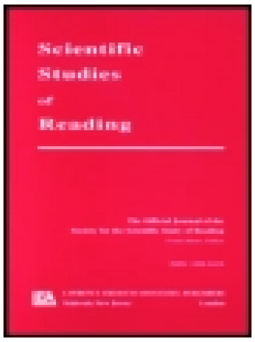

Routledge
Taylor & Francis Group

Scientific Studies of Reading

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/hssr20

Effects of Integrated Literacy and Content-area Instruction on Vocabulary and Comprehension in the Elementary Years: A Meta-analysis

HyeJin Hwang, Sonia Q. Cabell & Rachel E. Joyner

To cite this article: HyeJin Hwang, Sonia Q. Cabell & Rachel E. Joyner (2021): Effects of Integrated Literacy and Content-area Instruction on Vocabulary and Comprehension in the Elementary Years: A Meta-analysis, Scientific Studies of Reading, DOI: 10.1080/10888438.2021.1954005

To link to this article: https://doi.org/10.1080/10888438.2021.1954005

SCIENTIFIC STUDIES OF READING
https://doi.org/10.1080/10888438.2021.1954005

Effects of Integrated Literacy and Content-area Instruction on Vocabulary and Comprehension in the Elementary Years: A Meta-analysis

HyeJin Hwang [a], Sonia Q. Cabell [b], and Rachel E. Joyner[c]

[a]Department of Educational Psychology, College of Education and Human Development, University of Minnesota, Minneapolis, MN, USA; [b]School of Teacher Education and Florida Center for Reading Research, Florida State University, Tallahassee, FL, USA; [c]Florida State University, Tallahassee, FL, USA

ABSTRACT

This study synthesized impacts of integrated literacy and content-area instruction (i.e., science, social studies) on vocabulary and comprehension outcomes in the elementary years (i.e., kindergarten through fifth grade). A systematic search of the extant literature identified 35 (quasi)experimental studies. Random-effects models were used to combine effect sizes across studies. Results of meta-analysis revealed that the overall effects were positive and significant for vocabulary (effect size [ES] = 0.91) and comprehension (ES = 0.40). Moreover, a significant positive effect was observed for standardized comprehension outcomes (ES = 0.25), but not for standardized vocabulary outcomes. Supplementary analysis including studies with content knowledge outcomes demonstrated the positive and significant overall effect for content knowledge (ES = 0.89). In addition, no significant moderators of the effect sizes were found among features of research design and characteristics of interventions, perhaps partly due to the small number of studies. The results of our meta-analysis indicate that integrated literacy and content-area instruction has potential to enhance vocabulary words taught to students and comprehension in the elementary years, with the additional benefit of simultaneously cultivating science and social studies knowledge.

ARTICLE HISTORY
Received 18 September 2020
Revised 09 June 2021
Accepted 14 June 2021

Vocabulary and comprehension are foundational in academic achievement (Neuman & Wright, 2014; van den Broek et al., 2005), particularly since successful written and oral communication skills are in-demand 21st century competencies for career development (Rios, Ling, Pugh, Becker, & Bacall, 2020). However, a large gap in vocabulary has been reported among elementary-aged students (Marulis & Neuman, 2010). Biemiller and Slonim (2001) documented an approximately 4,000-word gap between students in second through fifth grade with high vocabulary knowledge (highest quartile in normative sample) and students with low vocabulary knowledge (lowest quartile). Students with limited vocabulary can encounter difficulties in comprehension because vocabulary is an essential component of comprehension (Wright & Cervetti, 2016). Moreover, many students display difficulties in comprehension in the elementary years, with only one third of fourth-grade students in the U.S. demonstrating a proficient level of reading comprehension (e.g., integrating information in text) and another third performing below the basic level (e.g., identifying information; National Assessment for Educational Progress [NAEP], 2019).

