EBP *briefs*

A scholarly forum for guiding evidence-based practices in speech-language pathology

Evidence-Based Vocabulary Instruction for Elementary Students via Storybook Reading

Carla J. Johnson and Erin Yeates
University of Toronto

EBP Briefs

A publication of AGS Publishing, now part of Pearson Assessements

Evidence-Based Vocabulary Instruction for Elementary Students via Storybook Reading

Carla J. Johnson and Erin Yeates
University of Toronto

Evidence-Based Vocabulary Instruction for Elementary Students via Storybook Reading

Carla J. Johnson and Erin Yeates
University of Toronto

Speech-language pathologists who work in educational settings recognize that students need strong vocabulary skills to succeed in school. Moreover, speech-language pathologists who are interested in evidence-based practice may wish to know whether there is scientific evidence to support particular methods of vocabulary instruction. In this brief, we evaluate the evidence on a specific educational question regarding vocabulary instruction. The question is posed using the PICO framework commonly used to guide inquiries in evidence-based practice (e.g., Centre for Evidence Based Medicine, 2001). In this framework, P stands for the patient group or population, I for the intervention being considered, C for a comparison intervention, and O for the outcome of interest. The specific PICO question addressed in this review is: Do elementary school students (kindergarten through grade 4) [P] learn new vocabulary [O] more effectively from hearing stories with brief explanations of unknown words [I] than from hearing stories without such explanations [C]?

Background and Rationale

Vocabulary knowledge is a cornerstone of language development and, eventually, of academic success. Children's vocabulary skills are strong predictors of syntactic development, reading comprehension, and overall intellectual ability (Miller, 1991).

Vocabulary acquisition is a mysterious and complex process. A single encounter with an unknown word in context may be enough to initiate a "fast mapping" between the sound of that word and a tentative meaning (Carey & Bartlett, 1978). With subsequent encounters, the quality of a child's

knowledge of a word becomes richer and more refined. It is difficult, however, to pinpoint precisely when a word becomes known. Is a word known when a child can: (a) choose a synonym for it from a list of multiple choices; (b) use the word correctly in a sentence; (c) provide an explicit definition of it; or (d) perform all of the above tasks? Is a word known when a child understands a single meaning for it (e.g., *down* as a spatial direction) or only when other common meanings are also acquired (e.g., *down* as feeling sad, *down* as goose feathers)? Is a word known when a child comprehends the meaning of its root form (e.g., smoke) or only after other inflected (e.g., smoking, smoked) and derived forms (e.g., smokeless, smokiness) are also learned?

Despite such complexities, it is clear that vocabulary growth occurs at an astonishing rate throughout childhood. On average, children learn nine new words per day in the preschool years (Carey, 1978) and this rate accelerates to more than 20 words per day during the early school years (Anglin, 1993). Young children probably acquire most of their new words implicitly by hearing them used in verbal contexts (Hart & Risley, 1995). For older children, independent reading becomes the main context for implicit learning of new words (Nagy & Anderson, 1984), because written language is much richer in rare vocabulary than is spoken language. Speech-language pathologists and other educators recognize the joint and reciprocal influences of both oral and written vocabulary in the development of reading proficiency:

Vocabulary growth occurs at an astonishing rate throughout childhood.

> **Oral vocabulary is a key to learning to make the transition from oral to written forms, whereas reading vocabulary is crucial to the comprehension processes of a skilled reader (National Reading Panel, 2000, Ch. 4, p. 15).**

Unfortunately, dramatic individual differences in early exposure to vocabulary input (Hart & Risley, 1995) are associated with corresponding large differences among children in resulting vocabulary knowledge. Thus, many students arrive at school with limited vocabularies that place them at high risk for poor academic performance (Walker, Greenwood, Hart, & Carta,

1994). In the early school years, the gap widens between students with low and high vocabulary skills (Biemiller & Slomin, 2001). If schools provided effective vocabulary instruction in the early grades, it might be possible to narrow this gap, or at least to improve the absolute level of vocabulary skill for children with the smallest vocabularies (Biemiller & Boote, 2006), perhaps alleviating their typical grade 4 slump in reading comprehension (Chall & Jacobs, 2003). Thus, for speech-language pathologists and other educators, enhancing vocabulary skills in the elementary years is an important objective. Accordingly, in this evidence review, we focus on vocabulary instruction methods for early elementary students.

Method

We initially conducted a general search of the scientific literature on vocabulary instruction methods for early elementary students to guide us in eventually narrowing our attention to a specific educational question. A general search seemed prudent because two recent systematic reviews of vocabulary instruction methods (Jitendra, Edwards, Sacks, & Jacobson, 2004; National Reading Panel, 2000, Ch. 4) both revealed that limited research had been directed toward vocabulary instruction methods for students in the early elementary years.

