

8-8-2018

The Effect of E-book Vocabulary Instruction on Spanish-English Speaking Children

Carla Wood
Florida State University

Lisa Fitton
University of South Carolina - Columbia, fittonl@mailbox.sc.edu

Yaacov Petscher
Florida State University

Estrella Rodriguez
Florida State University

Gretchen Sunderman
Florida State University

See next page for additional authors

Follow this and additional works at: https://scholarcommons.sc.edu/sph_communication_sciences_disorders_facpub



Part of the [Communication Sciences and Disorders Commons](#)

Publication Info

Published in *Journal of Speech, Language, and Hearing Research : JSLHR*, Volume 61, Issue 8, 2018, pages 1945-1969.

This Article is brought to you by the Communication Sciences and Disorders at Scholar Commons. It has been accepted for inclusion in Faculty Publications by an authorized administrator of Scholar Commons. For more information, please contact digres@mailbox.sc.edu.

Author(s)

Carla Wood, Lisa Fitton, Yaacov Petscher, Estrella Rodriguez, Gretchen Sunderman, and Taehyeong Lim

Research Article

The Effect of e-Book Vocabulary Instruction on Spanish–English Speaking Children

Carla Wood,^a Lisa Fitton,^a Yaacov Petscher,^a Estrella Rodriguez,^a
Gretchen Sunderman,^a and Taehyeong Lim^a

Purpose: This study aimed to examine the effect of an intensive vocabulary intervention embedded in e-books on the vocabulary skills of young Spanish–English speaking English learners (ELs) from low–socioeconomic status backgrounds.

Method: Children ($N = 288$) in kindergarten and 1st grade were randomly assigned to treatment and read-only conditions. All children received e-book readings approximately 3 times a week for 10–20 weeks using the same books. Children in the treatment condition received e-books supplemented with vocabulary instruction that included scaffolding through explanations in Spanish, repetition in English, checks for understanding, and highlighted morphology.

Results: There was a main effect of the intervention on expressive labeling ($g = 0.38$) and vocabulary on the Peabody Picture Vocabulary Test–Fourth Edition ($g = 0.14$; Dunn & Dunn, 2007), with no significant moderation effect of initial Peabody Picture Vocabulary Test score. There was no significant difference between conditions on children’s expressive definitions.

Conclusion: Findings substantiate the effectiveness of computer-implemented embedded vocabulary intervention for increasing ELs’ vocabulary knowledge.

Implications: Computer-assisted vocabulary instruction with scaffolding through Spanish explanations, repetitions, and highlighted morphology is a promising approach to facilitate word learning for ELs in kindergarten and 1st grade.

English learners (ELs) are a top education priority in the United States. The population characteristics of public schools have shifted over the last decade, as the enrollment of Hispanic students in PreK–12th grade increased from 8.6 to 12.1 million (24%) in 2012 (Kena et al., 2016). ELs comprised 9.3% of all public school students, with higher prevalence in urban educational settings (Kena et al., 2016). In the changing landscape of schools, Spanish is the most common language spoken by EL students in the United States (National Center for Education Statistics, 2016). Given population growth, schools face growing demand for instructional supports for ELs who may present with additional risks for poor academic achievement compared with their monolingual English-speaking peers (National Center for Education Statistics, 2011). Evidence of the risk for low achievement is demonstrated by the fact that foreign-born Hispanics have the highest school dropout rate (24%; Kena et al., 2015).

Multiple factors contribute to ELs’ high risk for reduced achievement. Importantly, a disproportionately large number of ELs live in poverty in the United States. Over 60% of school-age ELs, including 32% of Hispanic children, are reported to be from low–socioeconomic status backgrounds (Capps et al., 2005), a well-documented risk factor for diminished literacy and academic achievement (e.g., Kena et al., 2015). In addition, many parents of ELs have limited English proficiency (Capps et al., 2005), resulting in more restricted opportunities for their families to promote English language and literacy development. Given the well-established link between familial language support and child academic outcomes (Lonigan, 2015; Sénéchal & LeFevre, 2002), children from families with limited English proficiency are less likely to develop English as quickly as their monolingual English-speaking peers. This perceived risk has been corroborated by evidence from a recent Florida-based study, in which ELs’ average predicted score on an English receptive vocabulary test in kindergarten was nearly 2 *SDs* below the mean for monolingual English-speaking peers (Wood Jackson, Schatschneider, & Leacox, 2014). Early differences in English proficiency are not quickly remediated with current educational practices. ELs who begin kindergarten with limited English proficiency have

^aFlorida State University, Tallahassee

Correspondence to Carla Wood: carla.wood@cci.fsu.edu

Editor-in-Chief: Sean Redmond

Editor: Megan Dunn Davison

Received September 26, 2017

Revision received December 13, 2017

Accepted January 30, 2018

https://doi.org/10.1044/2018_JSLHR-L-17-0368

Disclosure: The authors have declared that no competing interests existed at the time of publication.

been observed to experience an achievement gap that persists into fifth grade (Kieffer, 2008).

Importance of Vocabulary

To improve academic achievement in ELs, vocabulary knowledge has an important role, particularly in supporting reading comprehension (August, Carlo, Dressler, & Snow, 2005; August & Shanahan, 2006; Baker et al., 2014; Scarborough, 2002). Vocabulary is fundamental to reading comprehension in both monolingual and EL students and serves as a link between oral and written language (Proctor, August, Carlo, & Snow, 2006; Scarborough, 2002). Specifically for Spanish-speaking ELs, English vocabulary knowledge is both a strong proximal and a strong distal predictor of reading comprehension (Proctor et al., 2006). Building English vocabulary knowledge has been specifically highlighted as a key factor for improving comprehension and metalinguistic skills among ELs (August & Shanahan, 2006; Baker et al., 2014).

ELs commonly show restricted depth of lexical knowledge in the second language (L2; Proctor et al., 2006) and have particular difficulty acquiring labels for words whose phonological representations differ from the sound patterns of their native language (Ordonez, Carlo, Snow, & McLaughlin, 2002). Deducing the meaning of novel words from incidental exposures in the context of typical reading exposures is particularly difficult for ELs. Novel words are difficult for ELs who have limited grammatical knowledge to leverage when inferring word meanings from the sentential context alone (Carlo et al., 2004). As a result, an expanding literature base substantiates the need for supplemental instruction for children who are at risk for poor reading and academic achievement (Justice, 2006; Justice, Meier, & Walpole, 2005).

Effective Vocabulary Instruction for Children

Previous research findings suggest that there are several active ingredients needed to enhance vocabulary acquisition for the school-age population at large. Providing expansions, or clear definitions and explanations of word meanings (Paul, 2007), has been shown to be an effective practice for increasing vocabulary knowledge (Apthorp et al., 2012; Dalton, Proctor, Uccelli, Mo, & Snow, 2011). Embedding these detailed expansions and scaffolding of word understanding in meaningful, authentic literacy contexts with repeated exposures has been associated with improved language and literacy outcomes (Beck & McKeown, 2007; Justice, 2006; Justice et al., 2005; Roberts & Neal, 2004). In particular, shared reading experiences containing elements of rich vocabulary instruction including expansions and definitions have been linked to accelerated vocabulary learning in monolinguals and ELs alike (Justice et al., 2005; Lugo-Neris, Wood Jackson, & Goldstein, 2010). In a seminal study by Beck and McKeown (2007), monolingual students in kindergarten and first grade received intense instruction including word contextualization, definitions, repetition of

targeted words, examples, and verbal choices of word meanings. Children who received the rich vocabulary instruction ($n = 52$) demonstrated significantly higher increases in word knowledge than children in the comparison group ($n = 46$). Children in the treatment classes gained an average of 5.58 words compared with an average of 1.04 words in the control group.

Embedding rich, explicit vocabulary instruction with scaffolding in the context of repeated shared storybook reading has consistently been shown to accelerate word learning for young children (Koskinen et al., 2000; Roberts & Neal, 2004). Much of the literature addressing shared storybook reading has focused primarily on monolinguals (Justice, 2006; Justice et al., 2005; Whitehurst et al., 1988). However, there is sufficient evidence to expect that shared storybook reading is also an effective and engaging approach to introduce novel vocabulary and develop emergent literacy knowledge in young ELs (Cena et al., 2013). Findings in the literature suggest that shared storybook reading facilitates learning when the context is meaningful, interesting, and motivating to young children (Honig, Diamond, Gutlohn, & Cole, 2008). In a study by Roberts and Neal (2004), young ELs who participated in instruction emphasizing interactive storybook reading, vocabulary instruction, and comprehension activities outperformed children in the comparison condition who did not receive storybook reading on vocabulary tasks. Furthermore, there is a well-established benefit of repeated readings for improving vocabulary comprehension (Koskinen et al., 2000), suggesting that shared reading can have a stronger impact with high dosage.

Effective Vocabulary Instruction for ELs

In addition to repeated readings, additional strategies specific for instructing ELs have been recommended as outlined in a systematic review (August, McCardle, & Shanahan, 2014). These recommendations include providing definitions or explanations for a focused, small number of words (Apthorp et al., 2012; Dalton et al., 2011); repeating concepts and scaffolding through linking new terminology to existing knowledge and/or first language (L1) skills (Burchinal, Field, Lopez, Howes, & Pianta, 2012; Duran & Dale, 2014); previewing target words (Liang, Peterson, & Graves, 2005); providing visual aids and cues (Roberts & Neal, 2004; R. Silverman & Hines, 2009); leveraging morphosyntactic knowledge (Goodwin & Ahn, 2013; R. D. Silverman et al., 2013); and checking for understanding frequently (August et al., 2014; August & Shanahan, 2006). The integration of these strategies is hypothesized to maximize EL word learning and vocabulary growth.

Of the strategies recommended, leveraging morphological awareness has been integrated in language interventions with young monolingual children (e.g., Apel & Werfel, 2014) and has been shown to predict vocabulary growth at Grade 2 (Sparks & Deacon, 2015). Recently, there has been attention to the role of morphological knowledge in facilitating vocabulary acquisition in ELs (Lo, Anderson, & Bunch-Crump, 2016; Nagy, Carlisle,

& Goodwin, 2014; R. D. Silverman et al., 2013). A meta-analysis that examined the impact of morphological instruction on school-age children's literacy achievement revealed that morphological instruction embedded in literacy contexts was positively associated with change in vocabulary for ELs (Goodwin & Ahn, 2010).

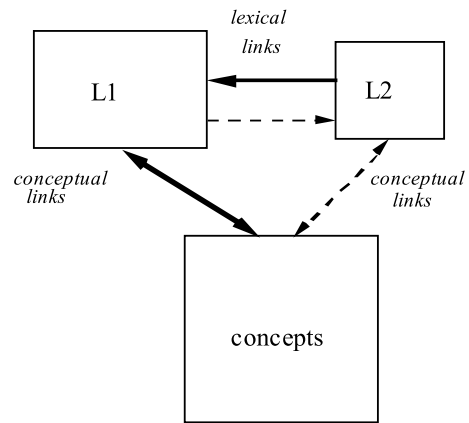
Bridging to the L1 is another promising strategy specific to supporting vocabulary acquisition in ELs. Bridging is a technique that leverages the theorized interconnectedness of ELs' two languages by explicitly connecting novel words presented in children's L2 (i.e., English) to their conceptual background knowledge in L1 (i.e., Spanish). Bilingual approaches to vocabulary instruction have yielded gains in ELs' English vocabulary that are at least as substantial as those obtained from English-only approaches (Gutiérrez-Clellen, Simon-Cerejido, & Sweet, 2012). Furthermore, multiple studies have demonstrated that incorporation of the L1 can simultaneously support L1 and English growth, resulting in a greater net increase in vocabulary than English-only instruction (Barnett, Yarosz, Thomas, Jung, & Blanco, 2007; Durán, Roseth, & Hoffman, 2010; Ebert, Kohnert, Pham, Disher, & Payesteh, 2014). This pattern has been observed among both typically developing ELs (e.g., Farver, Lonigan, & Eppe, 2009) and ELs with language impairment (e.g., Restrepo, Morgan, & Thompson, 2013). Incorporating L1 has been suggested to provide comprehensible input, promote self-esteem, and enhance academic achievement because of transferability of some language and literacy skills (Goodrich, Lonigan, Kleuver, & Farver, 2015; Gutierrez-Clellen, 1999; Kohnert, Yim, Nett, Kan, & Duran, 2005; Uccelli & Pérez, 2007). In a pilot study of Spanish-to-English bridging, a significant main effect was found for vocabulary growth given shared reading with expansions. Additional benefits of Spanish instruction were observed for improving children's ability to define words (Hedges's g , effect size = 0.22, $p = .026$), with Spanish instruction producing greater gains in definitions of targeted words (Lugo-Neris et al., 2010).

Theoretical Foundation

The use of bridging to accelerate word learning in L2 is theoretically supported by the Revised Hierarchical Model (RHM; Kroll & Stewart, 1994), a psycholinguistic model of the lexicon that emphasizes the interconnectedness between languages and the role of L1 during L2 learning. The RHM has been influential in providing scholars a better understanding of the underlying processes involved in the development of the bilingual lexicon (see Tokowicz, 2015). One of the critical assumptions of the model (and of our study) is that L2 learners are initially dependent on the L1 to access their conceptual store (refer to Figure 1).

The RHM depicts two separate and differentially sized lexicons for L1 and L2 words and one common conceptual store. The arrows in the model represent the lexical and conceptual links assumed to be active in bilingual memory. The relative strength of the links is indicated by the thickness of the arrows and is dependent on language proficiency. For example, for a Spanish-dominant EL, the associations

Figure 1. The Revised Hierarchical Model. L1 = first language; L2 = second language. Adapted from "Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations," by J. F. Kroll and E. Stewart, 1994, *Journal of Memory and Language*, 33, pp. 149–174. Copyright © 1994 by Elsevier.



between Spanish words and concepts will be very strong, whereas the associations between English words and concepts will be weaker. To access the conceptual store, the learner will use the L1 translation equivalent, because the L1 has privileged access to the conceptual store. In this sense, the L1 serves as a tool to enable initial conceptual processing and is a tool we leverage in the current study.

As depicted in the model, as the learners become more proficient in the L2, they will eventually begin to strengthen the link between the L2 words and concept, a process known as conceptual mediation, and a clear goal of L2 vocabulary learning. Conceptual mediation is a marker of proficient bilingual performance (e.g., La Heij, Kerling, & Van der Velden, 1996), but research suggests that conceptual mediation develops in stages, with the first stage being lexical mediation via the translation equivalent (e.g., Sunderman & Kroll, 2006). Therefore, this project takes the theoretical predictions of the RHM and applies them to vocabulary teaching methods by using bridging to the L1. Moreover, given the evidence indicating the L1 is always active and competing in bilingual memory (e.g., Van Heuven, Dijkstra, & Grainger, 1998), leveraging L1 knowledge during L2 instruction is a theoretically sound option.

