <li>T (<i>Repeated Story Book Reading</i>): Using same book for a week, researchers responded to children' responses by using a cloze procedure with expansion, a cloze procedure without expansion, or a contrast word procedure depending on interaction.</li> </ul>	<p>PCC, Percent Occurrence of Phonological Processes</p>	<p>During treatment condition, all participants produced more descriptive and interpretive utterances than labeling utterances.</p>
<p>Aram &amp; Biron (2004)</p> <ul style="list-style-type: none"> <li>T (<i>Joint Reading</i>): Open-ended questions, drawing, and drama were used to discuss read-aloud stories.</li> <li>C (<i>Comparison</i>): No description given.</li> </ul>	<p>Comprehension Word Recognition Phonological Awareness</p>	<p>T &lt; C ES = -0.42 T &gt; C ES = 1.56 T &gt; C ES = 1.26</p>
<p>Beck &amp; McKeown (2007) Study 1</p> <ul style="list-style-type: none"> <li>T (<i>Text Talk</i>): Read alouds followed with rich vocabulary instruction and discussion</li> <li>C: (<i>Read aloud</i>) Typical read aloud with similar trade-books</li> </ul>	<p>Vocabulary: Number of words</p>	<p>Kindergarten: T&gt;C ES = 0.41 First Grade: T&gt;C ES = 1.36</p>
<p>Bygrave (1994)</p> <ul style="list-style-type: none"> <li>T1 (<i>Music program</i>): Musical activities, such as singing and movement, which provided opportunities for children to acquire knowledge of musical skills and concepts</li> <li>T2 (<i>Storytelling program</i>): Read alouds aimed at improving children organization, comprehension, and memory skills</li> </ul>	<p>PPVT-R* LAC-R* Rhyme test Token Test* Basic language concepts screening test</p>	<p>Authors reported statistically significant pre to post test gains for children in the combination program (T3) on the Token Test.</p>

Design/Intervention	Measures	Results/Findings
<ul style="list-style-type: none"> <li><i>T3 (Combination program)</i>: Combination of music and story telling programs</li> <li><i>C (Control)</i>: No program</li> </ul>	Aural receptive sound concept test	
<p>Combs (1987)</p> <ul style="list-style-type: none"> <li><i>T (Modeled Approach)</i>: Modeled aspects of reading process using enlarged texts during read-aloud.</li> <li><i>C (Comparison)</i>: Read books aloud with emphasis on enjoying the story; few questions were asked.</li> </ul>	Probed story recall; free recall	T > C in listening comprehension. Below average readers doubled recall (43% at pretest to 86% at posttest)
<p>Coyne, McCoach &amp; Kapp (2007) Study 1</p> <ul style="list-style-type: none"> <li><i>T (Extended Word Instruction)</i>: Variety of interactive activities during and following read aloud focused on recognizing, discussing, and answering questions related to target words encountered during read alouds.</li> <li><i>C (Comparison)</i>: No instruction</li> </ul>	Vocabulary: Expressive Vocabulary: Receptive Vocabulary: Context	T > C ES = 2.66 T > C ES = 1.03 T > C ES = 1.19
<p>Coyne, Simmons, Kame'enui, &amp; Stoolmiller (2004)</p> <ul style="list-style-type: none"> <li><i>T (Storybook)</i>: 3 target vocabulary words taught before the read aloud, after the read aloud children were given opportunities to retell stories using target vocabulary and selected illustrations as prompts</li> <li><i>C (Control)</i>: Sounds and letters module of a commercial reading program</li> </ul>	Vocabulary: Taught Words Vocabulary: Untaught Words	T > C ES = 0.74 T > C ES = 0.20
<p>Hargrave &amp; Senechal (2000)</p> <ul style="list-style-type: none"> <li><i>T (Dialogic Reading)</i>: Asked wh-questions, followed correct responses with another questions, and emphasized importance of enjoying stories during read-alouds.</li> <li><i>C (Comparison)</i>: Typical reading instruction.</li> </ul>	PPVT - R*, EOWPVT - R*, Book Vocabulary	T < C ES = -0.12 T > C ES = 0.17 T > C ES = 0.07
<p>Justice &amp; Ezell (2002)</p> <ul style="list-style-type: none"> <li><i>T (Print-Focused)</i>: During read-aloud, prompts for print conventions, concepts of words, or alphabet knowledge were given.</li> <li><i>C (Picture-Focused)</i>: During read-aloud, prompts regarding pictures (such as characters, perceptual focus, or action focus) were given.</li> </ul>	Print Concepts Print Recognition Words in Print Letter Orient. & Discrimination Alphabet Knowledge Literacy Terms	T > C ES = 0.41 T > C ES = 3.28 T > C ES = 1.45 T > C ES = 0.72 T > C ES = 0.48 T > C ES = 0.67
<p>Justice, Meier &amp; Walpole (2005)</p> <ul style="list-style-type: none"> <li><i>T (Word Elaboration)</i>: Target words were explicitly defined as they occurred in the text followed by use of the word in context.</li> </ul>	Vocabulary: Elaborated words Vocabulary: Non-elaborated words	T > C ES = 1.59 T > C ES = 2.06