Inclusion Criteria

Prior to our search, we identified several inclusion criteria for selecting studies to be considered for this review (see Table 1). First, following the precedent set by McGinty and Justice (2006) in the first brief in this series, we selected only studies that employed an experimental design (randomized controlled trial; RCT), quasi-experimental design (QED), or regression discontinuity design (RD). Well-conducted studies with these types of designs provide the strongest evidence of treatment efficacy, that is, a causal relationship between an intervention and an outcome (Gersten et al., 2005). The second criterion was that the independent variable in the study was an identifiable method of vocabulary instruction; the third was that the dependent variable(s) included a measure of vocabulary knowledge. The fourth criterion was that participants in

the study were elementary students in the range from kindergarten to grade 4. The final criterion was that studies were published in English in peer-reviewed scientific journals.

Article Search

We first used the Scholar's Portal search engine to search a wide range of major databases in the social sciences, including PsychINFO and the Educational Resources Information Center (ERIC). The search terms included the keywords "vocabulary instruction" or "vocabulary teaching" or "vocabulary intervention", combined with the keywords "children" or "students". This initial search yielded 575 hits that were peer-reviewed journal articles published in English. This main search was also complemented by similar searches of the journal archive of the American Speech-Language-Hearing Association, the medical literature on PubMed, the What Works Clearinghouse, and the Web of Science. We also looked for additional relevant studies by examining reference lists of (a) articles meeting our criteria, and (b) systematic reviews on vocabulary instruction (Blok, Van Daalen-Kapteijns, Otter, & Overmaat, 2001; Bryant, Goodwin, & Bryant, 2003; Fukkink & de Glopper, 1998; Jitendra et al., 2004; National Reading Panel, 2000; Swanborn & de Glopper, 1999). Follow-up searches were also conducted to identify recent articles that had cited relevant studies or reviews. Abstracts for identified articles were then reviewed to select studies that warranted further evaluation of the full-text version.

Forty-eight full-text articles were retrieved and reviewed. Of these, 31 failed to meet one or more of our criteria: 12 did not meet the experimental design criterion; 8 described reading instruction rather than vocabulary instruction or included an unspecified mix of vocabulary instruction methods; 5 did not include a measure of vocabulary outcome; and 6 included participants outside of our target grade range.

The seventeen remaining studies were read and categorized according to the type of vocabulary instruction technique, the nature of the comparison condition (no treatment or another vocabulary technique), the grade of the participants, and whether or not the participants were "at-risk" for academic

difficulties (e.g., low SES, limited English proficiency, identified with language and/or learning impairments). Several general conclusions were apparent upon review of the categorized results. First, by far the most common vocabulary instruction technique was to have elementary school children listen to stories that were read to them by adults. Second, across individual studies, story readings were embellished in a variety of different ways in an effort to promote vocabulary acquisition. These embellishments included explanations for unfamiliar words, multiple readings of the stories, and questions or other methods to encourage active engagement of children with the story. Third, the most frequent embellishment studied was the provision of explanations for unfamiliar words compared with a condition in which the story was read without explanations. This specific contrast was studied in five articles which, therefore, became the basis for the PICO question of interest in this review: Do elementary school students (kindergarten through grade 4) [P] learn new vocabulary [O] more effectively from hearing stories with brief explanations of unknown words [I] than from hearing stories without such explanations [C]? Appendix 1 lists citations for the articles on vocabulary instruction that met our initial search criteria but did not address the specific PICO question.

> Five studies compared children's vocabulary gains after hearing stories read with and without explanations of target word meanings.

Results

Description of Included Studies

As shown in Table 2, the five studies included in this review consisted of two randomized controlled trials (RCTs; Justice, Meier, & Walpole, 2005; Penno, Wilkinson, & Moore, 2002) and three quasi-experimental designs (QEDs; Biemiller & Boote, 2006, Study 1; Brett, Rothlein, & Hurley, 1996; Elley, 1989, Study 2). A total of 573 elementary students participated in the five studies: 104 in kindergarten (Biemiller & Boote, 2006, Study 1; Justice et al., 2005); 84 in grade 1 (Biemiller & Boote, 2006, Study 1; Penno et al., 2002), 210 in grade 2 (Biemiller & Boote, 2006, Study 1; Elley, 1989, Study 2), and 175 in grade 4 (Brett et al., 1996). In two studies, a relatively large percentage of students were "at-risk" for academic problems because of low

socioeconomic status, limited English proficiency, or poor vocabulary skills (Biemiller & Boote, 2006, Study 1; Justice et al., 2005). No specific risk factors were identified for participants in the other three studies, in which participants were drawn from general education classrooms. All five studies employed experimenter-designed outcome measures of targeted vocabulary items, chosen from the books that were read in the interventions. Multiple-choice measures of vocabulary were used in three studies (Brett et al., 1996; Elley, 1989; Penno et al., 2002). Definitions of the target words were elicited in two studies (Biemiller & Boote, 2006, Study 1; Justice et al., 2005). Penno et al. (2002) also included a second outcome measure, the use of target vocabulary items in retelling of the stories.