Recognizing the difficulties that students experience in developing vocabulary and comprehension, and the resulting cumulative disadvantage in academic and career success, consensus documents have

CONTACT HyeJin Hwang hwang305@umn.edu; Department of Educational Psychology, University of Minnesota, 56 East River Rd, 151 Education Sciences Building, Minneapolis, MN 55455
Supplemental data for this article can be accessed on the publisher's website.

sought to identify effective instructional supports for vocabulary and comprehension in the elementary years (e.g., National Governors Association Center for Best Practices & Council of Chief State School Officers, 2010; National Institute for Literacy, 2008; National Institute of Child Health and Human Development, 2000; RAND Reading Study Group, 2002). These consensus documents have mostly focused on direct teaching of vocabulary and comprehension. For example, suggestions from the What Works Clearinghouse practice guides include explicit instruction on academic vocabulary (Foorman et al., 2016) and comprehension strategies and text structures (Shanahan et al., 2010). Yet, there remains a gap in the field's knowledge. Relatively little attention has been devoted to marshaling the empirical evidence for the potential benefits of simultaneously focusing on literacy and content knowledge (i.e., knowledge related to the natural and social world) by integrating literacy and content-area instruction (i.e., science and social studies) in supporting vocabulary and comprehension (Cabell & Hwang, 2020). This meta-analysis seeks to address this gap by examining the effects of integrated instruction on comprehension and vocabulary development in the elementary years (kindergarten to Grade 5).

Why might integrated literacy and content-area instruction support vocabulary and comprehension?

Characteristics of integrated literacy and content-area instruction

The literature contains varied definitions of integrated models of instruction without a clear consensus (Huck, 2019; Huntley, 1998; Thibaut, Knipprath, Dehaene, & Depaepe, 2018). For the purpose of this study, we define *integrated instruction* between literacy and the content areas as instruction in which literacy activities (reading and/or writing) serve as a tool to cultivate content knowledge (science and/or social studies) while, at the same time, content teaching serves as a lever to facilitate literacy skills (vocabulary and/or comprehension). This definition for the current study reflects conceptions offered by Dickinson and Young (1998), Huntley (1998), and Stoddart, Pinal, Latzke, and Canaday (2002) that the integrated approach blends ideas and skills in different subjects with balanced attention to each subject for the purpose of attaining the instructional goals for each subject, rather than focusing on one subject with the other subject only used as a tool.

Literacy activities can play a facilitative role in supporting content learning[1] because by interacting with texts, students can learn information about nature and society as well as how the information has been established by experts (Cervetti & Barber, 2008). Students can also learn how to interpret social and natural phenomenon and use their knowledge for problem solving by consulting texts (Cervetti & Barber, 2008). Moreover, engaging in discussion and writing about what they learned from text, students can learn how professionals communicate with one another to refine their understanding of nature and the social world (Mercer, Dawes, Wegerif, & Sams, 2004). Content teaching, on the other hand, can contribute to fostering literacy skills. Emphasis on information about social and natural phenomenon in content teaching can support students' comprehension because students can use the information to construct coherent mental representations of texts they read or that are read aloud to them (McCarthy & McNamara, 2021). Specifically, content knowledge in science and social studies can support processes for generating inferences about missing information in texts (Ozuru, Dempsey, & McNamara, 2009) and guides how different ideas within and across texts need to be combined with one another (Kintsch, 1998, 2013). Students can also develop a deeper interest in reading and writing in the context of content teaching as they view reading and writing as a means to grow their expertise in a given content area (Guthrie & Wigfield, 2000). Moreover, content teaching can support students' vocabulary learning. Words are often considered labels for content knowledge, and thus, in the context of content teaching, students can meaningfully learn words related to content (Pearson & Billman, 2016).

In order to leverage the potential of content teaching and literacy activities in supporting each other, the integrated approach is guided by the logical organization of science or social studies content

(Romance & Vitale, 2012b). Connected ideas from science or social studies content are presented to students explicitly and cumulatively within and across class sessions. Thus, the sequence of texts to read (or listen to) would reflect the connections among content ideas being studied (e.g., a text explaining the process of pollination followed by a text about animals assisting with plant pollination; Lupo, Berry, Thacker, Sawyer, & Merritt, 2019). Students regularly interact with content-related texts (i.e., multiple texts coherently connected to one another around science or social studies content; Lupo et al., 2019; Moss, 2005) and progress toward developing an in-depth understanding of content by leveraging what they learned previously from text (e.g., plants need to be pollinated) and integrating it with new information (e.g., how bees, birds, and bats help to pollinate plants). Moreover, in integrated literacy and content-area instruction, words related to content appear repeatedly within and across class sessions (e.g., Neuman & Dwyer, 2011). Students encounter and use content-related words repeatedly (e.g., pollen, flower, bees) while engaging in reading or writing in order to learn about the content (e.g., plants; Elleman & Compton, 2017). Meanings of words are often taught in relation to content, including the association among words around content, rather than in isolation (Barber & Cervetti, 2019).