Increased Access to Technology-Assisted Learning Opportunities

The use of electronic technology for delivery of book instruction was motivated by three main factors. Of primary importance, e-book delivery allows for the integration of word explanations in Spanish through prerecorded audio files in e-books that could be implemented by monolingual English-speaking educational personnel. A second potential advantage to e-books is that the digital format allows for inclusion of video segments to depict the functional or relational meaning of target words (e.g., *The man uses a ladder*

to climb up). Visual images through video and/or animation have been shown to support children's recall of word labels and relationships between words and their references (Verhallen & Bus, 2012). In addition, e-books can be presented using an interactive, game-like format, which may improve children's motivation to engage actively with the books.

The practicality of integrating L1 support with the explicit intensive vocabulary instruction strategies recommended may be increased using computer-assisted or technology-enhanced programs (White & Gillard, 2011). Feasibility and effective components of utilizing technology to implement vocabulary instruction are supported by several studies (e.g., Levy, 2009). In one study, ELs were provided tablets with electronic dictionaries to assist them in locating desired words quickly and hear a pronunciation model (Demski, 2011). In another study, Verhallen and Bus (2012) found that children learned more expressive labels from e-books that included moving images than e-books with static pictures alone. Specifically, 92 ELs from low-income immigrant backgrounds were randomly assigned to a comparison group or two treatment groups: repeated exposures to e-book readings with static pictures or video. Both treatment groups made significant gains in receptive and expressive vocabulary, but significantly more labels were retained in the video-enhanced e-book group.

Few, if any, e-book studies, however, have utilized L1 bridging supports and adaptive levels of responses based on the child's performance. It is posited that embedding both L1 bridging and rich instruction into e-book recorded presentations would result in larger effect sizes; however, more research is needed with ELs to extend findings of small-sample pilot projects (Leacox & Wood Jackson, 2014). Additional research is needed to rigorously evaluate this approach's effectiveness for improving ELs' vocabulary skills, assess potential moderators of ELs' outcomes, and examine the feasibility of implementing technology-enhanced intensive vocabulary instruction in English-based classrooms.

The primary aim of the current study was to examine the English vocabulary growth of kindergarten and first-grade EL students who participated in an intensive, L1-enhanced vocabulary intervention delivered via e-book three times a week. Secondarily, we explore one moderator, initial English vocabulary skills. The specific research questions included the following:

1. Is English vocabulary learning enhanced when an e-book vocabulary intervention that includes bridging to Spanish is included in the curriculum?
2. Are the effects of e-book vocabulary intervention moderated by students' English skills, such as initial vocabulary skills?
3. Is an e-book vocabulary intervention feasible for monolingual teachers to incorporate into the daily schedule of the classroom setting?

The question of feasibility involved two aspects of interest to the investigators: (a) if it was feasible to

administer technology-enhanced intervention in small groups within or outside the classroom for a 15- to 20-min dosage across 20 weeks in low-socioeconomic status schools, and (b) if an intervention that involved Spanish expansions could be administered by monolingual educational support staff.

Method

The current study was part of Project BLOOM (Bridging for Language Outcomes in the Classroom), funded through a grant from the Institute of Education Sciences, U.S. Department of Education, focusing on language and literacy interventions for bilingual elementary children. The study was reviewed and approved by the university's institutional review board for research involving human subjects (HSC 2016.18265). Children who returned consent were enrolled in the project to receive e-book sessions three times a week for 20 weeks. Half of the children were randomly assigned to the e-book read-only condition, and the other half received intensive vocabulary strategies embedded within the same e-books.

Students were randomly assigned to conditions within grade and schools. Because students were administered multiple assessments that measured their vocabulary, it was of interest to conceptualize vocabulary as a function of two theoretically meaningful components (Gross, Buac, & Kaushanskaya, 2014) during random assignment to condition. As such, a principal component analysis (PCA) was used to estimate a component score that leveraged the commonality between children's scores on a standardized measure of receptive vocabulary in English (the Peabody Picture Vocabulary Test–Fourth Edition [PPVT-4]; Dunn & Dunn, 2007) and an experimenter-created probe of expressive English vocabulary based on the targets of the intervention program. These assessments were selected for use in the PCA because both are measures of the construct targeted by the intervention program, English vocabulary. One, the PPVT-4, is distal relative to the intervention, and the other, the probe, is proximal to the intervention. A component was used rather than an estimated factor score as the former is useful for explaining variance across measured variables in the defined matrix and the latter requires evaluation of score determinacy. The underlying latent structure of expressive and receptive vocabulary was not germane to this project goal; thus, the PCA served as a useful outcome that produced an overall component score that maximized the individual differences between measures. The PCA extracted one component that accounted for 60% of the variance in the matrix. Students were stratified within each grade and school by this component score and then randomly assigned to conditions.

The study was conducted in public schools in Florida and Kansas City. ELs attended regular English-based education classroom settings. Although the partnering schools had a large percentage of ELs, participating classrooms did not exclusively serve ELs and did not necessarily have

a high proportion of ELs. Participating classrooms included 21 kindergarten classrooms and 25 first-grade classrooms. Three of the eight schools were considered to be in rural communities and served a high proportion of children of families from migrant farm-working backgrounds. Two of the schools were considered semirural, located on the outskirts of an urban area. The remaining three schools were in large urban areas. Additional descriptive information on the participating schools is provided in Table 1.

Participants

Children were enrolled in eight participating elementary schools located in Florida (196 children) and Kansas City, Kansas (95 children). Eligible participants were recruited through teachers in the participating classrooms who invited children in kindergarten and first grade whose parents had indicated that they spoke Spanish at home. The participant pool initially included 291 returned consents; however, there was attrition of 22 participants (7.7%) due to family relocation. An equal number of participants from kindergarten ($n = 11$) and Grade 1 dropped out of the study. Of these 22 total children, six had been assigned to the control group and 16 had been assigned to the treatment group, a differential attrition rate of 5.7%. The attrition rate was considered low according to the What Works Clearinghouse (2014) guidelines.

The final participant pool at the end of the school year included 147 boys (51.2% of the sample) and 140 girls (48.8%), with an average age of 72 months ($SD = 9.8$ months). On the basis of parents' report, 91.6% of the participants were eligible for free lunch and 3.4% were eligible for reduced lunch. None of the participants had any identified sensory impairments or other identified disorders based on parent and teacher reports. Children were screened for intellectual disabilities using the Primary Test of Nonverbal Intelligence (PTONI; Ehrler & McGhee, 2008). The mean standard score on the PTONI for the participants was considered to be within average range, with a mean of 92.38 ($SD = 19.75$). The variability in performance on the PTONI was higher than would have been expected for a sample of monolingual

children; however, the PTONI is not normed specifically on ELs. It is possible that ELs entering kindergarten did not have prior experience with the task demands of the PTONI.

Family Characteristics

After consent forms were returned, heritage Spanish speakers who were graduate students or advanced undergraduates majoring in the School of Communication Science and Disorders called parents to gather family information including their schooling level and percentage of Spanish and English use in the household. Investigators followed a script for the phone interview and obtained complete demographic data from 237 families, 86% of all participants. Missing data were largely attributable to disconnected phones or family's decline of interview. Most of the interviewed mothers were homemakers, and the most common occupation of the fathers was construction. Less than half of caregivers reported graduating from high school, and less than 10% attended any college. Additional descriptive information on participants and their families is provided in Table 2.

Of the total participant pool, 173 families of the children (63%) reported that Mexico was their country of origin, and 38 participating families reported that they were from Guatemala (18%). Other Spanish-speaking countries represented in the participating sample included Cuba with 18 families (8.2%), El Salvador with 17 families (7.6%), Honduras with nine families (4.2%), and Columbia with four families (2%). Countries/territories with fewer than four participants were the Dominican Republic, Nicaragua, Peru, Puerto Rico, and Venezuela, yielding a varied representation of Spanish-speaking dialects from Central and South America and the Caribbean region.

Instructional Materials

Book Selection

The BLOOM e-book sessions were supplemental to the regular classroom curriculum. The investigators selected

Table 1. Demographics of participating schools.

School	<i>N</i>	<i>n</i> classes	School population ^a	% F/R lunch	% Hispanic in school	% F/R lunch in district	% Hispanic in district	% EL in district	White–Hispanic gap ^b
A	35	5	305	82	52	83.1	19.2	6.2	0.229 ($SE = 0.161$)
B	36	2	145	99	93	39.1	25.9	12.2	0.708 ($SE = 0.036$)
C	27	7	1,042	86	63	51.9	31.4	11.3	0.614 ($SE = 0.009$)
D	28	5	358	82	52	31.1	19.0	8.3	0.774 ($SE = 0.014$)
E	67	6	1,250	78	63	83.0	45.3	28.8	0.255 ($SE = 0.016$)
F	55	15	907	82	52	51.9	31.4	11.3	0.614 ($SE = 0.009$)
G	34	5	944	94	53	65.9	67.6	19.6	0.516 ($SE = 0.009$)
H	6	1	258	77	25	83.1	19.2	6.2	0.229 ($SE = 0.161$)

Note. EL = English learner; F/R = free/reduced. District level data were retrieved from the Stanford Education Data Archive (Reardon, Kalogrides, & Shores, 2016).

^aSchool population represents the total number of children enrolled in the elementary school. ^bWhite–Hispanic gap reflects the achievement gap between average performance on state English Language Arts assessment for Hispanic students compared with White students in the district, which is reported in standard deviations (see Reardon et al., 2016).

Table 2. Demographics and characteristics of child participants.

Characteristic	<i>n</i>	%
Child gender		
Male	131	47.6
Female	144	52.3
Lunch status		
Free	269	97.8
Reduced	6	2.1
Child place of residence		
Florida	179	65.0
Kansas	96	34.0
Parent education (mothers and fathers)	237	86.0
Less than high school	165	35.0
High school diploma	183	39.0
College	26	5.5
	<i>M</i>	<i>SD</i>
Percentage of Spanish use	55.1	11.7
Percentage of English use	8.0	4.4
Both ^a	36.3	18.5
Child age (months)	72	9.8

Note. Demographic information was gathered through phone interviews with parents.

^aPercentage reporting use of both Spanish and English equally.

24 books per grade to ensure that there were sufficient materials for one book per week (administered three times per week) for at least 20 weeks. The same books were used for e-book sessions in both treatment and comparison conditions. For book selection, the authors considered children's books listed as supplemental for the classroom reading curriculum (Houghton Mifflin Harcourt, 2015) and employed Hargrave and Sénéchal's (2000) criteria for selection including that the books (a) were of high interest, (b) contained novel vocabulary, (c) contained multiple occurrences of the vocabulary, (d) had illustrations, and (e) were not excessively long. The final selection of 48 books (24 per grade) were between 12 and 32 pages in length, with an average of 21 pages per book ($SD = 5.02$ pages).

Word Selection

The authors selected four English words to target per book, consistent with the recommendation to teach a small set of vocabulary words intensively across several days (Baker et al., 2014). We included words that occurred multiple times in the books and that were likely to occur in the participants' environments. Nouns were selected for the purposes of the study because of the ease of illustration for measurement of word learning. On the basis of the child word frequency calculator (Bååth, 2010), the kindergarten word list had an average frequency of 33.81 occurrences per million ($SD = 5.90$ occurrences per million, ranging from 0 to 112 occurrences per million) in the 72- to 95-month corpus. First-grade words had an overall word frequency of 33.35 occurrences per million ($SD = 5.81$ occurrences per million, ranging from 0 to 265 occurrences per million). Word frequency is computed based on a corpus of over 3.5 million words included in a database of child language

(Bååth, 2010). Values above 20 occurrences per million are considered indicative of high-frequency words (e.g., Brysbaert & New, 2009). Consequently, it can be concluded that both lists included primarily high-frequency words (Appendix A).

Treatment Condition

All recorded readings of the written story occurred in English. The e-book display visually highlighted the English text of the book as the text was read aloud. In the treatment condition, the e-book session contained explicit instruction in Spanish and English on targeted words as they occurred in the book with consistent components: (a) preview of target words in English and Spanish; (b) expansion with bridging provided in Spanish and English; (c) visual representation of the target word's function; (d) a word map with tiered support, including visuals, a nonexample, and a morphology highlight; and (e) a final review. Through these instructional components, participants received repeated exposure to targeted words with intensive instruction on each word at least three times in the story. Treatment sessions were approximately 25 min in length.

At the beginning of each e-book, a preview of the four target words was provided in English and Spanish. The preview included pictures of the target words with recorded audio labels in English (Appendix D), followed by presentation of the translated label in Spanish. For the first occurrence of each target word within the story, the recorded audio presented an instruction directing the child to click on the image that represented the target word on the storybook page. To scaffold child progression through this task, the cursor image would change when the child moved the cursor to hover over the correct image. Upon the child clicking the target word image, the image animated and an audio expansion for the target word played first in English and then in Spanish. To incorporate bridging, the target word and explanation were presented in both English and Spanish (Burchinal et al., 2012; Duran & Dale, 2014). The second occurrence of each target word was also followed by a directive audio instruction for the child to click on the target word image on the book page. When the child selected the appropriate image, a video clip demonstrating the function of the target word played, and the book page then advanced. Each of the four target words was presented with the two instructional exposures in the book, and then the child continued to an interactive word map including tiered support. All directions in the intervention were presented in English first and then in Spanish.

The three-part word map was constructed as a check for understanding and as an opportunity for additional scaffolding and support (Appendix E). Each word map included three blank spaces in which the child accrued images associated with the target word in a game-like fashion. For the first space, the child was shown a four-picture array and presented with recorded instructions to select the image of the target word. If the child selected the correct image, it was added to the word map. If the child selected an alternative image, four representative images of the target word

appeared on the screen along with a more in-depth audio expansion, presented in English and then in Spanish. The second component included a two-picture array, from which the child was asked to select a nonexample, the picture that did not show the target word (e.g., “click on the picture that is NOT a bridge”), to highlight a lexical contrast. A correct response from the child resulted in another picture being added to the word map, whereas an incorrect response resulted in feedback. The final word map task highlighted morphology by prompting the child to add a prefix or suffix (e.g., “Here is one *bridge*. Find two *bridges*”). If the target word was *button*, the morphological task may require the child to find *unbutton* or *buttoning*. The final page of the e-book contained a review with images and audio-recorded labels of the word targets.