Design/Intervention	Measures	Results/Findings
<ul style="list-style-type: none"> <li>• C (<i>Comparison</i>): Regular kindergarten curriculum</li> </ul>	<p>Vocabulary Word Recognition Phonological Awareness</p>	<p>T1 &gt; C ES = 1.00 T2 &gt; C ES = 0.89 T1 &gt; C ES = 0.10 T2 &lt; C ES = -0.10 T1 &gt; C ES = 0.18 T2 &lt; C ES = -0.09</p>
<p>Korat &amp; Shamir (2007)</p> <ul style="list-style-type: none"> <li>• T1 (<i>E-book</i>): Children read story via computer in "read with dictionary," mode (oral reading with explanation of difficult words) as well as in "read story and play" mode (designed to enhance phonological awareness and comprehension).</li> <li>• T2 (<i>Adult Readers</i>): Adults completed read aloud with prescribed instructions that included pre-reading activities (predictions and introduction of vocabulary words and their definitions) as well as concrete and inferential questions during and after reading.</li> <li>• C (<i>Comparison</i>): Typical practice.</li> </ul> <p>Lonigan, Anthony, Bloomfield, Dyer, &amp; Samwel (1999)</p> <ul style="list-style-type: none"> <li>• T1 (<i>Typical Shared Reading</i>): Read text aloud, commented on pictures aloud, and asked children if there were any questions</li> <li>• T2 (<i>Dialogic Reading</i>): Read text aloud, ask children to name or describe objects or actions from text, and asks open-ended questions</li> <li>• C (<i>Control</i>): Typical preschool curriculum</li> </ul>	<p>WJ-LC* PA: Alliteration PA: Blending PA: Elision PA: Rhyme EOWPVT-R* ITPA-VE* PPVT-R*</p>	<p>T1 &lt; C ES = -0.23 T2 &lt; C ES = -0.07 T1 &lt; T2 ES = -0.15 T1 &gt; C ES = 0.98 T2 &gt; C ES = 1.43 T1 &lt; T2 ES = -0.32 T1 &gt; C ES = 0.03 T2 &gt; C 0.04 T1 &lt; T2 ES = -0.01 T1 &lt; C ES = -0.30 T2 &gt; C ES = 0.11 T1 &lt; T2 ES = -0.37 T1 &lt; C ES = -0.04 T2 &gt; C ES = 0.56 T1 &lt; T2 ES = -0.40 T1 &gt; C ES = 0.07 T2 &lt; C ES = -0.03 T1 &gt; T2 ES = 0.11 T1 &lt; C ES = -0.01 T2 &lt; C</p>