In contrast, in traditional literacy instructional design, the major consideration is deciding which literacy skills to teach (Hwang, Lupo, Cabell, & Wang, 2021). Texts and words are often selected independently from content and, as a result, are often disconnected from one another (e.g., reading about alpacas one day and galaxies the next) or lack sufficient depth in any particular topic. Traditional literacy instruction also tends to focus on activation of knowledge students already have in relation to a topic of a text, rather than systematically building content knowledge, and individual words are often taught without explicitly relating them to other words within content (Cervetti & Wright, 2020; Hwang et al., 2021).

Content-related texts and words in support of vocabulary and comprehension

Content-related texts (e.g., multiple texts coherently related to animal habitats) may support knowledge of content-related words (e.g., *burrow* and *den* in relation to animal habitats) because, while interacting with content-related texts, students are likely to encounter content-related words repeatedly, a facilitative condition to enhance vocabulary (August & Hakuta, 1997; Eller, Pappas, & Brown, 1988). When content-related words are explicitly taught, semantic networks among words and ideas are likely to be strengthened in long-term memory because students are provided with content knowledge to explain the reasons for these relations (Pritchard, 2019). This strengthened semantic network, in turn, can bolster incidental learning of new words because it can make retrieval and use of words more efficient (Steyvers & Tenenbaum, 2005). This efficient recall frees up cognitive resources to figure out the meanings of unknown words, resulting in increased incidental learning of new words (Pulido, 2007).

Comprehension involves a process of meaning construction in which readers or listeners integrate what the text says with what they already know (Kintsch, 1998). Opportunities to encounter texts on similar topics are vital, as meaning construction from new, but related, texts is likely to depend on students' ability to leverage what they have already learned from previous texts. By providing opportunities for an increased number of experiences with content-related texts (e.g., multiple texts coherently related to animal habitats), students may develop the ability to recognize meaningful relations (i.e., relational reasoning; Alexander & The Disciplined Reading and Learning Research Laboratory [DRLRL], 2012) between information they already know (e.g., bear cubs are safer in a den) and new information (e.g., joeys are safer in their mother kangaroo's pouch), which facilitates comprehension (Kendeou, Butterfuss, Van Boekel, & O'Brien, 2017).

In addition, content knowledge, developed by interacting with content-related texts, can facilitate the process of generating inferences about missing information (Cervetti & Wright, 2020) as well as the process of identifying important ideas in the text (Stahl, Hare, Sinatra, & Gregory, 1991). Moreover, in-depth content knowledge, which includes understandings of explanatory principles for phenomena and situations (e.g., animals have shelters to protect their babies; Alexander, Jetton, & Kulikowich, 1995), can facilitate comprehension by supporting efficient use of that content knowledge (Burgoon, Henderson, & Markman, 2013).

Integrated literacy and content-area instruction as motivating contexts to learn

Integrated literacy and content-area instruction can provide motivating contexts for students to engage in reading and writing. Crucially, literacy activities are not implemented in a vacuum, but rather in connection with the meaningful purpose of learning content (Pearson & Billman, 2016). Students are likely to be more interested in interacting with texts when the text is essential for learning science or social studies content (e.g., Guthrie et al., 1996). Use of content-related texts in integrated literacy and content-area instruction can support students' motivation because they can clearly see the relevance of new texts to their previous learning and can recognize progress in their learning (Meece, Anderman, & Anderman, 2006). When students are motivated to engage in literacy activities, they tend to increase their interactions with texts (Wang & Guthrie, 2004), which, in turn, can lead to better vocabulary knowledge (Swanborn & De Glopper, 1999) and comprehension (Guthrie, Wigfield, & You, 2012).