Comparison Condition

For the comparison condition, the children listened to the same recorded e-books three times a week in English but without any embedded instruction, directions, or additional language content other than a recording of the text as it appeared on the page. This design allows for the examination of the added impact of the enhanced instructional components embedded in the e-books, controlling for the impact of repeated reading alone. The comparison condition sessions were approximately 10–15 min in length.

Baseline Equivalence on Background Characteristics

Examination of background characteristics between the treatment and comparison conditions after randomization revealed similar language backgrounds between the groups. No differences were noted in reported mother educational level, $\chi^2(18, N = 211) = 19.15, p = .383$, or in father educational level, $\chi^2(18, N = 192) = 14.52, p = .695$. Similarly, no differences were observed in the reported Spanish/English use in the home, $t(212) = -0.93, p = .352$, or in the language dominance of the child, $\chi^2(21, N = 221) = 9.05, p = .989$.

Descriptive Measures

Research assistants administered standardized assessments of language, literacy, and nonverbal intelligence in September as descriptive measures of the participants' baseline skills. Overall means and standard deviations of the standardized assessments are provided in Table 3.

Spanish Receptive Vocabulary

The Test de Vocabulario en Imágenes Peabody (TVIP; Dunn, Lugo, Padilla, & Dunn, 1986) was administered in the fall as a descriptive measure of the participants' receptive Spanish vocabulary skills at the beginning of the study. The TVIP is a norm-referenced measure designed for ages 2;6–17;11 years;months. It has a mean reliability of .93. The TVIP requires the child to identify a picture representation of a word from a choice of four. The TVIP was normed on 2,707 monolingual Spanish-speaking children with a norm reference group from Mexico.

Sentence Repetition

The sentence repetition task from the morphosyntax subtest of the Bilingual English–Spanish Assessment (Peña, Gutiérrez-Clellen, Iglesias, Goldstein, & Bedore, 2014) was administered in the fall in English and Spanish. The Bilingual English–Spanish Assessment is designed for Spanish–English bilingual children 4–7 years old. For the sentence repetition task, the examiner asks the child to repeat sentences verbatim. Each sentence is presented individually, 7–14 words in length, with specific grammatical structures scored for accuracy. Children were administered the Spanish and English versions of the subtest separately.

Early Literacy Performance

The Woodcock Reading Mastery Tests–Third Edition (Woodcock, 2011) Letter Identification (LI), Phonological Awareness (PA), and Rapid Automatic Naming (RAN) subtests were administered in fall of the school year as descriptive measures. The test was not administered at all schools because of restrictions on assessment time. The Woodcock Reading Mastery Tests–Third Edition is a set of tests for measuring oral language and academic achievement normed on individuals 4–79 years old. The test was evaluated on a normative sample of 3,360 individuals (including 2,600 school-age participants) in 45 states in the United States. Split-half reliability for each subtest on Form A for kindergarten is as follows: .91 for LI, .92 for PA, and .83 for RAN. For first grade, split-half reliability for each subtest on Form A is as follows: .69 for LI and .91 for PA. Notably, the RAN subtest included RAN for digits, letters, pictures, and colors, consistent with other assessments of RAN (Denckla & Rudel, 1974). Following the standard procedures for administering and scoring, the pair of RAN subtests with the highest performance was used in calculating the standard score.

Nonverbal Intelligence

The PTONI (Ehrler & McGhee, 2008) was included as a baseline descriptive measure of participants' nonverbal cognitive abilities. This test uses pictures to assess reasoning ability in young children without requiring a verbal response. The authors report internal consistency reliability with a coefficient alpha of .93, test–retest reliability of .97, and interrater reliability of .99.

Experimental Measures

English Receptive Vocabulary

The PPVT-4 (Dunn & Dunn, 2007) was administered in the fall and spring. The PPVT-4 is an untimed, norm-referenced measure of receptive vocabulary normed through a sample of 3,540 participants for use with individuals 2–90 years old. The examiner presents a word and asks the child to point to the picture that best represents the word from a four-picture array. Split-half reliability by age for Forms A and B has a mean of .94 ($SD = 3.6$) and ranges from .90 to .97 for ages 5–11 years, based on normative data on monolingual English-speaking children.

Table 3. Descriptive data for standardized measures by group.

Assessment	Read 2 (Treatment; N = 155)			Read 1 (Control; N = 133)		
	M	SD	n	M	SD	n
TVIP			112			114
Standard score	87.24	18.19		80.66	17.26	
Raw score	28.63	19.73		26.48	19.67	
WRMT-III RAN			104			98
Standard score	96.81	12.56		96.27	13.06	
Raw score	19.89	7.10		19.71	7.19	
WRMT-III letter ID			76			54
Standard score	90.41	16.03		96.20	14.39	
Raw score	13.25	5.39		14.20	5.17	
WRMT-III PA			51			38
Standard score	96.80	16.72		91.47	20.30	
Raw score	21.20	6.72		19.26	7.35	
BESA sentence repetition			122			127
English percent correct	69.70	25.30		72.72	24.40	
English raw score	23.00	8.35		24.00	8.07	
BESA sentence repetition			148			128
Spanish percent correct	59.97	26.80		50.72	27.60	
Spanish raw score	22.19	9.91		18.77	10.19	
PPVT-4 pretest			151			127
Standard score	81.20	15.80		81.87	14.06	
Raw score	72.28	27.42		72.81	25.35	
PPVT-4 posttest			137			121
Standard score	88.72	13.50		86.85	11.46	
Raw score	92.66	24.74		89.66	21.11	
Labeling pretest			155			131
Raw score	9.09	7.18		9.45	6.43	
Labeling posttest			141			124
Raw score	18.57	9.96		16.23	8.42	
Definitions posttest			108			95
Raw score	20.64	7.32		20.67	6.79	

Note. The labeling task refers to a researcher-made probe in which participants labeled targeted words in English. The definition task refers to a researcher-made probe in which children were asked to explain the meaning of words. PPVT-4 = Peabody Picture Vocabulary Test—Fourth Edition (Dunn & Dunn, 2007); TVIP = Test de Vocabulario en Imagenes Peabody (Dunn et al., 1986); WRMT-III = Woodcock Reading Mastery Tests—Third Edition (Woodcock, 2011); RAN = Rapid Automated Naming; letter ID = letter identification subtest; PA = phonological awareness subtest; BESA = Bilingual English–Spanish Assessment (Peña et al., 2014).

Informal Probes

In addition to the standardized assessments, research assistants administered informal researcher-designed vocabulary probes as proximal measures of word learning through labeling and expressive definitions. Informal researcher-made probes for the set of targeted weekly vocabulary words were administered as benchmarks aligned with the curriculum to observe participants' responsiveness to the intervention. Probes included labeling and expressive definitions resembling those utilized in the relevant literature (Beck & McKeown, 2007; Nash & Donaldson, 2005; Ordonez et al., 2002). Refer to Appendix A for a list of words used in the creation of the informal probes.

Labeling

At the beginning and end of the school year, examiners administered labeling probes by displaying pictures of targeted vocabulary words electronically on a PowerPoint presentation (one image per slide) with the prompt, "What is this?" Children were then required to name the target word in English. The labeling probes consisted of 54 items

for both kindergarten and first grade. All items reflected target words from the administered e-books. If children provided the label in Spanish, they were encouraged to label the picture again in English. Misarticulations were not counted as errors unless the target word was unrecognizable (i.e., more than three phonemes in the word substituted or omitted). Items were administered in random order and scored with a 1 or 0.

Expressive Definitions

Testing children's ability to construct definitions has been widely used as a measure of the richness of children's semantic representations of words (Ordonez et al., 2002; Snow, 1990). At the end of the school year, examiners administered a definition probe that consisted of 13 items for kindergarten participants and 14 items for first-grade participants. Participants were asked to define the word given three prompts: (a) "Tell me what [*target word*] means," (b) "What else do you know about [*target word*]?" and (c) "Can you tell me anything else about [*target word*]?" Examiners recorded responses, which were scored by three

research assistants who were blind to the participants' assignment condition. Research assistants used a rubric to obtain a scaled score (Appendix B), consistent with those used in previous studies (Beck & McKeown, 2007; Justice et al., 2005; Lugo-Neris et al., 2010; Ordonez et al., 2002), employing a 0–3 scale. The probes were scored by two examiners with 85.8% agreement. A third scorer served as tie breaker to derive a final score.

Procedures

Computer station arrangements varied between schools based on space, teacher preferences, and existing classroom layouts. In four schools, clusters of six to eight computers were placed in a corner of the room reserved for small group activities. The e-book session was then integrated as a center with a rotation of students assigned. Students in some schools participated within their classrooms in small groups, whereas other schools used teacher offices in the back of the classroom as the designated “center” space for the e-books administered on laptop computers. In four schools, a room of computers in the media center, space adjacent to the classroom, and/or a portable unit were utilized. Regardless of physical layout of the implementation, children participated three times a week as a rotation during differentiated instruction time blocks.

Teachers, classroom assistants, paraeducators, volunteers, and researchers assisted children with logging into the computers. Variability with log-in assistance depended on the day of the week and availability of classroom support personnel. Although touchscreen laptop computers were used with most students, at one school, children participated using their school's iPads while seated at their desks. To avoid potential errors in implementation, the investigators created an individualized unique electronic log-in for each child. Although all children had access to the same book titles (regardless of assigned condition), the individual log-in granted access to the specific version of the selected book based on whether the child was assigned to the read-only (comparison) condition or the intensive intervention condition.

Fidelity of Implementation

Fidelity of implementation was assessed using two methods. First, implementers completed paper-based logs to track attendance and participation in the e-book. Second, Moodle (Dougiamas & Taylor, 2003), an open-source online platform and Learning Management System that was used to house the e-books, recorded data on the dates and times that each child was logged into an e-book in their assigned condition. The comparison, or read-only, condition participants completed 2.77 sessions per week on average ($SD = 0.85$ session). The intervention condition participants completed 2.82 sessions per week on average ($SD = 0.63$ session). There were no significant differences in the average number of sessions per week between the two groups, $F(1, 286) = 1.09, p = .297$. Additional tables of data disaggregated by group and school are provided in Appendix C.

There was a significant difference in the average number of treatment weeks completed by the participants between the two conditions. Because the homoscedasticity assumption was violated (Levene's test: $F = 6.34, p = .012$), a non-parametric comparison was conducted. A Mann–Whitney U test revealed that the average number of treatment weeks was greater for the comparison group ($Mdn = 19.5$) than for the intervention group ($Mdn = 17.0$), $U = 7435.50, p = .001$. The comparison group received 17.69 weeks of treatment on average ($SD = 3.89$ weeks), and the intervention group received 16.83 weeks on average ($SD = 4.35$ weeks). This difference can be partially attributed to when each school enrolled in the project. Several schools did not enroll in the project until later in the fall, resulting in fewer weeks of treatment being delivered to specific schools. This explanation is corroborated by the findings from the Levene's test, which revealed unequal variances ($F = 6.34, p = .012$) for the number of treatment weeks between the two conditions. Importantly, differences in treatment receipt were in favor of the comparison group. The intervention group did not receive more weeks of treatment than the comparison group.

Analyses

Linear mixed models (LMMs) and linear quantile mixed models (LQMMs) were used for testing the impacts of the intervention on two proximal and one distal language-related outcomes. The LMM was used to estimate the average treatment effect on the selected outcomes while adjusting the standard errors for the covariates for the non-independence of units due to the nesting of students in classrooms and schools. The LMM analysis allows for estimation of effect sizes of the intervention compared with the control condition while accounting for variation attributable to classroom- and school-specific characteristics. In a similar manner, LQMM estimates treatment effects accounting for the multilevel structure of the data. Unlike LMMs that estimate regression coefficients conditional on the mean of the posttest distribution, quantile models estimate regression coefficients conditional on many points of the posttest distribution (Koenker & Bassett, 1978; Petscher, 2016). A quantile may be viewed as conceptually similar to a percentile or fractile, and quantile regression may be viewed as a special case of median regression (Koenker & Bassett, 1978). Although traditional mixed effects models that are based on the conditional mean produce coefficients that reflect the average relation between a set of covariates and a selected outcome, this approach may be potentially limited because of the lack of strict normality in the measured variables or circumstances where one has a hypothesis that the relationship between the independent and dependent variables differs for students at one end of the distribution compared with another. Moreover, an average treatment effect may mask stronger or weaker treatment effects that may exist at other points of the conditional posttest distribution relative to the mean. In this way, where an LMM tests for average treatment effects, LQMM tests for local treatment effects

(Wanzek et al., 2016). For example, the difference between treatment and control groups may differ at the conditional mean, but larger or smaller treatment effects may exist at the 0.25 quantile (similar to the 25th percentile) of the posttest distribution or that at the 0.75 quantile (similar to the 75th percentile) of the posttest distribution. In such instances, quantile regression is a useful analysis to estimate the conditional relationship between the independent and dependent variables at selected points of the outcome distribution. Its lack of assumptions about the shape of the outcome distribution (Koenker & Bassett, 1978) and ability to estimate individual conditional effects are key features that make this possible.

A useful mechanism for understanding the similarities and differences between the LMM and LQMM may be viewed through a formulaic expression of each. The given LMM for the primary impact analysis in this study is

$$Y_{ijk} = \gamma_{000} + \gamma_{100}(Pretest_{ijk}) + \gamma_{200}(BLOOM_{ijk}) + e_{ijk} + u_{0jk} + r_{0k}, \quad (1)$$

where Y_{ijk} is the posttest score on measure Y for student i in classroom j in school k ; γ_{000} is the conditional mean posttest score for the control group; $\gamma_{200}(BLOOM_{ijk})$ is the fitted deflection of the treatment group mean from the control on the posttest conditional mean, controlling for the student pretest [i.e., $\gamma_{100}(Pretest_{ijk})$]; e_{ijk} is the student level residual; r_{0k} is the school level residual; and each residual has a respective variance. In a similar manner, the LQMM equation for the primary impact analysis is

$$Y_{j\tau} = \gamma_{00\tau} + \gamma_{10\tau}(Pretest_{j\tau}) + \gamma_{20\tau}(BLOOM_{j\tau}) + e_{j\tau} + r_{0\tau}. \quad (2)$$

Note the equivalence between the two expressions such that each contains terms for the parameters in the model with subscripts noting clustering units; coefficients for means of the control group, pretest, and experimental group; and residuals with associated variances. The primary difference between the two equations is that the LQMM expression includes the subscript τ , denoting the quantile at which the coefficients are estimated, and the LMM includes the random effect for classrooms and schools, whereas the LQMM only allows for two random effects (i.e., students and classrooms) with cluster-corrected standard errors for schools.