All five studies compared children's vocabulary gains after hearing stories read in two different conditions, with and without explanations of target word meanings. Three studies also included a control condition in which the books were not read to students (Brett et al., 1996; Elley, 1989, Study 2; Justice et al., 2005). Appendix 2 gives examples of the books read to children in the different grades, the target vocabulary items taught, and the types of explanations given for target meanings. Vocabulary instruction took place in small groups (Justice et al., 2005; Penno et al., 2002) or in entire classrooms (Biemiller & Boote, 2006, Study 1; Brett et al., 1996; Elley, 1989, Study 2). The number of readings of each book also varied across studies, from one (Brett et al., 1996) to as many as four (Biemiller & Boote, 2006, Study 1; Justice et al., 2005). In most cases, the vocabulary interventions were relatively short, lasting one to two weeks; the intervention in Justice et al. (2005) was the longest, conducted over a period of ten weeks.

Quality of Evidence

A key element of evidence-based practice is the critical appraisal of the methodological quality of studies that forms the basis for clinical decision-making. Accordingly, we adopted, with some slight modifications, the quality criteria proposed by Troia (1999) to appraise the five studies included in our review. The results of that appraisal are shown in Table 3. We independently scored each study on 29 attributes, which assessed both internal and external

validity. Inter-rater reliability was 86%. Discrepancies were resolved via discussion to arrive at the final quality assessments.

Examination of Table 3 shows that the five studies varied considerably in their methodological rigor. Following Troia (1999), we weighted each criterion on a three-point scale according to its importance to overall methodological quality, with 1 representing factors that were desirable but not essential and 3 representing factors that were critical for drawing causal conclusions. For each study, we summed these weightings to arrive at a total methodological quality score, which was also expressed as a percentage relative to the maximum possible score for the applicable criteria. The five studies ranged from 39% (Elley, 1989, Study 2) to 72% (Justice et al., 2005) in the extent to which they met the quality criteria. Troia (1999) reported a similar range of quality ratings for studies of phonological awareness interventions. In conclusion, then, the five vocabulary studies reviewed here were not as methodologically sound as might be desired for making evidence-based decisions. Nonetheless, as a group, they represented the best evidence currently available to speech-language pathologists and educators regarding this PICO question.

The five studies varied considerably in their methodological rigor.

Estimated Effects: Listening to Storybooks With and Without Explanations of Words

For each study, Table 2 also includes calculations of standardized posttest effect size estimates (Cohen, 1988) for the key comparison of vocabulary learning after hearing stories with and without brief explanations of target words. Also included are 95% confidence intervals for the effect sizes based on posttest data provided by the study authors. The effect size estimates represent, in standard deviation units, the mean difference between the two conditions in which books were heard with and without explanations of target words. For example, an effect size of .50 would represent a difference of one-half of a standard deviation unit, in this case favoring the condition with target word explanations over that without explanations. By convention (Cohen, 1988), effect size estimates of .20 are generally interpreted as small; those of .50 are medium; and those of .80 are large.

With one exception, the effect sizes calculated in this review showed that students learned more vocabulary items from story readings with explanations of target words than from readings without explanations. The exception was for children with high vocabulary skills in Justice et al. (2005), who showed a slight disadvantage in learning from hearing words with explanations as compared to those without explanations, *d* = -.22; *CI* = -1.64 to 1.45. In all other cases, effect sizes suggested a benefit from hearing explanations of target words. However, as shown in Table 2, the size of that benefit varied widely across studies, from a medium-sized effect for students with low vocabulary skills in Justice et al. (2005), *d* = .48; *CI* = -1.62 to 2.91, to large effects in Brett et al. (1996) for Book 1, *d* = 1.47; *CI* = .73 to 1.83, and for Book 2, *d* = 1.30; *CI* = .65 to 1.66. Small sample sizes and large standard deviations, however, meant that many confidence intervals included zero, signaling that this advantage of word explanations may not be reliable.

Two studies also included a control condition in which the stories were not heard and for which sufficient information was reported to compute effect sizes. Compared to a no story condition, hearing stories with explanations of target words conferred a large and consistent advantage in vocabulary learning [Brett et al., (1996), Book 1, *d* = 1.35, *CI* = .60 to 1.65; Book 2, *d* = 1.09, *CI* = .45 to 1.52; Justice et al. (2005), *d* = 1.22]. However, when stories were read without explanations of target words, benefits for vocabulary learning were not found consistently relative to the control condition without stories [Brett et al., (1996), Book 1, *d* = -.33, *CI* = -.70 to -.03; Book 2, *d* = -.17, *CI* = -.53 to .26; Justice et al. (2005), *d* = .53]. Taken together, these effect size comparisons indicated that, of the three conditions studied, vocabulary learning was best when explanations of target words were included in the stories read to children.