Design/Intervention	Measures	Results/Findings
<p>Lonigan &amp; Whitehurst (1998)</p> <ul style="list-style-type: none"> <li><i>T1 (Dialogic reading)</i>: Read text aloud, ask children to name or describe objects or actions from text, and asks open-ended questions. High compliance center.</li> <li><i>T2 (Dialogic reading)</i>: Read text aloud, ask children to name or describe objects or actions from text, and asks open-ended questions.</li> </ul>	<p>Semantic Complexity: Number of Different Words* Semantic Complexity:</p>	<p>ES = -0.01 T1 = T2 ES = 0.00 T1 &gt; C ES = 0.04 T2 &lt; C ES = -0.49 T1 &gt; T2 ES = 0.49</p> <p>T1 &gt; C1 ES = 0.96 T2 &gt; C2 ES = 0.97 T1 &gt; C1 ES = 0.18</p>
<p>Low compliance center.</p> <ul style="list-style-type: none"> <li><i>C1 (Control)</i>: No specific instructions or activities. High compliance center.</li> <li><i>C2 (Control)</i>: No specific instructions or activities. Low compliance center.</li> </ul>	<p>Number of Modifiers* Semantic Complexity: Number of Nouns Semantic Complexity: Number of Verbs Language Production: Number of Words Language Production: Mean Length of Utterance PPVT-R* EOWPVT* ITPA - VE*</p>	<p>T2 &gt; C2 ES = 1.04 T1 &lt; C1 ES = -0.10 T2 &gt; C2 ES = 1.38 T1 &gt; C1 ES = 0.73 T2 &gt; C2 ES = 0.13 T1 &gt; C1 ES = 1.17 T2 &gt; C2 ES = 1.31 T1 &gt; C1 ES = 0.27 T2 &gt; C2 ES = 0.18 T1 &lt; C1 ES = -0.74 T2 &gt; C2 ES = 0.02 T1 &lt; C1 ES = -0.86 T2 &gt; C2 ES = 0.29 T1 &lt; C1 ES = -0.54 T2 &gt; C2 ES = 0.34</p>
<p>Morrow (1984)</p> <ul style="list-style-type: none"> <li><i>T1 (Structural Discussion)</i>: Questions posed to enhance listening comprehension before and after read aloud focused on setting, theme, characters, episodes, and resolution.</li> <li><i>T2 (Traditional Discussion)</i>: Questions posed to enhance listening comprehension before and after read aloud focused on recalling facts and details, cause and effect relationships,</li> </ul>	<p>Traditional Questions Story Structure Questions</p>	<p>Middle and high achieving students' scores increased. None of the treatments increased the performance of low achieving students.</p>