A particular benefit of the quantile regression model is that it does not make assumptions about the distribution of the outcome or predictor variables. The estimation process for intercept and slope coefficients in LQMM is similar to LMM in that it uses a loss function; the chief difference between the two approaches is that the latter conditions only on the mean whereas the LQMM conditions at as many points of the posttest that the user is interested in testing. LMMs would require splitting posttest data into

quartiles, quintiles, deciles, or the like to look at evaluated associations conditional on posttest performance. Such methods have been shown to be problematic (Petscher, 2016) due to a truncated range of scores and diminished power for estimating associations. Quantile regression models do not truncate the sample but rather use the full data matrix with differential weights to estimate the intercept and regression coefficients. Subsequently, there is no loss of power or truncation of the range of the dependent variable.

For both LMM and LQMM effects, Hedges's g was estimated as a measure of effect size. Hedges's g provides a less biased estimate of intervention effects compared with Cohen's d , which can overestimate effect sizes, and is computed using a correction factor (Hedges, 1981). After the primary impact LMM and LQMM analyses, exploratory analyses tested the extent to which treatment effects were moderated by grade level or baseline performance.

Because the present design was a cluster randomized trial, it was necessary to use an LQMM (Geraci & Bottai, 2014), which estimates unique random and fixed effects for each specified quantile (see Geraci & Bottai, 2014, for technical details). Analyses were performed with the LQMM package (Geraci, 2014) for the R environment (R Core Team, 2017); however, the package only allows for two-level models. As previously noted, the standard errors were cluster-corrected to account for school level effects. Before running the LQMM, it was necessary to specify the number of quantiles to estimate. Although it is possible to specify as many quantiles as you have equal to the range of scores on the dependent variable, the specification should be based on the sample size and number of parameters in the number (Cade & Noon, 2003). Given the relatively small cluster size and overall participant pool, we opted to select five specific quantiles (i.e., 0.10, 0.25, 0.50, 0.75, and 0.90 quantiles). We chose these values to reflect the variation of approximate treatment effects along lower levels of the conditional posttest distribution (0.10 and 0.25), at the median of the conditional distribution (0.50), and at higher levels of the conditional distribution (0.75 and 0.90).

Results

Descriptive Statistics

The means and standard deviations of all assessments administered are reported in Table 3, disaggregated by the treatment and control groups. Examination of the experimental measures (i.e., PPVT-4, labeling probes, and definitions) revealed that both the treatment and control children's average PPVT-4 standard scores at baseline (i.e., 81.20 and 81.87, respectively) were more than 1 SD below the normative mean. The raw scores for both groups demonstrated that students' average vocabulary scores increased from the fall to spring assessments. For the labeling probes, both the treatment and control participants performed similarly on average in the fall (i.e., 9.09 and 9.45, respectively). However, in the spring, treatment participants achieved slightly higher scores on average (18.57) compared with the control

participants' average (16.23) for the labeling probe. Results from the labeling probes indicate that both groups achieved higher scores in the spring than in the fall. The definition probes were only administered in the spring and revealed similar average scores for the treatment (20.64) and control (20.67) children.

Correlations among the experimental measures (Table 4) ranged from $r = .40$ between the spring PPVT-4 standard score and the Spring Definitions task to $r = .93$ between the fall PPVT-4 raw and standard scores. Missing data rates ranged from 2% to 31% across the measures included in the impact models for the full sample, from 1% to 20% for the control group, and from 0% to 30% for the treatment group. Differential attrition was less than or equal to 3% for all measures with the exception of the Spring Definitions (i.e., 10%). Little's test of data missing completely at random resulted in a significant effect, $\chi^2(48) = 100.10$, $p < .001$, suggesting that the data were not missing completely at random; however, an evaluation of the data did not point to the mechanism for missingness being due to the observed measures. As such, the data were assumed to be missing at random.

Impact Model Results

Before the testing of the intervention effects, baseline equivalence was evaluated on the vocabulary assessments. Results from the baseline LMM analyses indicated that no significant differences were observed between the two groups on either the PPVT-4 ($t = 0.65$, $p > .500$, Hedges's $g = 0.07$) or labeling ($t = -0.27$, $p > .500$, Hedges's $g = -0.008$). The initial unconditional model suggests that variances in each of the outcomes could be reasonably attributed to between-student, classroom, and school differences for the PPVT-4 (student intraclass correlation coefficient [ICC] = 65%, classroom ICC = 8%, school ICC = 27%), definitions (student ICC = 50%, classroom ICC = 2%, school ICC = 48%), and labeling (student ICC = 72%, classroom ICC = 11%, school ICC = 17%). Subsequently, a three-level model was used to

test LMM primary and moderated impacts. Table 5 summarizes the results from the primary impact analyses. Statistically significant effects were initially observed for the PPVT-4 ($t = 2.39$, $p = .018$) and labeling ($t = 5.76$, $p < .001$) assessments even after applying a linear step-up to correct for Type I errors (i.e., critical p value = .033). Hedges's g for the PPVT-4 was calculated as 0.18, and that for labeling the effect was 0.38, both modest effects, with the effect on labeling being small–moderate and that for the PPVT-4 being small (Cohen, 1988).

LQMM results (Table 6) show that the effect of the intervention was stronger at the upper end of the distribution of the PPVT-4. That is, although the LMM showed a small average effect of the intervention, the LQMM shows that, at low levels of the PPVT-4 posttest (e.g., the 0.10 and 0.25 quantiles), there was no practically important effect ($g = -0.01$ and 0.04, respectively) but that, at the 0.90 quantile of the PPVT-4, the effect was small (0.11). It should be noted that only a few quantiles of the conditional post-test distribution were selected for testing and should not be averaged and compared with the LMM. The LQMM further showed that, despite a lack of statistical significance on the definitions outcome, a small effect was observed at the 0.50 and 0.75 quantiles (i.e., $g = 0.08$ and 0.09, respectively). LQMM for the definitions outcome presented with small yet not statistically significant effects (range of $g = 0.07$ –0.19). Moreover, when considering the labeling probe outcome, it can be seen that the effect of the intervention was significant across all of the selected quantiles ($p < .001$), with effect sizes ranging from $g = 0.33$ to 0.44. Exploratory LMM analyses showed that no significant moderation of treatment effects for the PPVT-4 or labeling existed based on grade or pretest scores (Appendix F). The definition outcome presented with a significant interaction between treatment and baseline PPVT-4 at the 0.50 quantile (.15, $p = .032$). Despite the interaction, simple slopes were not evaluated because there was no systematic moderation across a range of quantiles conditional on the distribution of posttest definition.

Discussion

The primary aim of the current intervention study with random assignment was to test whether a vocabulary intervention embedded in shared e-book readings would cause more vocabulary gains than shared reading alone without embedded intensified instruction. Overall, the findings indicated that there were modest, significant effects of the intervention on one proximal (labeling) and one distal measure (PPVT-4; Dunn & Dunn, 2007), with no significant effects on children's skills at constructing word definitions. The e-book vocabulary intervention with embedded word explanations in Spanish was feasible to implement in classrooms without bilingual implementers.

Comparison With Existing Literature

Findings revealed a consistent small impact of group assignment on the experimenter-created proximal measure

Table 4. Correlations and missing data.

Variable	1	2	3	4	5	6	7
1. Fall PPVT-4 raw	1.00						
2. Fall PPVT-4 SS	.93	1.00					
3. Fall Labeling	.81	.78	1.00				
4. Spring PPVT-4 raw	.82	.75	.76	1.00			
5. Spring PPVT-4 SS	.64	.72	.70	.90	1.00		
6. Spring labeling	.68	.70	.81	.73	.71	1.00	
7. Spring definitions	.61	.49	.56	.58	.40	.56	1.00
% Missing data full	.05	.05	.02	.12	.12	.09	.31
% Missing data control	.04	.04	.01	.08	.08	.06	.20
% Missing data treatment	.03	.03	.00	.12	.12	.09	.30

Note. The labeling task refers to a researcher-made probe in which participants labeled targeted words in English. The definition task refers to a researcher-made probe in which children were asked to explain the meaning of words. PPVT-4 = Peabody Picture Vocabulary Test–Fourth Edition (Dunn & Dunn, 2007); SS = standard score.

Table 5. Linear mixed model primary impact and moderation results.

Model	Parameter	Coefficient	SE	df	t value	p value
PPVT-4 raw impact	Intercept	82.18	2.16	242	38.11	< .001
	Fall PPVT-4 raw	1.00	0.06	242	15.78	< .001
	Grade	12.79	2.14	242	5.99	< .001
	Treatment	4.11	1.72	242	2.39	.018
Definition impact	Intercept	17.20	0.99	189	17.21	< .001
	Fall PPVT-4 raw	0.15	0.02	189	5.92	< .001
	Grade	5.21	1.21	189	4.23	< .001
	Treatment	0.93	0.69	189	1.33	.184
Labeling impact	Intercept	15.06	1.01	250	14.92	< .001
	Fall labeling	1.04	0.05	250	20.48	< .001
	Grade	-0.52	0.78	250	-0.67	.504
	Treatment	3.50	0.61	250	5.76	< .001
PPVT-4 raw grade moderation	Intercept	40.07	3.89	130	10.31	< .001
	Fall PPVT-4 raw	0.69	0.04	243	18.12	< .001
	Treatment	2.86	2.10	241	1.36	.176
	Grade	-2.24	2.78	100	-0.81	.423
	Grade × Treatment	0.70	3.15	244	0.22	.824
Definition grade moderation	Intercept	13.99	1.68	107	8.34	< .001
	Fall PPVT-4 raw	0.10	0.02	186	6.18	< .001
	Treatment	1.48	0.90	185	1.64	.103
	Grade	-2.55	1.49	48	-1.71	.093
	Grade × Treatment	-1.92	1.41	186	-1.36	.176
Labeling grade moderation	Intercept	4.31	1.22	18	3.54	.002
	Fall labeling	1.05	0.05	257	20.66	< .001
	Treatment	4.06	0.80	246	5.02	< .001
	Grade	1.07	1.03	69	1.04	.304
	Grade × Treatment	-1.21	1.19	250	-1.02	.310
PPVT-4 raw baseline moderation	Intercept	41.45	4.81	175	8.61	< .001
	Fall PPVT-4 raw	0.67	0.05	246	12.53	< .001
	Treatment	0.82	4.84	246	0.17	.866
	Fall PPVT-4 × Treatment	0.03	0.06	242	0.51	.608
Definition baseline moderation	Intercept	16.36	2.05	146	7.98	< .001
	Fall PPVT-4 raw	0.08	0.02	187	3.42	.001
	Treatment	-3.07	2.18	187	-1.41	.161
	Fall PPVT-4 × Treatment	0.05	0.03	185	1.82	.071
Labeling baseline moderation	Intercept	5.11	1.30	24	3.91	< .001
	Fall labeling	1.00	0.07	256	13.59	< .001
	Treatment	2.65	1.05	251	2.52	.012
	Fall Label × Treatment	0.09	0.09	248	1.00	.317

Note. This analysis allows for estimation of the intervention effect sizes compared with the control condition while accounting for nesting of children within classrooms and schools. Unlike the linear quantile mixed model, the effect sizes estimated at a single quantile, the mean (i.e., 0.50 quantile), are similar to linear regression. For further reading, see Koenker and Bassett (1978) and Petscher (2016). PPVT-4 = Peabody Picture Vocabulary Test–Fourth Edition.

of vocabulary labeling. This finding substantiated previously reported findings in the literature of intervention studies targeting similar skills with ELs, which show treatment effects on measures closely aligned with the intervention (e.g., Carlo et al., 2004; Lugo-Neris et al., 2010; R. Silverman & Hines, 2009). The significant effects on the distal measure (i.e., the PPVT-4, the standardized English vocabulary assessment), however, were surprising compared with previous findings in the literature. In a recent meta-analysis of the impact of shared reading on EL language and literacy growth, the average effect size on distal measures of language and literacy obtained from peer-reviewed studies was $g < 0.01$ (Fitton, McIlraith, & Wood, 2016). Notably, the words included on the PPVT-4 were not directly taught in the intervention. Subsequently, it is reasonable to suggest that the effect of the intervention may generalize to more distal outcomes.

Although the cause of the effect cannot be derived from the data, we offer several possible explanations. Although maturation and classroom or teacher effects are common suspects, the design of the current study minimized these potential influences by randomly assigning participants to treatment and comparison groups within classrooms. In other words, children in both conditions received instruction in the same classrooms across the same period of maturation. An alternative explanation, considering the predictions of the RHM (Kroll & Stewart, 1994) for lexical level processing via the translation equivalent at the beginning stages of vocabulary development, is it is likely that the bridge to the L1 served as a crutch to help link the corresponding word to the L2, English. Potentially, as children developed better semantic “webs,” from intervention, it strengthened their semantic networks enabling them to grasp other new words

Table 6. Linear quantile mixed model primary impact results.