In evidence-based practice, standardized effect sizes are useful to compare results across studies, but their meaning in terms of the possible impact of interventions is not always transparent. To judge impact, it is often useful to

Compared to a no story condition, hearing stories with explanations of target words conferred a large advantage in vocabulary learning.

think in terms of units that are more easily understood. Biemiller and Boote (2006), for example, summarized the results of vocabulary interventions with a more interpretable measure, namely, the total number of words that students learned. Adopting that measure here, we compared our studies in terms of how many words children learned in the relevant instructional conditions (E = explanation; NE = without explanation; and C = control, no books). For Justice et al. (2005), low vocabulary students acquired 2.3 (E), 1.2 (NE), and .75 (C) words, whereas high vocabulary students acquired 2.7 (E), 2.45 (NE), and 1.48 (C) words, respectively. The numbers of words learned were 5.28 (E) and 2.88 (NE) in Biemiller and Boote (2006, Study 1); 2.7 (E) and .6 (NE) in Penno et al. (2002); 5.8 (E), 2.0 (NE), and .2 (C) in Elley (1989, Study 2, Book 1); and 6.6 (E), -.2 (NE), and .4 (C) in Brett et al. (1996). These values may seem small, but it must be remembered that, in most studies, only a small number of target words were selected for intervention and assessment. Students may have learned additional words that were not target words just from hearing them read without explanations in the stories. However, if such learning occurred, it was not captured on the outcome measures used in the studies.

Secondary analyses from some studies in our review also provided preliminary evidence on other issues of interest to speech-language pathologists and educators who might be considering implementation of vocabulary interventions. For example, students retained new vocabulary items for periods of at least six weeks following story readings (Brett et al., 1996; Elley, 1989, Study 2). Evidence was inconsistent, however, on whether vocabulary instruction showed differential benefits for children with high versus low levels of vocabulary abilities (Elley, 1989, Study 2; Justice et al., 2005; Penno et al., 2002). Finally, there was a hint that repeated readings of storybooks might be more beneficial for younger than for older students in enhancing vocabulary learning (Biemiller & Boote, 2006, Study 1).

Conclusions

Do elementary school students (kindergarten through grade 4) [P] learn new vocabulary [O] more effectively from hearing stories with brief

explanations of unknown words [I] than from hearing stories without such explanations [C]? Based on our evidence review, we concluded that this PICO question should be answered "Yes", albeit with some qualifications. Three key considerations influenced our conclusion: scientific control, clinical (educational) significance, and feasibility (Cleave, 2001).

Our review identified five studies, with varying degrees of scientific control (i.e., methodological quality), which specifically addressed our PICO question. Of these, two employed experimental designs (RCTs) that permit inferences about causal links between an intervention and an outcome. Three others employed quasi-experimental designs (QEDs) that do not support causal inferences, but may provide evidence to support practice if they are reasonably well designed and executed (Gersten et al., 2005). All of the studies had some important methodological limitations (no study is perfect!), but we judged this combination of two RCTs and three QEDs to be sufficient to warrant consideration of this vocabulary intervention as a promising, but not yet established, evidence-based practice. Moreover, the results of the five studies generally converged on the conclusion that hearing stories with brief explanations of target words provided a consistent benefit in vocabulary learning relative to hearing stories without such explanations. This small, but relatively consistent, body of evidence is valuable, given the paucity of research on effective vocabulary instruction for children in early elementary grades (National Reading Panel, 2000). Of course, at this time, we only have evidence that hearing stories with explanations of target words promotes vocabulary development more effectively than hearing stories without such explanations. We do not yet have evidence that this is the best possible technique for promoting vocabulary growth.

Vocabulary development plays a critical role in students' educational achievement and reading comprehension. Therefore, effective vocabulary instruction should be a topic of educational significance to speech-language pathologists and other educators. The studies we reviewed, however, were rather small in scope, which is not surprising given that this is a relatively new area of scientific inquiry. However, in order for school-based vocabulary instruction to make a meaningful impact, students with low vocabulary skills

would need to acquire hundreds of new words each year (Nagy & Anderson, 1984). Larger scale studies are needed to determine whether this vocabulary intervention can be effective in teaching sufficient numbers of words to make an important difference for these students. Although not reviewed for this brief, Biemiller and Boote (2006) attempted to increase the potency of the explanation intervention in their Study 2. The results were promising, but such efforts need to be evaluated further using methodologically sound designs that permit causal inferences. Additional studies might also assess whether or not such vocabulary interventions can ultimately result in generalization to other critical skills, such as improved reading comprehension. It is also possible that stories with explanations can be combined with other vocabulary instruction methods (e.g., computerized instruction, semantic mapping) to produce a more effective instructional package, although there is currently no direct evidence to support this suggestion.

A third consideration in favor of the vocabulary intervention reviewed here is that it is feasible for implementation in elementary schools, either as a small group or whole class intervention. Many teachers already read stories to their students on a regular basis. With some additional planning, it seems possible to introduce brief explanations of unknown words into these readings in a systematic fashion. Appendix 2 and the studies reviewed in this brief are valuable resources to help speech-language pathologists and teachers in selecting appropriate books, vocabulary items, and meaning explanations. If several educators worked together, it might be possible to develop a collection of materials that could be used regularly with elementary students of different ages. Ideally, this collection would include multiple-choice tests of vocabulary that could be used to monitor outcomes of the interventions. Although other methods of measuring vocabulary, such as story retellings or verbal definitions, might be better at tapping into the depth of students' vocabulary knowledge, multiple-choice tests would be easiest for educators to develop, administer, and score.