Design/Intervention	Measures	Results/Findings
<p>classifying, interpreting characters' feelings, problem solving, and relating the story to children's experiences.</p> <ul style="list-style-type: none"> <li>T3 (<i>Combined Structural and Traditional Discussion</i>): Questions posed to enhance listening comprehension before and after read aloud focused on both story structure elements and traditional discussion elements.</li> <li>C (<i>Comparison</i>): No questions or discussion initiated to enhance listening comprehension before or after read aloud.</li> </ul>		
<p>Morrow (1988)</p> <ul style="list-style-type: none"> <li>T1 (<i>Different Book Reading</i>): Listened to a different story read aloud each week.</li> <li>T2 (<i>Repeated Book Reading</i>): Listened to 3 different books read 3 times each.</li> <li>C (<i>Reading Readiness Activities</i>): Focused on visual and auditory discrimination skills using a commercial program.</li> </ul>	<p>Comprehension: Question Answering            Comprehension: Focus on Meaning            Comprehension: Story Structure            Comprehension: Focus on Print            Comprehension: Illustration Related</p>	<p>T1 &gt; C            ES = 0.74            T2 &gt; C            ES = 3.18            T1 &gt; C            ES = 2.04            T2 &gt; C            ES = 2.84            T1 &gt; C            ES = 1.60            T2 &gt; C            ES = 2.76            T1 &gt; C            ES = 1.47            T2 &gt; C            ES = 2.27            T1 &gt; C            ES = 3.10            T2 &gt; C            ES = 0.38</p>
<p>Peskin &amp; Astington (2004)</p> <ul style="list-style-type: none"> <li>T (<i>Metacognitive language</i>): Read aloud texts included numerous metacognitive verbs; teachers encouraged to use metacognitive verbs during text discussions</li> <li>C (<i>Comparison</i>): Read aloud texts were edited to remove metacognitive verbs</li> </ul>	<p>TELD = 3*            False belief prediction            False belief explanation            Metacognitive comprehension            Storytelling metacognitive types            Storytelling metacognitive tokens            Storytelling total words</p>	<p>T &gt; C            ES = 0.03            T &gt; C            ES = 0.22            T &lt; C            ES = -0.51            T &gt; C            ES = 0.27            T &gt; C            ES = 0.06            T &gt; C            ES = 0.10            T &gt; C            ES = 0.34</p>
<p>Talley, Lancy, &amp; Lee (1997)</p> <ul style="list-style-type: none"> <li>T (<i>CD-ROM Storybook Program</i>): Allowed children to select and control the pace of the story read aloud. Included pre- and post-reading activities</li> <li>C (<i>Comparison</i>): Children spent the same amount of time on the computer but did not use the CD-ROM storybook program</li> </ul>	<p>Print Awareness            Concepts about Print            Pictic</p>	<p>T &gt; C            ES = 0.62            T &lt; C            ES = -0.35            T &gt; C            ES = 0.26</p>
<p>Ulanoff &amp; Pucci (1999)</p>	<p>Vocabulary</p>	<p>T1 &lt; C            ES = -0.44</p>

Design/Intervention	Measures	Results/Findings
<ul style="list-style-type: none"> <li><i>T1 (Con-current Translation)</i>: During read-aloud, two languages used interchangeable or concurrently; care taken to avoid use of direct translation between languages</li> <li><i>T2 (Preview/Review)</i>: Used two languages separately. Preview and review activities conducted in dominant language; read-aloud conducted in target language</li> <li><i>C (Comparison)</i>: Read-aloud conducted in target language without explanation of stories.</li> </ul>		<p>T2 &gt; C ES = 0.64</p>
<p>Verhallen, Bus &amp; de Jong (2006)</p> <p>T1 (<i>Multimedia</i>): Story was read via “computer-produced voice” and related “action” pictures and sounds were included with a focus on narrative story elements. Story read 4X.</p> <p>T3 (<i>Multimedia</i>): Same as T1. Story read 1X</p> <p>T2 (<i>Static</i>): Related picture were displayed while “computer-produced voice” read story (no sounds) to children. Story read 4X</p> <p>T4; (<i>Static</i>): Same as T2. Story read 1X</p> <p>C1 (<i>Midnight Play</i>): Computer game composed of a series of unrelated screens was activated by clicking on screen (no spoken language)</p> <p>C2 (<i>Comparison</i>): Typical practice</p>	<p>Comprehension: Retelling a story Language: Syntax Vocabulary: Expressive</p>	<p>T1&gt;C1 ES = 3.25 T1 &gt; C2 ES = 3.60 T2 &gt; C2 ES = 2.38 T3 &gt; C1 ES = 1.37 T3 &gt; C2 ES = 1.68 T1&gt;C1 ES = 1.76 T1 &gt; C2 ES = 1.04 T2 &lt; C2 ES = -0.04 T3 &gt; C1 ES = 0.97 T3 &gt; C2 ES = 0.27 T1 &gt; C1 ES = 2.28 T1 &gt; C2 ES = 1.53 T2 &gt; C2 ES = 0.11 T3 &gt; C1 ES = 0.79 T3 &gt; C2 ES = 0.04</p>
<p>Wasik &amp; Bond (2001)</p> <ul style="list-style-type: none"> <li><i>T (Interactive Book Reading Techniques)</i>: Defined target vocabulary, used objects as visual representations, provided opportunities for children to use target vocabulary, and asked open-ended and reflective questions before, during, and after read-aloud.</li> <li><i>C (Comparison)</i>: Typical read alouds; teachers used same texts but were not trained in the interactive book reading techniques</li> </ul>	<p>PPVT – III* Vocabulary: Receptive Vocabulary: Expressive</p>	<p>T &gt; C ES = 1.59 T &gt; C ES = 2.06</p>
<p>Wasik, Bond, Hindman (2006)</p> <ul style="list-style-type: none"> <li><i>T</i>: (Book Reading and Oral Language Strategies): Teachers promoted reflective discussion during and after reading by introducing target vocabulary words, objects that represented target words, and open-ended questions regarding the objects. Active listening, rich language, and expanded feedback were also promoted.</li> </ul>	<p>PPVT-III* EQPVT-III* Alphabet Knowledge</p>	<p>T &gt; C ES = 0.74 T &gt; C ES = 0.44 T &lt; C ES = -0.35</p>