Outcome	Quantile	Parameter	Coefficient	SE	LB	UB	p value	Hedges's g
PPVT-4 raw	0.10	Intercept	80.01	1.76	76.46	83.54	< .001	-0.02
		Fall PPVT-4 raw	0.66	0.07	0.53	0.81	< .001	
		Grade	-1.12	2.53	-6.22	3.97	.659	
	0.25	Treatment	-0.44	1.15	-2.76	1.87	.705	
		Intercept	84.88	1.47	51.93	87.85	< .001	
		Fall PPVT-4 raw	0.68	0.05	0.58	0.77	< .001	
	0.50	Grade	-2.44	1.46	-5.36	0.49	.100	
		Treatment	0.00	0.88	-1.77	1.80	.998	
		Intercept	87.88	1.34	85.19	90.58	< .001	
	0.75	Fall PPVT-4 raw	0.67	0.04	0.57	0.76	< .001	
		Grade	-2.89	1.52	-5.94	0.17	.063	
		Treatment	1.78	1.13	-0.49	4.04	.121	
	0.90	Intercept	91.82	1.60	88.60	95.04	< .001	
		Fall PPVT-4 raw	0.67	0.05	0.56	0.78	< .001	
		Grade	-1.09	1.89	-4.90	2.71	.567	
Definition	0.10	Treatment	2.05	1.13	-0.23	4.32	.077	0.09
		Intercept	96.11	1.77	92.55	99.68	< .001	
		Fall PPVT-4 raw	0.70	0.07	0.55	0.84	< .001	
	0.25	Grade	-3.70	2.21	-8.16	0.75	.101	
		Treatment	3.00	1.77	-1.08	6.03	.167	
		Intercept	10.37	1.81	7.20	14.47	< .001	
	0.50	Fall PPVT-4 raw	0.15	0.04	0.07	0.23	< .001	
		Grade	5.54	1.78	1.96	9.12	.003	
		Treatment	1.33	1.32	-1.33	3.99	.319	
	0.75	Intercept	15.78	1.62	12.51	19.04	< .001	
		Fall PPVT-4 raw	0.17	0.04	0.09	0.24	< .001	
		Grade	2.50	1.98	-1.47	6.47	.212	
	0.90	Treatment	0.67	1.19	-1.73	3.06	.578	
		Intercept	17.62	1.32	14.97	20.26	< .001	
		Fall PPVT-4 raw	0.16	0.04	0.08	0.24	< .001	
Probe	0.10	Grade	4.06	1.52	1.01	7.11	.009	0.10
		Treatment	0.69	1.07	-1.46	2.84	.519	
		Intercept	20.22	1.30	17.62	22.83	< .001	
	0.25	Fall PPVT-4 raw	0.14	0.04	0.05	0.22	.002	
		Grade	5.98	1.38	3.19	8.86	< .001	
		Treatment	1.27	1.02	-0.78	3.32	.218	
	0.50	Intercept	24.40	1.80	20.76	28.03	< .001	
		Fall PPVT-4 raw	0.12	0.05	0.02	0.21	.013	
		Grade	3.94	1.97	-0.02	7.91	.051	
	0.75	Treatment	0.55	1.20	-1.85	2.97	.643	
		Intercept	9.44	1.77	5.88	12.99	< .001	
		Fall probe	1.00	0.08	0.08	1.16	< .001	
	0.90	Grade	-0.99	2.06	-5.15	3.15	.630	
		Treatment	3.99	1.12	1.76	6.25	< .001	
		Intercept	13.08	1.46	10.14	16.01	< .001	
Probe	0.10	Fall probe	1.08	0.09	0.90	1.26	< .001	0.44
		Grade	-0.89	1.48	-3.86	2.09	.551	
		Treatment	2.99	0.89	1.21	4.79	.002	
	0.25	Intercept	15.67	1.12	13.41	17.93	< .001	
		Fall probe	1.10	0.08	0.93	1.26	< .001	
		Grade	-1.33	1.33	-4.01	1.34	.321	
	0.50	Treatment	3.00	0.97	1.05	4.96	.003	
		Intercept	18.48	1.06	16.35	20.60	< .001	
		Fall probe	1.09	0.08	0.93	1.26	< .001	
	0.75	Grade	-1.44	1.13	-3.72	0.82	.208	
		Treatment	3.72	1.28	1.15	6.30	.005	
		Intercept	20.42	1.49	17.43	23.41	< .001	
	0.90	Fall probe	1.07	0.11	0.79	1.23	< .001	
		Grade	0.96	1.56	-2.18	4.11	.541	
		Treatment	3.01	1.28	0.43	5.59	.023	

Note. This analysis allows for estimation of the intervention effect sizes compared with the control condition while accounting for nesting of children within classrooms and schools, like the linear mixed model. Importantly, however, the effect sizes are estimated based on children's initial PPVT-4 scores. Separate effect sizes are reported for the five quantiles of interest to determine if children respond differently to the intervention based on their starting receptive English vocabulary. For further reading, see Koenker and Bassett (1978) and Petscher (2016). PPVT-4 = Peabody Picture Vocabulary Test-Fourth Edition; LB = lower bound; UB = upper bound.

not taught. This explanation is consistent with the literature on the intricate semantic relationships conceptualized as a lexico-semantic network (Aitchison, 2012) and spreading activation that is thought to occur along the network (Neely, 1991). It is also possible that children who received embedded vocabulary support with scaffolding throughout the school year became better at inferring the meaning of new words from embedded everyday exposures. This explanation is consistent with word learning models in which children are active learners with attentional word learning biases, learning how to be better word learners through experience (Smith, 2000). Finally, leveraging morphology in the word maps may have improved children's ability to bootstrap the meanings of words, which improved their overall word learning. The contributions of morphological knowledge on reading vocabulary have been supported by other studies (e.g., Goodwin, Huggins, Carlo, August, & Calderon, 2013).

The finding that no statistically significant effect sizes were identified on definition skills was unanticipated, given that other vocabulary intervention studies involving ELs have demonstrated effects on expressive definitions (e.g., Lugo-Neris et al., 2010) at similar ages. In a teacher-implemented explicit vocabulary intervention in Spanish with first-grade ELs, researchers reported a main effect on definition skills (Cena et al., 2013) using a similar rubric. Although explaining the cause of the differences in findings is beyond the scope of this study, it is possible that the task of formulating thorough, well-constructed definitions of words was beyond the language ability of young participating ELs. Further examination of the data on expressive definition is planned to allow for comparative analyses across multiple indexes of definition scoring. It is also possible that explicit instruction on defining words may be necessary to stimulate outcomes on definitions. For example, other studies have integrated active practice for children in formulating definitions, with the example: "Your turn, what is the definition for the word 'instructions'?" (Cena et al., 2013, p. 1310), which may be beneficial and necessary to produce impacts on definition skills.

Quantile regression results provided additional depth in understanding the effect sizes obtained for each of the vocabulary outcome measures. Effect sizes on the standardized English vocabulary assessment (i.e., PPVT-4) were evidently larger conditional at higher levels of vocabularies in English outcome compared with weaker effects conditional at lower levels of English vocabularies at outcome. At the 90th quantile of posttest performance on the PPVT-4, the effect size of the intervention was .19 ($p = .004$), whereas effect sizes were at .06 or below for posttest performance below the 50th quantile on the PPVT-4. Similarly, proximal measures also produced larger effect sizes at the upper quantiles compared with the lower quantiles. These findings suggest that there were differences in how ELs' vocabulary grew within and between the two reading groups. There was a significant overall impact of the intervention compared with the control on proximal and distal measures of vocabulary, but the effect was different for ELs with varying levels

of English vocabulary. Certain subgroups of ELs appeared to respond differentially to the intervention, which was considered in terms of the RHM (Kroll & Stewart, 1994). Within the RHM, learners transition from reliance on the translation equivalent to conceptual mediation with increasing skill in the L2. A plausible interpretation might be that upper quantile reflects those who have made the transition to conceptual mediation and are thus better able to take advantage of the other components of the intervention. Thus, the benefit of L1 bridging may be more apparent once learners have moved beyond the reliance on the L1 lexical links that is necessary at initial stages of L2 vocabulary development. This interpretation is speculative, and the nature of the differential effects should be explored more directly in future research.

It is beyond the scope of the current design to identify a cause for the effects observed, but it is possible that dosage or intensity contributed to impacts on the two vocabulary measures. The longer duration of the current intervention, compared with the shorter duration typically employed in similar programs (e.g., 8 weeks in Cena et al., 2013; 15 weeks in Carlo et al., 2004), may have contributed to the observed differences in the distal measure. However, not all of the participating schools were able to include all 20 weeks in the current study, so this is not a conclusive explanation. It is feasible that the interplay of active ingredients (e.g., scaffolding with Spanish and highlighting morphology) was responsible for the effects. The relationship between morphological awareness and vocabulary gains has been substantiated in other studies with similar-aged monolinguals (Sparks & Deacon, 2015).

Limitations

Given that multiple intervention strategies were packaged to intensify the vocabulary instruction, critical components cannot be determined from the current study. It is possible that not all of the instructional components were necessary or that one component was the critical active ingredient or agent of change. Although the differential effects of individual strategies cannot be considered from the bundle implementation, the cluster of word learning strategies employed is consistent with established best practices to vocabulary instruction for ELs.

Although the fidelity of implementation data supports the feasibility of implementation in early elementary level classes, it should be noted that there were challenges to implementation that should be acknowledged. Implementation obstacles primarily related to scheduling, absences, and competing demands (e.g., field trips, picture day, assemblies, or misbehavior). Similar to difficulties reported in other randomized controlled trial language intervention projects in classrooms (LaRusso, Donovan, & Snow, 2016), the schedule for intervention required flexibility to accommodate field trips and special activities. In addition, absences were a common barrier to implementation, with missing data and reduced frequency of readings. Another barrier to implementation was the location of the computer laboratories

or centers. Although many schools utilized space within the classroom, for classrooms that chose an adjacent space or a separate media center or portable unit, implementation required a transition of walking to and from the location of the e-book sessions. The transition time took away instructional time and risked potential reentry or engagement delays when children settled back into their classroom.

In addition to implementation challenges, the results should be interpreted cautiously because of limitations in the methods and design, with particular consideration of (a) variability in classroom characteristics, (b) retention of gains, and (c) social validity. Ideally, ELs would have been equally distributed in participating classrooms; however, there was variability between schools as to how ELs were placed within classrooms. As a result, some participating classrooms had a high proportion of ELs, and other schools had eight participating classrooms with a balanced number in each. Furthermore, the current design did not allow for measurement of retention over a longer period. Examination of the children's word knowledge after the summer or a school year later would inform the broader importance of this work. In addition, although we examined effects on word learning through proximal and distal measures, measures of social validity at the end of the school year would be beneficial to include. In the current design, teachers were blind as to which students received read-only versus the intensive instruction on the computer. It would be interesting in future studies to assess teachers' perceptions of the children's progress and word learning to determine if teachers felt that the intervention made a noticeable impact on their performance.

Implications

The significant effects of the vocabulary intervention on labeling and word understanding provide some evidence for the effectiveness of computer-assisted intensive vocabulary instruction that includes definitions, bridging to L1, repetitions, and morphology for ELs in kindergarten and first grade. A strength of the computer-assisted delivery was that it was feasible for monolingual personnel to leverage Spanish-language scaffolding without bilingual implementers. A feasible, effective approach for employing rich supplemental vocabulary instruction could have positive implications for elementary schools facing challenges in providing effective vocabulary instruction for ELs. Currently, there is a great deal of variability in the utilization of ELs' L1 to scaffold instruction warranting flexible, innovative curricular strategies (Slavin & Cheung, 2005). Although there is consensus that instruction should be sensitive to students' cultural backgrounds and linguistic proficiencies (Brown & Doolittle, 2008), programs may lack feasible options and/or expertise on how to interweave LI bootstrapping. Because most modern classrooms have access to computers in school buildings, the intensified e-book instruction provides a practical approach for classroom teachers to implement recommended practices for intensive small-group interventions for rich

vocabulary instruction with comprehensible scaffolding with L1 supports (Gersten et al., 2007).

Considerations for Future Research

The finding that the intervention did not produce significant effects on constructing definitions in the sample of young ELs warrants closer examination in future work (Cena et al., 2013). Follow-up study should integrate explicit instruction on constructing definitions and compare the effect on definitions. Although we expected that definitions would improve with instruction, it is also possible that the definition rubric was scaled to capture sophisticated definitions that young ELs are not developmentally ready to compose.

In addition, further examination of the students' click data on readings nested within weeks might provide insight into how children's interactions with the books impact word learning. The authors plan to further examine the students' click data within e-books to explore differential effects or trends in word types learned quickly. Children may acquire new words with a fewer number of readings as the school year progresses. Some of the books contained one or more target words that were semantically related (e.g., knight, castle, sword), and in other cases, the four target words had little in common with each other. Potential differential learning of the semantically related words could inform selection and grouping of words for instruction.

Finally, although this study focused specifically on how ELs' initial vocabulary knowledge in English moderated their English vocabulary growth during the intervention, attention to cross-language moderation would further inform educational practice. The present findings indicate that ELs with at least basic English vocabulary skills benefited more strongly from the e-books than children with lower levels of English. However, the intervention included native-language bridging, which is theorized to be beneficial for children learning an L2 (Kroll & Stewart, 1994). Examination of children's initial Spanish vocabulary as a moderator would facilitate assessment of the effectiveness of this approach for ELs with varying Spanish proficiencies.

Conclusions

Intensive vocabulary instruction with bridging to L1 embedded in e-books yielded significant effects on ELs' vocabulary growth compared with shared reading without embedded vocabulary instruction. The proximal effect on English labeling was robust across all quantiles regardless of children's initial English vocabulary. The distal effect on the standardized assessment of receptive English vocabulary was stronger for children who started with greater English knowledge. There was no overall effect on children's ability to generate definitions. E-book-delivered intervention with bridging to Spanish was feasible for monolingual personnel to implement in classrooms without access to bilingual implementers.

Acknowledgments

The research reported here was supported by the Institute of Education Sciences, U.S. Department of Education, through Grant R305A130460 to Florida State University. The authors are especially grateful to partnering schools, participating families, and support from research assistants Dana Brown, Kristina Bustamante, Clare Gabas, Rachel Hoge, Angie Joseph, Mary Martin, Vanessa Peña, Amenda Shapiro, Jaclyn Suveg, Jennifer Vamos, and Claire Wofford who were essential in providing research assistance. The opinions expressed are those of the authors and do not represent views of the Institute or the U.S. Department of Education.