Speech-language pathologists and educators should keep in mind several other considerations in implementing a vocabulary intervention that involves reading with explanation of target words. One factor is careful

consideration of possible costs versus benefits. For example, it could be that such interventions would take valuable time away from other instructional activities that would be more beneficial for students. Another concern is that none of the studies we reviewed focused on vocabulary instruction for children with identified language/learning disorders, although some included students with low levels of vocabulary skills (Biemiller & Boote, 2006, Study 1; Justice et al., 2005). Thus, there is currently no direct evidence on the viability of these interventions for children with identified disorders. This should not necessarily preclude use of these interventions with such children, as there is also no clear reason to believe that the intervention would not generalize to this population (Johnson, 2006). Finally, there are indications that vocabulary learning varies as a function of factors other than the nature of the intervention, such as the books chosen (Elley, 1989), the teacher who does the reading (Biemiller & Boote, 2006), and the frequency with which unknown words are encountered in the story (Elley, 1989). These factors may also need to be considered in implementation and evaluation of vocabulary interventions.

Storybook reading with brief explanations of target words may be a promising evidence-based practice for improving vocabulary learning.

This review has highlighted that storybook reading with brief explanations of target words may be a promising evidence-based practice for improving vocabulary learning in young school-aged children. Speech-language pathologists and teachers can, therefore, add this practice to a growing repertoire of possible evidence-based techniques. Evidence, however, is only one part of the decision making process in evidence-based practice. The practitioner must also carefully integrate clinical experience and client needs and preferences into such decisions (Sackett, Rosenberg, Gray, Haynes, & Richardson, 2000).

References

Anglin, J. M. (1993). Vocabulary development: A morphological analysis. *Monographs of the Society for Research in Child Development, 10* (238), 1-166.

Biemiller, A., & Boote, C. (2006). An effective method for building meaning vocabulary in primary grades. *Journal of Educational Psychology, 98* (1), 44-62.

Biemiller, A., & Slomin, N. (2001). Estimating root vocabulary growth in normative and advantaged populations: Evidence for a common sequence of vocabulary acquisition. *Journal of Educational Psychology, 93,* 498-520.

Blok, H., Van Daalen-Kapteijns, M. M., Otter, M. E., & Overmaat, M. (2001). Using computers to learn words in the elementary grades: An evaluation framework and a review of effect studies. *Computer Assisted Language Learning, 14* (2), 99-128.

Brett, A., Rothlein, L., & Hurley, M. (1996). Vocabulary acquisition from listening to stories and explanations of target words. *The Elementary School Journal, 96* (4), 415-422.

Bryant, D. P., Goodwin, M., & Bryant, B. R. (2003). Vocabulary instruction for students with learning disabilities: A review of the research. *Learning Disability Quarterly, 26* (2), 117-128.

Carey, S. (1978). The child as word learner. In M. Halle, J. Bresnan, & G. A. Miller (Eds.), *Linguistic theory and psychological reality.* Cambridge, MA: MIT Press.

Carey, S. & Bartlett, E. J. (1978). Acquiring a single new word. *Papers and Reports on Child Language, 15,* 17-29.

Centre for Evidence Based Medicine (2001). Levels of evidence and grades of recommendation. Retrieved June 15, 2006 from *http://www.cebm.net/levels_of_evidence.asp.*

Chall, J. S., & Jacobs, V. A. (2003, Spring). The classic study on poor children's fourth grade slump. *American Educator, 44,* 14-15.

Cleave, P. L. (2001). Design issues in treatment efficacy research for child language intervention: A review of the literature. *Journal of Speech-Language Pathology and Audiology, 25,* 24-34.

Cohen, J. S. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Elley, W.B. (1989). Vocabulary acquisition from listening to stories. *Reading Research Quarterly, 24* (2), 174-187.

Fukkink, R.G., & deGlopper, K. (1998). Effects of instruction in deriving word meaning from context: A meta-analysis. *Review of Educational Research, 68* (4), 450-469.

Gersten, R., Fuchs, L. S., Compton, D., Coyne, M., Greenwood, C., & Innocenti, M. S. (2005). Quality indicators for group experimental and quasi-experimental research. *Exceptional Children, 71,* 137-149.

Hart, B., & Risley, T. R. (1995). *Meaningful differences in the everyday experiences of young American children.* Baltimore: Brookes.

Jitendra, A.K., Edwards, L.L., Sacks, G., & Jacobson, L.A. (2004). What research says about vocabulary instruction for students with learning disabilities. *Exceptional Children, 70* (3), 299-322.

Johnson, C.J. (2006). Getting started in evidence-based practice for childhood speech-language disorders. *American Journal of Speech-Language Pathology, 15,* 20-35.

Justice, L.M., Meier, J., & Walpole, S. (2005). Learning new words from storybooks: An efficacy study with at-risk kindergartners. *Language, Speech, and Hearing Services in Schools, 36,* 17-32.

McGinty, A. S., & Justice, L. M. (2006). Classroom-based versus pull-out interventions: A review of the experimental evidence. *EBP Briefs, 1* (1), 1-25.