Design/Intervention	Measures	Results/Findings
<ul style="list-style-type: none"> <li>• C: Head Start centers were provided with a list of the books used in T. The books were displayed and used by teachers on an independent basis.</li> </ul> <p>Whitehurst, Arnold, Epstein, Angell, Smith, &amp; Fischel (1994)</p> <ul style="list-style-type: none"> <li>• T (<i>Dialogic reading</i>): Dialogic reading; books chosen because of their potential to support vocabulary growth</li> <li>• C (<i>Control</i>): Play activities in small groups with toys</li> </ul>	<p>PPVT-R* ITPA* EOWPVT* Our Word</p>	<p>T &gt; C ES = 0.13 T &lt; C ES = -0.03 T &gt; C ES = 0.18 T &gt; C ES = 0.06</p>

Note. T = Treatment; C = Comparison; PPVT = Peabody Picture Vocabulary Test; WPPSI = Wechsler Preschool and Primary Scale of Intelligence; PPVT-R = Peabody Picture Vocabulary Test – Revised; EOWPVT-R = Expressive One Word Picture Vocabulary Test – Revised; LMAC – R = Lindamood Auditory Conceptualization Test – Revised; PCC = Percentage of Consonants Correct;

\* norm-referenced measure; + This group did not meet inclusion criteria for use in synthesis or meta-analysis.

**Table 3**

## Descriptors for Article Categorization.

<b>Category</b>	<b>Descriptors</b>
Dialogic Reading	<ul style="list-style-type: none"> <li>• Authors call the intervention “dialogic reading.”</li> <li>• The storytelling role is gradually shifted from the adult reader to the child through various techniques (e.g. open ended questions, repetition, modeling; Ezell &amp; Justice, 2005).</li> <li>• Conversational turn-taking occurred. Turns may consist of verbal, vocal, or nonverbal responses (Ezell &amp; Justice, 2005).</li> <li>• The teacher asks questions and prompts the child to increase the quality of descriptions with the goal of the child learning to become the storyteller (Lonigan, 1998).</li> </ul>
Repeated Reading	<ul style="list-style-type: none"> <li>• Same book is read aloud on several occasions.</li> <li>• Repeated reading is the main component of the intervention.</li> </ul>
Limited Questioning	<ul style="list-style-type: none"> <li>• Questioning before, during, and/or after reading.</li> <li>• Does not engage in extended dialogue.</li> </ul>
Computer Assisted	<ul style="list-style-type: none"> <li>• Read aloud is conducted through the use of a computer</li> </ul>
Extended Vocabulary	<ul style="list-style-type: none"> <li>• Activities before, during and after reading focused on vocabulary words and/or concepts.</li> <li>• More than simply defining words. Activities must be conducted with the purpose of increasing word knowledge.</li> </ul>
Other	<ul style="list-style-type: none"> <li>• Read aloud of text is provided</li> <li>• No focus on dialogic reading, repeated reading, limited questioning, computer assisted, or extended vocabulary instruction.</li> </ul>

**Table 4**

Quality of treatment-comparison studies.