References

- Aitchison, J. (2012). *Words in the mind: An introduction to the mental lexicon*. Hoboken, NJ: Wiley.
- Apel, K., & Werfel, K. (2014). Using morphological awareness instruction to improve written language skills. *Language, Speech, and Hearing Services in Schools, 45*, 251–260. https://doi.org/10.1044/2014_LSHSS-14-0039
- Apthorp, H., Randel, B., Cherasaro, T., Clark, T., McKeown, M., & Beck, I. (2012). Effects of a supplemental vocabulary program on word knowledge and passage comprehension. *Journal of Research on Educational Effectiveness, 5*(2), 160–188. <https://doi.org/10.1080/19345747.2012.6602.40>
- August, D., Carlo, M., Dressler, C., & Snow, C. (2005). The critical role of vocabulary development for English language learners. *Learning Disabilities Research and Practice, 20*, 50–57. <https://doi.org/10.1111/j.1540-5826.2005.00120>
- August, D., McCardle, P., & Shanahan, T. (2014). Developing literacy in English language learners: Findings from a review of the experimental research. *School Psychology Review, 43*, 490–498. <https://doi.org/10.17105/SPR-14-0088.1>
- August, D., & Shanahan, T. (Eds.). (2006). *Developing literacy in second-language learners: Report of the National Literacy Panel on Language-Minority Children and Youth*. Mahwah, NJ: Erlbaum.
- Bääth, R. (2010). ChildFreq: An online tool to explore word frequencies in child language. *LUCS Minor, 16*, 1104–1609.
- Baker, S., Lesaux, N., Jayanthi, M., Dimino, J., Proctor, C. P., Morris, J., . . . Newman-Gonchar, R. (2014). *Teaching academic content and literacy to English learners in elementary and middle school* (NCEE 2014-4012). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. Retrieved from <https://ies.ed.gov/ncee/>
- Barnett, W. S., Yarosz, D. J., Thomas, J., Jung, K., & Blanco, D. (2007). Two-way and monolingual English immersion in pre-school education: An experimental comparison. *Early Childhood Research Quarterly, 22*(3), 277–293. <https://doi.org/10.1016/j.ecresq.2007.03.003>
- Beck, I. L., & McKeown, M. G. (2007). Increasing young low-income children's oral vocabulary repertoires through rich and focused instruction. *The Elementary School Journal, 107*(3), 251–271. <https://doi.org/10.1086/511706>
- Brown, J. E., & Doolittle, J. (2008). A cultural, linguistic, and ecological framework for response to intervention with English language learners. *TEACHING Exceptional Children, 40*(5), 66–72. <https://doi.org/10.1177/004005990804000509>
- Brybaert, M., & New, B. (2009). Moving beyond Kučera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods, 41*(4), 977–990. <https://doi.org/10.3758/BRM.41.4.977>
- Burchinal, M., Field, S., Lopez, M. L., Howes, C., & Pianta, R. (2012). Instruction in Spanish in pre-kindergarten classrooms and child outcomes for English language learners. *Early Childhood Research Quarterly, 27*(2), 188–197. <https://doi.org/10.1016/j.ecresq.2011.11.003>
- Cade, B. S., & Noon, B. R. (2003). A gentle introduction to quantile regression for ecologists. *Frontiers in Ecology and the Environment, 1*(8), 412–420. [https://doi.org/10.1890/1540-9295\(2003\)001\[0412:AGITQR\]2.0.CO;2](https://doi.org/10.1890/1540-9295(2003)001[0412:AGITQR]2.0.CO;2)
- Capps, R., Fix, M., Murray, J., Ost, J., Passel, J., & Herwanto, S. (2005). *The new demography of America's schools: Immigration and the No Child Left Behind Act*. Washington, DC: Urban Institute.
- Carlo, M. S., August, D., McLaughlin, B., Snow, C. E., Dressler, C., Lippman, D. N., . . . White, C. (2004). Closing the gap: Addressing the needs of English language learners in bilingual and mainstream classrooms. *Reading Research Quarterly, 39*(2), 188–215. <https://doi.org/10.1598/RRQ.39.2.3>
- Cena, J., Luft Baker, D., Kame'enui, E. J., Baker, S. K., Park, Y., & Smolkowski, K. (2013). The impact of a systematic and explicit vocabulary intervention in Spanish with Spanish-speaking English learners in first grade. *Reading and Writing, 26*, 1289–1315. <https://doi.org/10.1007/s11145-012-9419-y>
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.
- Dalton, B., Proctor, C. P., Uccelli, P., Mo, E., & Snow, C. E. (2011). Designing for diversity: The role of reading strategies and interactive vocabulary in a digital reading environment for fifth-grade monolingual English and bilingual students. *Journal of Literacy Research, 43*(1), 68–100. <https://doi.org/10.1177/1086296X10397872>
- Demski, J. (2011). ELL to go: Two schools transform their ELL programs by giving students around-the-clock-access to some of the latest mobile devices. *Technological Horizons in Education Journal, 38*(5), 28–32.
- Denckla, M. B., & Rudel, R. (1974). Rapid “automatized” naming of pictured objects, colors, letters and numbers by normal children. *Cortex, 10*, 186–202. [https://doi.org/10.1016/S0010-9452\(74\)80009-2](https://doi.org/10.1016/S0010-9452(74)80009-2)
- Dougiamas, M., & Taylor, P. (2003). *Moodle: Using learning communities to create an open source course management system*. Paper presented at the World Conference on Educational Multimedia, Hypermedia and Telecommunications (EDMEDIA) 2003, Chesapeake, VA.
- Dunn, L. M., & Dunn, L. M. (2007). *Peabody Picture Vocabulary Test—Fourth Edition*. Circle Pines, MN: AGS.
- Dunn, L. M., Lugo, S., Padilla, R., & Dunn, L. (1986). *Test de Vocabulario en Imágenes Peabody* [Peabody Picture Vocabulary Test]. Circle Pines, MN: AGS.
- Durán, L. K., Roseth, C. J., & Hoffman, P. (2010). An experimental study comparing English-only and transitional bilingual education on Spanish-speaking preschoolers' early literacy development. *Early Childhood Research Quarterly, 25*(2), 207–217. <https://doi.org/10.1016/j.ecresq.2009.10.002>
- Duran, N. D., & Dale, R. (2014). Perspective-taking in dialogue as self-organization under social constraints. *New Ideas in Psychology, 32*, 131–146. <https://doi.org/10.1016/j.newideapsych.2013.03.004>
- Ebert, K. D., Kohnert, K., Pham, G., Disher, J. R., & Payesteh, B. (2014). Three treatments for bilingual children with primary language impairment: Examining cross-linguistic and cross-domain effects. *Journal of Speech, Language, and Hearing*

- Research*, 57, 172–186. [https://doi.org/10.1044/1092-4388\(2013/12-0388\)](https://doi.org/10.1044/1092-4388(2013/12-0388))
- Ehrler, D. J., & McGhee, M. L.** (2008). *Primary Test of Nonverbal Intelligence (PTONI)*. Austin, TX: Clocktower Hill Research and Publishing Group, LLC.
- Farver, J. A. M., Lonigan, C. J., & Eppe, S.** (2009). Effective early literacy skill development for young Spanish-speaking English language learners: An experimental study of two methods. *Child Development*, 80(3), 703–719. <https://doi.org/10.1111/j.1467-8624.2009.01292.x>
- Fitton, L., McIlraith, A. L., & Wood, C.** (2016, July). *The impact of shared book reading on young ELs' English outcomes: A meta-analysis*. Poster presentation at the Society for the Scientific Study of Reading Conference, Porto, Portugal.
- Geraci, M.** (2014). Linear quantile mixed models: The lqmm package for Laplace quantile regression. *Journal of Statistical Software*, 57(13), 1–29. <https://doi.org/10.18637/jss.v057.i13>
- Geraci, M., & Bottai, M.** (2014). Linear quantile mixed models. *Statistics and Computing*, 24(3), 461–479. <https://doi.org/10.1007/s11222-013-9381-9>
- Gersten, R., Baker, S. K., Shanahan, T., Baker, S. K., Linan-Thompson, S., Collins, P., & Scarcella, R.** (2007). *Effective literacy and English language instruction for English learners in the elementary grades: A practical guide* (NCEE 2007-4011). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. Retrieved from <http://ies.ed.gov/ncee>
- Goodrich, J. M., Lonigan, C. J., Kleuver, C. G., & Farver, J. M.** (2015). Development and transfer of vocabulary knowledge in Spanish-speaking language minority preschool children. *Journal of Child Language*, 43(5), 969–992. <https://doi.org/10.1017/S0305000032X>
- Goodwin, A. P., & Ahn, S.** (2010). A meta-analysis of morphological interventions: Effects on literacy achievement of children with literacy difficulties. *Annals of Dyslexia*, 60(2), 183–208. <https://doi.org/10.1007/s11881-010-0041-x>
- Goodwin, A. P., & Ahn, S.** (2013). A meta-analysis of morphological interventions in English: Effects of literacy outcomes for school-age children. *Scientific Studies of Reading*, 17, 257–285. <https://doi.org/10.1080/10888438.2012.68791>
- Goodwin, A. P., Huggins, A. C., Carlo, M. S., August, D., & Calderon, M.** (2013). Minding morphology: How morphological awareness relates to reading for English language learners. *Reading and Writing*, 26(9), 1387–1415. <https://doi.org/10.1007/s11145-012-9412-5>
- Gross, M., Buac, M., & Kaushanskaya, M.** (2014). Conceptual scoring of receptive and expressive vocabulary measures in simultaneous and sequential bilingual children. *American Journal of Speech-Language Pathology*, 23(4), 574–586. https://doi.org/10.1044/2014_AJSLP-13-0026
- Gutiérrez-Clellen, V. F.** (1999). Language choice in intervention with bilingual children. *American Journal of Speech-Language Pathology*, 8(4), 291–302. <https://doi.org/10.1044/1058-0360.0804.291>
- Gutiérrez-Clellen, V., Simon-Cerejido, G., & Sweet, M.** (2012). Predictors of second language acquisition in Latino children with specific language impairment. *American Journal of Speech-Language Pathology*, 21, 64–77. [https://doi.org/10.1044/1058-0360\(2011\)10-0090](https://doi.org/10.1044/1058-0360(2011)10-0090)
- Hargrave, A. C., & Sénéchal, M.** (2000). A book reading intervention with preschool children who have limited vocabularies: The benefit of regular reading and dialogic reading. *Early Childhood Research Quarterly*, 15, 75–90. [https://doi.org/10.1016/S0885-2006\(99\)00038-1](https://doi.org/10.1016/S0885-2006(99)00038-1)
- Hedges, L. V.** (1981). Distribution theory for Glass's estimator of effect size and related estimators. *Journal of Educational Statistics*, 6(2), 107–128. <https://doi.org/10.2307/1164588>
- Honig, B., Diamond, L., Gutlohn, L., & Cole, C. L.** (2008). *Teaching reading sourcebook*. Washington, DC: Arena Press.
- Houghton Mifflin Harcourt.** (2015). *Journeys*. Boston, MA: Houghton Mifflin Harcourt Company. Retrieved from <http://www.hmhco.com/journeys>
- Justice, L.** (2006). Evidence-based practice, response to intervention, and the prevention of reading difficulties. *Language, Speech, and Hearing Services in Schools*, 37, 284–297. [https://doi.org/10.1044/0161-1461\(2006\)033](https://doi.org/10.1044/0161-1461(2006)033)
- Justice, L., 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(1), 17–32. [https://doi.org/10.1044/0161-1461\(2005\)003](https://doi.org/10.1044/0161-1461(2005)003)
- Kena, G., Hussar, W., McFarland, J., de Brey, C., Musu-Gillette, L., Wang, X., . . . Dunlop Velez, E.** (2016). *The condition of education 2016* (NCES 2016-144). Washington, DC: National Center for Education Statistics, U.S. Department of Education. Retrieved from <http://nces.ed.gov/pubsearch>
- Kena, G., Musu-Gillette, L., Robinson, J., Wang, X., Rathbun, A., Zhang, J., . . . Dunlop Velez, E.** (2015). *The condition of education 2015* (NCES 2016-144). Washington, DC: National Center for Education Statistics, U.S. Department of Education. Retrieved from <http://nces.ed.gov>
- Kieffer, M. J.** (2008). Catching up or falling behind? Initial English proficiency, concentrated poverty, and the reading growth of language minority learners in the United States. *Journal of Educational Psychology*, 100(4), 851–868. <https://doi.org/10.1037/0022-0663.100.4.851>
- Koenker, R., & Bassett, G., Jr.** (1978). Regression quantiles. *Econometrica*, 46, 33–50. <https://doi.org/10.2307/1913643>
- Kohnert, K., Yim, D., Nett, K., Kan, P. F., & Duran, L.** (2005). Intervention with linguistically diverse preschool children: A focus on developing home language(s). *Language, Speech, and Hearing Services in Schools*, 36, 251–263. [https://doi.org/10.1044/0161-1461\(2005\)025](https://doi.org/10.1044/0161-1461(2005)025)
- Koskinen, P. S., Blum, I. H., Bisson, S. A., Philips, S. M., Creamer, T. S., & Baker, T. K.** (2000). Book access, shared reading, and audio models: The effects of supporting the literacy learning of linguistically diverse students in school and at home. *Journal of Educational Psychology*, 92(2), 23–36. <https://doi.org/10.1037/0022-0663.92.1.23>
- Kroll, J. F., & Stewart, E.** (1994). Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. *Journal of Memory and Language*, 33, 149–174.
- La Heij, W., Kerling, R., & Van der Velden, E.** (1996). Nonverbal context effects in forward and backward translation: Evidence for concept mediation. *Journal of Memory and Language*, 35(5), 648–665. <https://doi.org/10.1006/jmla.1996.034>
- LaRusso, M., Donovan, S., & Snow, C.** (2016). Implementation challenges for tier one and tier two school-based programs for early adolescents. *New Directions for Child and Adolescent Development*, 154, 11–30. <https://doi.org/10.1002/cad.20179>
- Leacox, L., & Wood Jackson, C.** (2014). Spanish vocabulary-bridging technology enhanced instruction for young English language learners' word learning. *Journal of Early Childhood Literacy*, 14(2), 175–197. <https://doi.org/10.1177/1468798412458518>
- Levy, M.** (2009). WEB 2.0 implications on knowledge management. *Journal of Knowledge Management*, 13(1), 120–134. <https://doi.org/10.1108/13673270910931215>