Miller, G. A. (1991). *The science of words.* New York: Scientific American Library.

Nagy, W., & Anderson, R. (1984). The number of words in printed English. *Reading Research Quarterly, 19,* 304-330.

National Reading Panel (2000). *Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction.* Washington, DC: National Institute of Child Health and Human Development.

Penno, J.F., Wilkinson, I.A.G., & Moore, D.W. (2002). Vocabulary acquisition from teacher explanation and repeated listening to stories: Do they overcome the Matthew effect? *Journal of Educational Psychology, 94* (1), 23-33.

Sackett, D. L., Straus, S. E., Richardson, W. S., Rosenberg, W., & Haynes, R. B. (2000). *Evidence-based medicine: How to practice and teach EBM*. Edinburgh, United Kingdom: Churchill Livingstone.

Swanborn, M. S. L., & de Glopper, K. (1999). Incidental word learning while reading: A meta-analysis. *Review of Educational Research, 69,* 261-285.

Troia, G.A. (1999). Phonological awareness intervention research: A critical review of the experimental methodology. *Reading Research Quarterly, 34* (1), 28-52.

Walker, D., Greenwood, C., Hart, B., & Carta, J. (1994). Prediction of school outcomes based on early language production and socioeconomic factors. *Child Development, 65,* 606-621.

Table 1. Initial Inclusion Criteria for Survey of Vocabulary Studies.

Design
Randomized Controlled Trial (RCT)   OR   Quasi-Experimental Design (QED)   OR   Regression Discontinuity Design (RD)
Independent Variable(s)
An identifiable method of vocabulary instruction   as the treatment condition in the study
Dependent Variable(s)
An outcome assessment of vocabulary acquisition   as the dependent measure in the study
Participants
Elementary school-aged children   AND   Grade between kindergarten and grade 4
Publication
Peer-reviewed journal   AND   English language

Table 2. *Comparison of Outcomes in Vocabulary Interventions Involving Reading of Storybooks with and without Explanations of Target Words*

Grade	Study	Design	Condition	N	No. of books read	No. of words taught	Outcome Measures	Pre-test	Post-test	Gain	Effect Size & 95% CI
K	Justice, Meier & Walpole, 2005	RCT	Reading with explanation	29	10	30	Researcher designed word definition test; scored 0-2 for depth of knowledge	High[a] 4% Low[b] 5%	High[a] 13% Low[b] 12%	High[a] 9% Low[b] 7%	High vocab[a] $d = -.22$ $CI = -1.64$ to 1.45 Low vocab[b] $d = .48$, $CI = -1.62$ to 2.91
			Reading without explanation	29	10	30		High[a] 6% Low[b] 5%	High[a] 14% Low[b] 9%	High[a] 8% Low[b] 4%	
			No books	28	0	0		High[a] 3% Low[b] 4%	High[a] 8% Low[b] 7%	High[a] 5% Low[b] 3%	
K-2	Biemiller & Boote, 2006 (Study 1)	QED	Reading with explanation	112[c]	3	24	Researcher designed word definition test; scored 0-1 for depth of knowledge	23%	45%	22%	$d = .59^{d}$, $CI = -2.93$ to 3.37
			Reading without explanation	112[c]	3	24		27%	40%	12%	
1	Penno, Wilkinson & Moore, 2002	RCT	Reading with explanation	23	2	20	Researcher designed multiple-choice picture-pointing vocabulary test & use of words in story retelling	28%[e]	53%[e]	25%[e]	Multiple-choice $d = .71$, $CI = -.04$ to 1.29; Retelling $d = .94$, $CI = -1.46$ to 2.36
			Reading without explanation	24	2	30		32%[e]	37%[e]	5%[e]	

Grade	Study	Design	Condition	N	No. of books read	No. of words taught	Outcome Measures	Pre-test	Post-test	Gain	Effect Size & 95% CI
2	Elley, 1989 (Study 2)	QED	Reading with explanation	72	2	20[f]	Researcher designed multiple-choice vocabulary test	33%[f]	73%[f]	40%[f]	Book 1 $d = 1.46$[g], $CI = -6.37$ to 8.44 Book 2 Insufficient information for ES calculation
			Reading without explanation	55	2	20		40%[f]	55%[f]	15%[f]	
			No books	61	0	0		n/r[h]	n/r[h]	2%[h]	
4	Brett, Rothlein & Hurley, 1996	QED	Reading with explanation	58	2	20	Researcher designed multiple-choice vocabulary test	24%	57%	33%	Book 1 $d = 1.47$, $CI = .73$ to 1.83; Book 2 $d = 1.30$, $CI = .66$ to 1.66
			Reading without explanation	56	2	20		29%	28%	-1%	
			No books	61	0	0		30%	32%	2%	

Notes. Grade: K = kindergarten; 1 = grade 1; 2 = grade 2; 4 = grade 4; Design: RCT=Randomized Controlled Trial; QED=Quasi-Experimental Design; Pre-test: *M* performance expressed as a percentage of the maximum possible score; Post-test: *M* performance expressed as a percentage of the maximum possible score; Gain: difference between pre-test and post-test performance; Effect size & 95% CI: Standardized post-test effect size and 95% confidence interval for the outcome difference between reading with and without explanations of target words.