<b>Element</b>	<b>Number of Studies</b>
Random assignment to conditions	11
Fidelity of treatment reported	12
Standardized dependent measures	14
Random assignment, treatment fidelity, and standardized measures	2

Table 5

Type of intervention by study design

Intervention Type	Study Design				Marginal Totals
	Treatment-Comparison	Multiple Treatment	Single Group	Single Subject	
Dialogic	10	2	1	0	13
Extended Vocabulary	3	1	1	0	5
Repeated Reading	1	1	0	0	2
Limited Questioning	2	0	0	0	2
Computer Assisted	2	0	0	0	2
Other	2	1	0	1	4
Marginal Totals	20	5	2	1	28*

\* Korat & Shamir (2007) included a computer assisted treatment and a limited questioning treatment in their multiple treatment design study and are not included in the study totals.



Table 6

Grand Mean Weighted Effect Sizes from Unconditional Model

Outcome	Fixed Effects			Random Effects			
	Coefficient	SE	t	p-value	Variance Component	SD	p-value
Language	0.29	0.09	3.16	0.005	0.003	0.06	>.50
PA	0.78	0.15	5.23	<.001	0.42	0.65	<.001
PC	0.86	0.27	3.17	0.01	0.71	0.84	<.001
RC	0.70	0.14	4.94	<.001	0.28	0.53	<.001
Voc	1.02	0.18	5.83	<.001	1.48	1.21	<.001
WR	0.23	0.37	0.62	0.55	1.00	0.99	<.001

Note. PA = Phonological Awareness; PC = Print Concepts; RC = Reading Comprehension; Voc = Vocabulary; WR = Word Recognition

**Table 7**

Moderator Analysis for Outcomes

Outcome	Moderator	# of ES	Coefficient	SE	t	p-value	Variance Component	SD	p-value	R <sup>2</sup> Red.
PA	DR	19	0.84	0.15	5.49	<.001	0.41	0.64	<.001	0.02
	CA	2	-0.39	0.70	-0.56	0.58				
	LQ	1	-1.07	0.70	-1.53	0.14				
PC	DR	3	0.70	0.43	1.65	0.14	0.66	0.82	<.001	0.06
	RR	5	0.60	0.59	1.02	0.34				
	CA	3	-0.52	0.74	-0.71	0.50				
RC	DR	15	0.60	0.16	3.81	0.00	0.27	0.52	<.001	0.04
	CA	5	0.67	0.39	1.74	0.09				
	Other	1	-0.18	0.62	-0.30	0.77				
Voc	DR	30	0.57	0.19	2.97	0.005	1.06	1.03	<.001	0.28
	CA	6	1.15	0.50	2.32	0.03				
	LQ	2	1.43	1.07	1.33	0.19				
WR	Voc	10	-0.42	0.65	-0.65	0.52				
	Other	3	1.61	0.38	4.18	<.001				
	DR	4	0.04	0.49	0.08	0.94	1.16	1.07	<.001	-
RR	RR	2	1.95	1.31	1.49	0.21				
	LQ	1	-0.11	1.21	-0.09	0.93				
	CA	1	0.03	1.21	0.03	0.98				

Note. PA = Phonological Awareness; PC = Print Concepts; RC = Reading Comprehension; Voc = Vocabulary; WR = Word Recognition; CA = Computer Assisted; LQ = Limited Questioning; RR = Repeated Reading; DR = Dialogic Reading