- Liang, L. A., Peterson, C. A., & Graves, M. F. (2005). Investigating two approaches to fostering children's comprehension of literature. *Reading Psychology, 26*, 387–400. <https://doi.org/10.1080/02702710500285748>
- Lo, Y., Anderson, A., & Bunch-Crump, K. (2016). Building vocabulary of English learners with reading disabilities through computer-assisted morphology instruction. *Intervention in School and Clinic, 52*(3), 113–140. <https://doi.org/10.1177/1053451216644829>
- Lonigan, C. J. (2015). Early literacy. In L. Liben & U. Miller (Eds.), *Handbook of child psychology and developmental science. Volume 2: Cognitive processes*. Hoboken, NJ: Wiley.
- Lugo-Neris, M. J., Wood Jackson, C., & Goldstein, H. (2010). Facilitating vocabulary acquisition of young English language learners. *Language, Speech, and Hearing Services in Schools, 41*, 314–327. [https://doi.org/10.1044/0161-1461\(2009/07-0082\)](https://doi.org/10.1044/0161-1461(2009/07-0082))
- Nash, M., & Donaldson, M. L. (2005). Word learning in children with vocabulary deficits. *Journal of Speech, Language, and Hearing Research, 48*, 439–458. [https://doi.org/10.1044/1092-4388\(2005/030\)](https://doi.org/10.1044/1092-4388(2005/030))
- Nagy, W. E., Carlisle, J. F., & Goodwin, A. P. (2014). Morphological knowledge and literacy acquisition. *Journal of Learning Disabilities, 47*, 3–12. <https://doi.org/10.1177/002221943509967>
- National Center for Education Statistics. (2011). *The nation's report card: Reading 2011* (NCES 2012-457). Washington, DC: Author, Institute of Education Sciences, U.S. Department of Education. Retrieved from <http://www.nces.ed.gov>
- National Center for Education Statistics. (2016). *Table 204.27: English language learner (ELL) students enrolled in public elementary and secondary schools, by grade, home language, and selected student characteristics: Selected years, 2008–2009 through 2013–2014*. Retrieved from <http://www.nces.ed.gov>
- Neely, J. H. (1991). Semantic priming effects in visual word recognition: A selective review of current findings and theories. In D. Besner, & G. W. Humphreys (Eds.), *Basic processes in reading: Visual word recognition* (pp. 264–336). Hillsdale, NJ: Erlbaum.
- Ordonez, C. L., Carlo, M. S., Snow, C. E., & McLaughlin, B. (2002). Depth and breadth of vocabulary in two languages: Which vocabulary skills transfer? *Journal of Educational Psychology, 94*(4), 719–728. <https://doi.org/10.1037/0022-0663.94.4.719>
- Paul, R. (2007). *Language disorders from infancy through adolescence assessment and intervention*. New Haven, CT: Elsevier Mosby.
- Peña, E. D., Gutiérrez-Clellen, V. F., Iglesias, A., Goldstein, B. A., & Bedore, L. M. (2014). *BESA: Bilingual English–Spanish Assessment*. San Rafael, CA: AR-Clinical Publications.
- Petscher, Y. (2016). Do our means of inquiry match our intentions? *Frontiers in Psychology, 7*, 1048. <https://doi.org/10.3389/fpsyg.2016.01048>
- Proctor, C. P., August, D., Carlo, M. S., & Snow, C. (2006). The intriguing role of Spanish language vocabulary knowledge in predicting English reading comprehension. *Journal of Educational Psychology, 98*(1), 159–169. <https://doi.org/10.1037/0022-0663.98.1.159>
- R Core Team. (2017). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from <https://www.R-project.org/>
- Reardon, S. F., Kalogrides, D., & Shores, K. (2016). *The geography of racial/ethnic test score gaps* (CEPA Working Paper No. 16-10). Retrieved from <https://cepa.stanford.edu/sites/default/files/wp16-10-v201701.pdf>
- Restrepo, M. A., Morgan, G. P., & Thompson, M. S. (2013). The efficacy of a vocabulary intervention for dual-language learners with language impairment. *Journal of Speech, Language, and Hearing Research, 56*, 748–765. [https://doi.org/10.1044/1092-4388\(2012/11-0173\)](https://doi.org/10.1044/1092-4388(2012/11-0173))
- Roberts, T., & Neal, H. (2004). Relationships among preschool English language learners' oral proficiency in English, instructional experience and literacy development. *Contemporary Educational Psychology, 29*(3), 283–311. <https://doi.org/10.1016/j.cedpsych.2003.08.001>
- Scarborough, H. (2002). Young bilingual children and early literacy development. In S. Neuman & D. K. Dickinson (Eds.), *Handbook of early literacy research* (pp. 159–178). New York, NY: Guilford.
- Sénéchal, M., & LeFevre, J. (2002). Parental involvement in the development of children's reading skill: A five-year longitudinal study. *Child Development, 73*, 445–460. <https://doi.org/10.1111/1467-8624.00417>
- Silverman, R., & Hines, S. (2009). The effects of multimedia-enhanced instruction on the vocabulary of English-language learners and non-English-language learners in pre-kindergarten through second grade. *Journal of Educational Psychology, 101*(2), 305–314. <https://doi.org/10.1037/a0014217>
- Silverman, R. D., Proctor, C. P., Harring, J. R., Doyle, B., Mitchell, M. A., & Meyer, A. G. (2013). Teachers' instruction and students' vocabulary and comprehension: An exploratory study with English monolingual and Spanish–English bilingual students in grades 3–5. *Reading Research Quarterly, 49*, 31–60. <https://doi.org/10.1002/rrq.63>
- Slavin, R. E., & Cheung, A. (2005). A synthesis of research on language of reading instruction for English language learners. *Review of Educational Research, 75*(2), 247–284. <https://doi.org/10.3102/00346543075002247>
- Smith, L. B. (2000). Learning how to learn words: An associative crane. In R. Michnick Golinkoff et al. (Eds.), *Becoming a word learner: A debate on lexical acquisition* (pp. 51–80). New York, NY: Oxford University Press.
- Snow, C. E. (1990). The development of definitional skill. *Journal of Child Language, 17*(3), 697–710. <https://doi.org/10.1017/S0305000900010953>
- Sparks, E., & Deacon, H. (2015). Morphological awareness and vocabulary acquisition: A longitudinal examination of their relationship in English-speaking children. *Applied Psycholinguistics, 36*, 299–321. <https://doi.org/10.1017/S0142716413000246>
- Sunderman, G., & Kroll, J. F. (2006). First language activation during second language lexical processing: An investigation of lexical form, meaning, and grammatical class. *Studies in Second Language Acquisition, 28*, 387–422. <https://doi.org/10.1017/S0272263106060177>
- Tokowicz, N. (2015). *Lexical processing and second language acquisition*. New York, NY: Routledge.
- Uccelli, P., & Pérez, M. M. (2007). Narrative and vocabulary development of bilingual children from kindergarten to first grade: Developmental changes and associations among English and Spanish skills. *Language, Speech, and Hearing Services in Schools, 38*, 225–236. [https://doi.org/10.1044/0161-1461\(2007/024\)](https://doi.org/10.1044/0161-1461(2007/024))
- Van Heuven, W. J. B., Dijkstra, A., & Grainger, J. (1998). Orthographic neighborhood effects in bilingual word recognition. *Journal of Memory and Language, 39*(3), 458–483. <https://doi.org/10.1006/jmla.1998.2584>
- Verhallen, M. J., & Bus, A. G. (2012, July). *Beneficial effects of illustrations in picture storybooks for storing and retaining story text*. Paper presented at the annual meeting of the Society for the Scientific Studies of Reading, Montreal, Canada.
- Wanzek, J., Petscher, Y., Al Otaiba, S., Kent, S. C., Schatschneider, C., Haynes, M., . . . Jones, F. G. (2016). Examining the average and local effects of a standardized treatment for fourth graders

with reading difficulties. *Journal of Research on Educational Effectiveness*, 9, 45–66. <https://doi.org/10.1080/19345747.2015.1116032>

What Works Clearinghouse. (2014). *Assessing attrition bias addendum (version 3.0)*. Washington, DC: Author, Institute of Education Sciences, U.S. Department of Education. Retrieved from <https://ies.ed.gov>

White, E., & Gillard, S. (2011). Technology-based literacy instruction for English learners. *Teaching and Learning*, 8(6), 1–5.

Whitehurst, G. J., Falco, F. L., Lonigan, C. J., Fischel, J. E., DeBaryshe, B. D., Valdez-Menchaca, M. C., & Caulfield, M.

(1988). Accelerating language development through picture book reading. *Developmental Psychology*, 24(4), 552–559. <https://doi.org/10.1037/0012-1649.24.4.552>

Woodcock, R. W. (2011). *Woodcock Reading Mastery Tests—Third Edition (WRMT)*. San Antonio, TX: Pearson.

Wood Jackson, C., Schatschneider, C., & Leacox, L. (2014). Longitudinal analysis of receptive vocabulary growth in young Spanish English-speaking children from migrant families. *Language, Speech, and Hearing Services in Schools*, 45, 40–51. https://doi.org/10.1044/2013_LSHSS-12-0104

Appendix A (p. 1 of 2)

Quantile Regression Grade Moderation

Outcome	Quantile	Parameter	Coefficient	SE	LB	UB	p value	
PPVT-4 raw grade moderation	0.10	Intercept	27.27	5.70	15.83	38.72	< .001	
		Fall PPVT-4 raw	0.63	0.07	0.48	0.78	< .001	
		Treatment	0.31	1.27	-2.25	2.86	.881	
		Grade	0.94	2.83	-4.75	6.63	.741	
	0.25	Grade × Treatment	0.93	2.31	-3.72	5.58	.689	
		Intercept	27.70	4.54	18.58	36.82	< .001	
		Fall PPVT-4 raw	0.67	0.06	0.55	0.78	< .001	
		Treatment	0.25	1.12	-2.00	2.50	.824	
	0.50	Grade	1.59	1.73	-1.89	5.07	.363	
		Grade × Treatment	1.24	1.71	-2.29	4.69	.471	
		Intercept	30.84	3.67	23.47	38.22	< .001	
		Fall PPVT-4 raw	0.67	0.05	0.57	0.77	< .001	
	0.75	Treatment	1.32	1.81	-2.32	4.96	.469	
		Grade	1.71	1.85	-2.00	5.42	.358	
		Grade × Treatment	1.02	2.18	-3.37	5.40	.643	
		Intercept	35.81	5.07	25.61	46.00	< .001	
	0.90	Fall PPVT-4 raw	0.68	0.06	0.55	0.81	< .001	
		Treatment	0.15	1.66	-1.88	4.80	.383	
		Grade	0.27	1.70	-3.14	3.68	.874	
		Grade × Treatment	1.32	2.59	-3.88	6.53	.611	
	Definition grade moderation	0.10	Intercept	35.92	6.77	22.31	39.53	< .001
			Fall PPVT-4 raw	0.69	0.08	0.53	0.85	< .001
			Treatment	2.23	2.30	-2.29	6.85	.337
			Grade	3.54	2.40	-1.28	8.36	.147
0.25		Grade × Treatment	0.46	3.69	-6.95	7.87	.901	
		Intercept	7.55	4.17	-0.82	15.92	.076	
		Fall PPVT-4 raw	1.56	1.68	-1.81	4.93	.357	
		Treatment	0.12	0.05	0.02	0.21	.014	
0.50		Grade	-6.08	2.21	-10.53	-1.64	.008	
		Grade × Treatment	-0.42	2.33	-5.10	4.27	.859	
		Intercept	7.43	3.26	0.88	13.99	.027	
		Fall PPVT-4 raw	1.25	1.26	-1.29	3.79	.328	
0.75		Treatment	0.14	0.04	0.07	0.22	< .001	
		Grade	-5.55	2.13	-9.83	-1.27	.012	
		Grade × Treatment	-1.69	2.11	-5.93	2.54	.426	
		Intercept	8.40	3.72	0.92	15.87	.028	
0.90		Fall PPVT-4 raw	2.36	1.35	-0.35	5.08	.086	
		Treatment	0.16	0.04	0.08	0.25	< .001	
		Grade	-3.59	2.31	-8.23	1.06	.127	
		Grade × Treatment	-2.38	2.24	-6.89	2.14	.295	
0.10		Intercept	13.83	3.63	6.54	21.13	< .001	
		Fall PPVT-4 raw	2.55	1.37	-0.21	5.31	.069	
		Treatment	0.12	0.04	0.03	0.20	.008	
		Grade	-3.19	2.42	-8.06	1.68	.194	
0.25	Grade × Treatment	-3.41	2.73	-8.89	2.08	.218		
	Intercept	20.17	5.17	9.78	30.56	< .001		
	Fall PPVT-4 raw	1.25	1.60	-1.96	4.47	.438		
	Treatment	0.09	0.06	-0.02	0.21	.106		
0.50	Grade	-3.10	2.12	-7.35	1.15	.149		
	Grade × Treatment	-3.06	2.78	-8.65	2.53	.277		

Appendix A (p. 2 of 2)

Quantile Regression Grade Moderation

Outcome	Quantile	Parameter	Coefficient	SE	LB	UB	p value
Labeling probe grade moderation	0.10	Intercept	-0.15	1.30	-2.76	2.46	.907
		Fall PPVT-4 raw	1.00	0.09	0.83	1.17	< .001
		Treatment	3.15	1.38	0.38	5.92	.027
		Grade	0.15	1.95	-3.77	4.08	.938
	0.25	Grade × Treatment	0.85	1.87	-2.91	4.60	.652
		Intercept	2.00	0.91	0.18	3.82	.032
		Fall PPVT-4 raw	1.00	0.08	0.83	1.17	< .001
		Treatment	3.00	1.10	0.79	5.21	.009
	0.50	Grade	0.93	1.09	-1.25	3.11	.397
		Grade × Treatment	-0.93	1.38	-3.70	1.84	.504
		Intercept	3.65	0.78	2.08	5.23	< .001
		Fall PPVT-4 raw	1.13	0.06	1.00	1.26	< .001
	0.75	Treatment	3.78	0.99	1.80	5.76	< .001
		Grade	1.57	1.08	-0.60	3.73	.152
		Grade × Treatment	-2.00	1.21	-4.44	0.44	.106
		Intercept	5.43	1.27	2.88	7.98	< .001
	0.90	Fall PPVT-4 raw	1.14	0.09	0.96	1.32	< .001
		Treatment	5.00	1.38	2.23	7.77	< .001
		Grade	2.57	1.03	0.50	4.64	.016
		Grade × Treatment	-2.43	1.76	-5.96	1.10	.173
		Intercept	8.05	1.82	4.38	11.71	< .001
		Fall PPVT-4 raw	1.08	0.09	0.89	1.26	< .001
		Treatment	5.77	1.55	2.65	8.89	< .001
		Grade	1.77	1.87	-1.99	5.54	.348
		Grade × Treatment	-2.55	1.90	-6.36	1.27	.186

Note. PPVT-4 = Peabody Picture Vocabulary Test–Fourth Edition; LB = lower bound; UB = upper bound.