[a] For students with high vocabulary skills.

[b] For students with low vocabulary skills.

[c] Kindergarten *n* = 43, Grade 1 *n* = 37, Grade 2 *n* = 32.

[d] Effect size based on gain scores.

[e] Based on multiple choice test; retelling measure could not be expressed as percentage.

[f] Based on Book 1 only; results for Book 2 were not reported fully but were smaller in magnitude.

[g] Based on item rather than participant means and standard deviations

[h] n/r = not reported.

Table 3. Assessment of Methodological Quality for each Study, Based on Criteria and Weightings by Troia (1999).

Criterion[a]	Weight[b]	Study				
		Justice et al.[c,d]	Biemiller & Boote, Study 1[d]	Penno et al.[d]	Elley, Study 2[c]	Brett et al.[c]
INTERNAL VALIDITY						
1) Random Assignment	3	3	0	3	0	3
2) Control Group Intervention	3	0	0	3	3	3
3) Control Group Materials	1	0	0	1	1	1
4) Counterbalancing of Instructors	2	0	0	2	0	0
5) Treatments Explicitly Described	2	2	2	2	2	2
6) Equivalent Instructional Time	3	n/a	3	3	3	3
7) Equivalent Mortality Rates	1	1	0	0	0	0
8) Operationalized Measures	3	3	3	3	3	3
9) Reliability of Measures Reported[e]	2	2	2	2	0	2
10) Treatment Fidelity	3	3	3	0	0	0
11) Absence of Floor and Ceiling Effects	3	3	3	0	0	3
12) Sufficiently Large N	1	1	1	1	1	1
13) Type I Error Controlled[f]	2	0	0	0	0	0
14) Appropriate Unit of Analysis	2	2	0	2	0	0
15) Satisfactory Statistical Analyses	3	3	3	3	0	3
16) Effect Sizes Reported	1	1	1	0	0	0
Internal validity subtotal	35	24	21	25	13	24
EXTERNAL VALIDITY						
17) Rationale for Study	3	3	3	3	3	3
18) Participant Selection	3	3	3	3	3	3
19) Age/Grade	3	3	3	3	3	3
20) Gender	2	2	2	2	0	0

Criterion[a]	Weight[b]	Justice et al.[c,d]	Biemiller & Boote, Study 1[d]	Penno et al.[d]	Elley, Study 2[c]	Brett et al.[c]
				Study		
21) Race	2	2	0	2	0	2
22) Socioeconomic Status	3	3	3	0	0	0
23) Locale	1	1	1	1	1	1
24) IQ	3	0	0	0	0	0
25) Achievement	3	3	3	3	0	3
26) History of Special Education[g]	1	n/a	n/a	n/a	n/a	n/a
27) Disability Criteria[g]	3	n/a	n/a	n/a	n/a	n/a
28) Task or Stimulus Transfer[h]	3	0	0	3	0	0
29) Maintenance Probes	2	0	0	0	2	2
External validity subtotal	32	20	18	20	12	17
Total Score	67	44	39	45	25	41
% Applicable Criteria Met		72%	64%	70%	39%	64%

[a] For all criteria, if information is not explicitly stated, assume NO.

[b] Weighting based on importance of criterion, ranging from 1 = factor is desirable but not essential to 3 = factor is critical for drawing causal conclusions.

[c] Study included a no treatment control group.

[d] Study included a within-participants comparison of reading with and without explanations.

[e] Criterion 9. Reliability of well-known standardized tests (e.g., PPVT, EOWPVT) can be assumed.

[f] Criterion 13. Type I error must be controlled experiment-wide with a specific procedure (e.g., Bonferroni, MANOVA) mentioned explicitly to control for Type I error.

[g] Criteria 26 & 27. If general education students are participants, these categories should be scored not applicable (n/a) unless specifically reported.

[h] Criterion 28. An assessment of transfer of learning to novel tasks or novel stimuli was included.

Appendix 1

Vocabulary Intervention Studies that met Initial Inclusion Criteria but were not Relevant to the PICO Question.

Beck, I.L., Perfetti, C.A., & McKeown, M.G. (1982). Effects of long-term vocabulary instruction on lexical access and reading comprehension. *Journal of Educational Psychology, 74,* (4), 506-521.

Boling, C., Martin, S. H., & Martin, M. A. (2002). The effects of computer-assisted instruction on first grade students' vocabulary development. *Reading Improvement, 39* (2), 79-88.

Brabham, E.G., & Lynch-Brown, C. (2002). Effects of teachers' reading-aloud styles on vocabulary acquisition and comprehension of students in the early elementary grades. *Journal of Educational Psychology, 94* (3), 465-473.

Ewers, C.A., & Brownson, S.M. (1999). Kindergartners' vocabulary acquisition as a function of active vs. passive storybook reading, prior vocabulary, and working memory. *Journal of Reading Psychology, 20,* 11-20.