Appendix B

Rubric for Scoring Expressive Definitions

Points	Criterion	Example: <u>Chameleon</u>
0	<ul style="list-style-type: none"> No response Response of “I don’t know” or shrug of shoulders, only gestures the word Inappropriate definition Definitions of homophone 	<p>“Bug”</p> <p>“They eat lettuce”</p>
1	<ul style="list-style-type: none"> Mentions <i>only</i> features in the book (parts of the story) Vague, imprecise, or partial definition Example of word in context (phrase or sentence) but does not define meaning A description with / example of the word or item/person/object within word category An example of something it is not, or an antonym Mentions only one attribute of the target word 	<p>“Green”</p> <p>“Climb on branches”</p> <p>“Eat bugs”</p> <p>“Not an iguana, but it has a tail and looks like one”</p>
2	<ul style="list-style-type: none"> At least two different attributes of the word/item person/object within word category are listed and context in which the word is used Unambiguous synonym alone or used in context which defines meaning 	<p>“Chameleons are like lizards (category) that change colors (feature distinct to chameleon)”</p> <p>“The chameleon is climbing the tree outside”</p>
3	<ul style="list-style-type: none"> Complete and precise definition At least three or more descriptors Narrows possibility of confusing word with another word that is similar in meaning or shape 	<p>“Chameleons are animals (category) that can have a long tail and big eyes (size), they hide from people by changing colors (distinct feature) and sometimes can be a pet in a cage at a house (distinct feature).”</p>

Appendix C

Average Number of Sessions Per Week by School

School	Read-only	Intervention	<i>N</i>	Total
A	2.65 (0.15)	2.69 (0.21)	35	2.67 (0.18)
B	3.00 (0.04)	3.00 (0.05)	36	3.00 (0.05)
C	3.18 (0.32)	3.35 (0.18)	27	3.26 (0.27)
D		2.62 (0.68)	28	2.62 (0.68)
E	2.48 (0.10)	2.48 (0.19)	67	2.48 (0.15)
F	3.02 (0.28)	3.04 (0.24)	55	3.03 (0.26)
G	2.46 (0.68)	2.99 (0.27)	34	2.74 (0.57)
H	2.79 (0.24)	2.91 (0.03)	6	2.85 (0.17)
Total	2.77 (0.40)	2.82 (0.43)	288	2.80 (0.42)

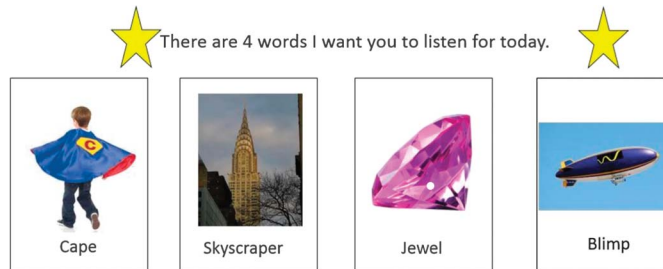
Average Number of Weeks for Treatment by School

School	Read-only	Intervention	<i>N</i>	Total
A	18.06 (1.66)	16.59 (2.60)	35	17.34 (2.26)
B	20.89 (0.32)	21.00 (0.00)	36	20.94 (0.23)
C	13.86 (4.04)	14.92 (0.28)	27	14.37 (2.91)
D		10.39 (2.92)	28	10.39 (2.92)
E	20.56 (0.50)	19.27 (4.08)	67	19.93 (2.94)
F	14.36 (2.26)	14.07 (2.32)	55	14.22 (2.27)
G	16.31 (6.01)	17.61 (5.01)	34	17.00 (5.46)
H	19.00 (0.00)	18.33 (0.58)	6	18.67 (0.52)
Total	17.69 (3.89)	16.83 (4.35)	288	16.10 (4.60)

Appendix D

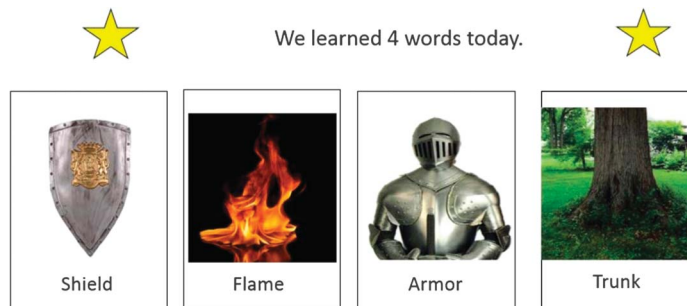
Instructional Components: Preview and Review

Preview



An example of the initial component of the e-book treatment condition providing a preview of target words in the story.

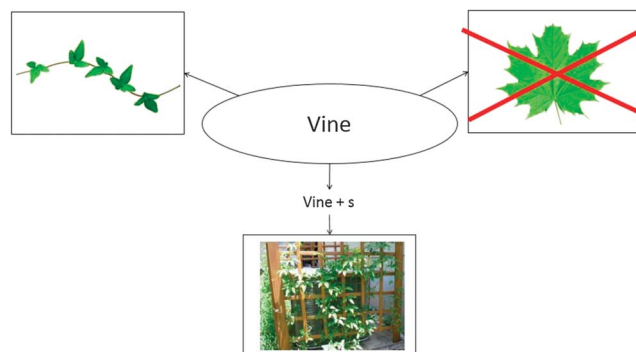
Review



An example of the final component of the e-book session, depicting a review of the four target words from the story.

Appendix E

Sample Word Map



Example word map of an e-book session that provided a check for understanding (select the appropriate picture), a nonexample, and a morphological component adding a word ending to the target word (e.g., vine + s = vines).

Appendix F (p. 1 of 3)

LQMM Moderation

Quantile regression baseline moderation

Outcome	Quantile	Parameter	Coefficient	SE	LB	UB	p value
PPVT-4 raw grade moderation	0.10	Intercept	27.27	5.70	15.83	38.72	< .001
		Fall PPVT-4 raw	0.63	0.07	0.48	0.78	< .001
		Treatment	0.31	1.27	-2.25	2.86	.881
		Grade	0.94	2.83	-4.75	6.63	.741
	0.25	Grade × Treatment	0.93	2.31	-3.72	5.58	.689
		Intercept	27.70	4.54	18.58	36.82	< .001
		Fall PPVT-4 raw	0.67	0.06	0.55	0.78	< .001
		Treatment	0.25	1.12	-2.00	2.50	.824
	0.50	Grade	1.59	1.73	-1.89	5.07	.363
		Grade × Treatment	1.24	1.71	-2.29	4.69	.471
		Intercept	30.84	3.67	23.47	38.22	< .001
		Fall PPVT-4 raw	0.67	0.05	0.57	0.77	< .001
	0.75	Treatment	1.32	1.81	-2.32	4.96	.469
		Grade	1.71	1.85	-2.00	5.42	.358
		Grade × Treatment	1.02	2.18	-3.37	5.40	.643
		Intercept	35.81	5.07	25.61	46.00	< .001
	0.90	Fall PPVT-4 raw	0.68	0.06	0.55	0.81	< .001
		Treatment	0.15	1.66	-1.88	4.80	.383
		Grade	0.27	1.70	-3.14	3.68	.874
		Grade × Treatment	1.32	2.59	-3.88	6.53	.611
Definition grade moderation	0.10	Intercept	35.92	6.77	22.31	39.53	< .001
		Fall PPVT-4 raw	0.69	0.08	0.53	0.85	< .001
		Treatment	2.23	2.30	-2.29	6.85	.337
		Grade	3.54	2.40	-1.28	8.36	.147
	0.25	Grade × Treatment	0.46	3.69	-6.95	7.87	.901
		Intercept	7.55	4.17	-0.82	15.92	.076
		Fall PPVT-4 raw	1.56	1.68	-1.81	4.93	.357
		Treatment	0.12	0.05	0.02	0.21	.014
	0.50	Grade	-6.08	2.21	-10.53	-1.64	.008
		Grade × Treatment	-0.42	2.33	-5.10	4.27	.859
		Intercept	7.43	3.26	0.88	13.99	.027
		Fall PPVT-4 raw	1.25	1.26	-1.29	3.79	.328
	0.75	Treatment	0.14	0.04	0.07	0.22	< .001
		Grade	-5.55	2.13	-9.83	-1.27	.012
		Grade × Treatment	-1.69	2.11	-5.93	2.54	.426
		Intercept	8.40	3.72	0.92	15.87	.028
	0.90	Fall PPVT-4 raw	2.36	1.35	-0.35	5.08	.086
		Treatment	0.16	0.04	0.08	0.25	< .001
		Grade	-3.59	2.31	-8.23	1.06	.127
		Grade × Treatment	-2.38	2.24	-6.89	2.14	.295
0.10	Intercept	13.83	3.63	6.54	21.13	< .001	
	Fall PPVT-4 raw	2.55	1.37	-0.21	5.31	.069	
	Treatment	0.12	0.04	0.03	0.20	.008	
	Grade	-3.19	2.42	-8.06	1.68	.194	
0.25	Grade × Treatment	-3.41	2.73	-8.89	2.08	.218	
	Intercept	20.17	5.17	9.78	30.56	< .001	
	Fall PPVT-4 raw	1.25	1.60	-1.96	4.47	.438	
	Treatment	0.09	0.06	-0.02	0.21	.106	
0.50	Grade	-3.10	2.12	-7.35	1.15	.149	
	Grade × Treatment	-3.06	2.78	-8.65	2.53	.277	

Appendix F (p. 2 of 3)

LQMM Moderation

Quantile regression baseline moderation

Outcome	Quantile	Parameter	Coefficient	SE	LB	UB	p value
Labeling probe grade moderation	0.10	Intercept	-0.15	1.30	-2.76	2.46	.907
		Fall PPVT-4 raw	1.00	0.09	0.83	1.17	< .001
		Treatment	3.15	1.38	0.38	5.92	.027
		Grade	0.15	1.95	-3.77	4.08	.938
	0.25	Grade × Treatment	0.85	1.87	-2.91	4.60	.652
		Intercept	2.00	0.91	0.18	3.82	.032
		Fall PPVT-4 raw	1.00	0.08	0.83	1.17	< .001
		Treatment	3.00	1.10	0.79	5.21	.009
	0.50	Grade	0.93	1.09	-1.25	3.11	.397
		Grade × Treatment	-0.93	1.38	-3.70	1.84	.504
		Intercept	3.65	0.78	2.08	5.23	< .001
		Fall PPVT-4 raw	1.13	0.06	1.00	1.26	< .001
	0.75	Treatment	3.78	0.99	1.80	5.76	< .001
		Grade	1.57	1.08	-0.60	3.73	.152
		Grade × Treatment	-2.00	1.21	-4.44	0.44	.106
		Intercept	5.43	1.27	2.88	7.98	< .001
	0.90	Fall PPVT-4 raw	1.14	0.09	0.96	1.32	< .001
		Treatment	5.00	1.38	2.23	7.77	< .001
		Grade	2.57	1.03	0.50	4.64	.016
		Grade × Treatment	-2.43	1.76	-5.96	1.10	.173
PPVT-4 raw baseline moderation	0.10	Intercept	8.05	1.82	4.38	11.71	< .001
		Fall PPVT-4 raw	1.08	0.09	0.89	1.26	< .001
		Treatment	5.77	1.55	2.65	8.89	< .001
		Grade	1.77	1.87	-1.99	5.54	.348
	0.25	Grade × Treatment	-2.55	1.90	-6.36	1.27	.186
		Intercept	26.38	10.38	5.53	47.24	.014
		Fall PPVT-4 raw	0.65	0.13	0.38	0.92	< .001
		Treatment	-1.05	11.53	-24.23	22.13	.923
	0.50	Fall PPVT-4 × Treatment	0.01	0.15	-0.28	0.30	.94
		Intercept	25.54	7.01	11.45	39.63	< .001
		Fall PPVT-4 raw	0.70	0.09	0.52	0.87	< .001
		Treatment	8.19	7.59	-7.06	23.44	.286
	0.75	Fall PPVT-4 × Treatment	-0.09	0.09	-0.28	0.10	.338
		Intercept	27.24	5.48	16.24	38.24	< .001
		Fall PPVT-4 raw	0.72	0.07	0.59	0.85	< .001
		Treatment	11.50	6.64	-1.84	24.85	.089
	0.90	Fall PPVT-4 × Treatment	-0.12	0.08	-0.29	0.05	.158
		Intercept	37.75	6.14	25.41	50.10	< .001
		Fall PPVT-4 raw	0.65	0.08	0.50	0.81	< .001
		Treatment	0.47	6.62	-12.83	13.77	.943
	Fall PPVT-4 × Treatment	0.02	0.08	-0.15	0.19	.803	
	Intercept	46.49	8.52	29.37	63.61	< .001	
	Fall PPVT-4 raw	0.59	0.10	0.38	0.80	< .001	
	Treatment	-4.88	6.94	-18.82	9.06	.485	
		Fall PPVT-4 × Treatment	0.11	0.08	-0.06	0.28	.192

Appendix F (p. 3 of 3)

LQMM Moderation

Quantile regression baseline moderation

Outcome	Quantile	Parameter	Coefficient	SE	LB	UB	p value
Definition baseline moderation	0.10	Intercept	8.81	4.23	0.30	17.32	.043
		Fall PPVT-4 raw	-10.51	6.08	-22.71	1.70	.090
		Treatment	0.09	0.05	-0.01	0.19	.064
	0.25	Fall PPVT-4 × Treatment	0.12	0.07	-0.02	0.26	.084
		Intercept	8.34	3.60	1.10	15.57	.025
		Fall PPVT-4 raw	-7.50	5.37	-18.30	3.30	.169
	0.50	Treatment	0.11	0.04	0.04	0.18	.004
		Fall PPVT-4 × Treatment	0.10	0.06	-0.03	0.22	.135
		Intercept	11.76	4.70	2.32	21.21	.016
	0.75	Fall PPVT-4 raw	-11.57	6.24	-24.12	0.97	.070
		Treatment	0.09	0.05	-0.01	0.20	.086
		Fall PPVT-4 × Treatment	0.15	0.07	0.01	0.28	.033
	0.90	Intercept	18.36	5.23	7.85	28.87	< .001
		Fall PPVT-4 raw	-10.21	6.05	-22.36	1.94	.098
		Treatment	0.05	0.07	-0.08	0.18	.434
Labeling probe baseline moderation	0.10	Fall PPVT-4 × Treatment	0.13	0.07	0.00	0.27	.057
		Intercept	16.83	6.48	3.82	29.85	.012
		Fall PPVT-4 raw	-4.65	6.23	-17.17	7.88	.460
	0.25	Treatment	0.14	0.07	-0.01	0.28	.070
		Fall PPVT-4 × Treatment	0.05	0.07	-0.08	0.19	.431
		Intercept	0.003	1.22	-2.44	2.45	.998
	0.50	Fall PPVT-4 raw	0.10	0.11	0.78	1.22	< .001
		Treatment	2.96	1.30	0.36	5.57	.027
		Fall PPVT-4 × Treatment	0.03	0.15	-0.27	0.33	.829
	0.75	Intercept	2.62	1.12	0.38	4.87	.023
		Fall PPVT-4 raw	1.00	0.11	0.77	1.23	< .001
		Treatment	1.80	1.61	-1.44	5.03	.270
	0.90	Fall PPVT-4 × Treatment	0.11	0.16	-0.22	0.44	.504
		Intercept	5.50	1.16	3.16	7.84	< .001
		Fall PPVT-4 raw	1.00	0.10	0.80	1.20	< .001
0.25	Treatment	1.50	1.39	-1.29	4.29	.285	
	Fall PPVT-4 × Treatment	0.17	0.16	-0.15	0.48	.288	
	Intercept	8.00	1.14	5.72	10.28	< .001	
0.50	Fall PPVT-4 raw	1.00	0.07	0.86	1.14	< .001	
	Treatment	1.00	1.23	-1.46	3.46	.419	
	Fall PPVT-4 × Treatment	0.28	0.13	0.01	0.54	.041	
0.75	Intercept	9.73	1.51	6.70	12.76	< .001	
	Fall PPVT-4 raw	1.09	0.12	0.84	1.34	< .001	
	Treatment	2.55	1.88	-1.23	6.32	.181	
0.90	Fall PPVT-4 × Treatment	0.09	0.13	-0.17	0.36	.493	

Note. LQMM = linear quantile mixed model; PPVT-4 = Peabody Picture Vocabulary Test–Fourth Edition; LB = lower bound; UB = upper bound.