Juel, C., & Deffes, R. (2004). Making words stick. E*ducational Leadership, March,* 30-34.

McKeown, M. G., Beck, I. L., Omanson, R. C., & Pople, M. T. (1985). Some effects of the nature and frequency of vocabulary instruction on the knowledge and use of words. *Reading Research Quarterly, 20,* 522-535.

Morrow, L. M., & Smith, J. K. (1990). The effects of group size on interactive storybook reading. *Reading Research Quarterly, 25* (3), 213-231.

Robbins, C. & Ehri, L. C. (1994). Reading storybooks to kindergartners helps them to learn new vocabulary words. *Journal of Educational Psychology, 86,* 54-64.

Segers, E., & Verhoeven, L. (2003). Effects of vocabulary training by computer in kindergarten. *Journal of Computer Assisted Learning, 19,* 557-566.

Segers, E., Takke, L., & Verhoeven, L. (2004). Teacher-mediated versus computer-mediated storybook reading to children in native and multicultural kindergarten classrooms. *School Effectiveness and School Improvement, 15,* 215-226.

Terrell, S. L., & Daniloff, R. (1996). Children's word learning using three modes of instruction. *Perceptual and Motor Skills, 83,* 779-787.

Tomesen, M. & Aarnoutse, C. (1998). Effects of an instructional programme for deriving word meanings. *Educational Studies, 24,* 107-128.

Appendix 2

Examples of Books, Explanations, and Target Vocabulary Items for Different Grades.

Justice, L.M., Meier, J., & Walpole, S. (2005). Learning new words from storybooks: An efficacy study with at-risk kindergartners. *Language, Speech, and Hearing Services in Schools, 36,* 17-32.	
Grade	Kindergarten
Books	*Book! Book! Book!* (Bruss, 2001) *The Caterpillar that Roared* (Lawrence, 2000) *Harry and the Terrible Whatzit* (Gackenbach, 1977) *Imogene's Antlers* (Small, 1985) *Otis* (Bynum, 2000) *Possum and the Peeper* (Hunter, 1998) *Shy Charles* (Wells, 1988) *Swimmy* (Lionni, 1963) *The Bear Under the Stairs* (Cooper, 1993) *What do You do with a Kangaroo?* (Mayer, 1973)
Explanation	"A *marsh* is a very wet place where there are wetlands covered with grasses. Like, 'We took a boat through the marsh and we saw lots of birds and alligators.'"
Target Words	Heaved Gulp Furnace Stale
Biemiller, A., & Boote, C. (2006). An effective method for building meaning vocabulary in primary grades. *Journal of Educational Psychology, 98,* 44-62.	
Grade	Kindergarten, Grade 1, Grade 2

Books	Kindergarten: *There is More, Much More* (Alexander, 1987)
	            Mud (Ray, 1996)
	            Something From Nothing (Gilman, 1992)
	Grade 1: *Going Down the Road* (Schertle, 1995)
	        The Seashore Book (Zolotow, 1992)
	        The Tree that Grew to the Moon (Fernandes, 1994)
	Grade 2: *Drac and the Gremlin* (Baillie & Tanner, 1991)
	        Alexander and the Wind-Up Mouse (Lionni, 1969)
	        Brenda and Edward (Kovalski, 1984)

Explanation	"It seemed like a good *solution*. A *solution* is the answer to a problem."
Target Words	Kindergarten: circus, cackle, holler Grade 1: obstacle, actually, dwindle Grade 2: saucer, quiver, barely

Penno, J.F., Wilkinson, I.A.G., & Moore, D.W. (2002). Vocabulary acquisition from teacher explanation and repeated listening to stories: Do they overcome the Matthew effect? *Journal of Educational Psychology, 94,* 23-33.	
Grade	Grade 1
Books	*Anak the Brave* (Ling, 1990) *No Place Like Home* (Elliott, 1990)
Explanation	"A *dugout* is another name for a canoe that Anak would get in and paddle down the river."
Target Words	Dugout Bunyip Verandah
Elley, W.B. (1989). Vocabulary acquisition from listening to stories. *Reading Research Quarterly, 24,* 174-187.	
Grade	Grade 2
Books	*Rapscallion Jones* (Marshall, 1983) *The White Crane* (Smith, 1983)
Explanation	The teacher might explain *pressing engagements* by giving the synonym *important things to do*.
Target Words	Redistribute Summoned Dingy
Brett, A., Rothlein, L., & Hurley, M. (1996). Vocabulary acquisition from listening to stories and explanations of target words. *The Elementary School Journal, 96,* 415-422.	
Grade	Grade 4
Books	*Bunnicula* (Howe & Howe, 1979) *The Reluctant Dragon* (Grahame, 1953)
Explanation	"*Despondency* is a feeling of being discouraged or hopeless. As the boy went back to the village, he felt hopeless."
Target Words	Indulgent Disdain Reverie

Evidence-Based Vocabulary Instruction for Elementary Students via Storybook Reading

Carla J. Johnson and Erin Yeates
University of Toronto