Past reviews

Many syntheses have focused on supporting vocabulary and comprehension in students with or without learning difficulties/disabilities or reading difficulties/disabilities in literacy or content-area instruction (see Figure S1 in Supplementary Material for examples of previous syntheses). However, scholars have published only a handful of reviews exploring integrated literacy and content-area instruction, which aims to support both literacy skills (vocabulary, comprehension) and content knowledge. Most of these have been chapters in literacy-related handbooks that have included sections summarizing the potential of the integrated approach for improving literacy development, including vocabulary and comprehension. As the focus of these reviews was to explicate the landscape of the literature on the role of prior knowledge in reading development (e.g., Cervetti & Wright, 2020) or to illustrate promising practices implemented in classrooms (e.g., Cervetti, 2013; Halvorsen, Alleman, & Brugar, 2013), these reviews did not systematically search and examine intervention studies that tested the effects of integrated literacy and content-area instruction on vocabulary and comprehension, nor did they provide estimates of overall effect sizes via meta-analysis.

Other reviews have focused on the effects of integrated literacy and science instruction on vocabulary or reading comprehension in the elementary years. For example, taking a closer look at younger learners, Guo, Wang, Hall, Breit-Smith, and Busch (2016) conducted a meta-analysis of seven studies on the effects of integrated literacy and science instruction on vocabulary outcomes in pre-kindergarten and kindergarten children. The results of their meta-analysis indicated that integrated literacy and science instruction enhanced vocabulary in pre-kindergarten and kindergarten children (Hedges' $g = 0.66$). Romance and Vitale (2012a) included seven studies in a review that examined the effects of the *Science IDEAS* curriculum on reading comprehension. Guthrie, McRae, and Klauda (2007) summarized 11 studies investigating *Concept-Oriented Reading Instruction* and its effects on reading comprehension. Both of these reviews revealed that treatment-group students receiving integrated literacy and science instruction consistently displayed better reading comprehension outcomes compared to control-group students.[2] Despite similar conclusions to the handbook chapters (e.g., Cervetti & Wright, 2020; Halvorsen et al., 2013), these reviews did not summarize the effects of integrated literacy and science instruction on both comprehension and vocabulary. Also, these reviews focused exclusively on projects integrating literacy and science, so integrated instruction between literacy and social studies was not included.

In a broader review, Hartzler (2000) conducted a meta-analysis to examine integrated models of instruction between/among different subject areas (e.g., literacy, science, social studies, mathematics, music) for K-12 and college students. Overall, the results of the meta-analysis of 19 studies indicated that integrated models of instruction between/among literacy and other subject areas enhanced literacy achievement (e.g., grammar, writing, reading), compared to separate literacy and subject-area instruction (average effect size was 0.42, based on Glass' formula [1976]). Despite including a comprehensive review of the literature, the findings of Hartzler's meta-analysis cannot answer whether integrated literacy and content-area instruction (science and social studies) supports vocabulary and comprehension, as the effect size for integrated literacy and content-area instruction cannot be disentangled from those for other subject integrations (e.g., integration among music, literacy, mathematics, science), and the effect size was calculated combining all literacy measures (without separate analyses for vocabulary or comprehension).

In summary, to our knowledge, no scholars have conducted a meta-analysis on the effects of integrated literacy and content-area instruction (science and social studies) on vocabulary and comprehension in the elementary years. Meta-analytic studies on this topic are needed, given persistent struggles with vocabulary and comprehension achievement among elementary-aged students as measured by various state- and national-level assessments (e.g., Biemiller & Slonim, 2001; NAEP, 2019) and the importance of supporting vocabulary and comprehension from early grades (Duke & Carlisle, 2011). Moreover, extant reviews have not evaluated features of research designs (e.g., experimental vs. quasi-experimental) or characteristics of interventions (e.g., length of the intervention). Meta-analyses considering these factors can help clarify our understanding of the current state of evidence on the potential of integrated literacy and content-area instruction in enhancing vocabulary and comprehension as well as informing research, policy, and practice.

Current meta-analysis

We conducted a meta-analysis to investigate the effects of integrated literacy and content-area instruction on vocabulary and comprehension in the elementary years. We also examined whether variability in the effects on vocabulary and comprehension are associated with features of research designs or characteristics of interventions. We asked the following research questions: 1) Does integrated literacy and content-area instruction enhance vocabulary and comprehension?, 2) Are features of research designs associated with the effect sizes for vocabulary and comprehension?, and 3) Are characteristics of interventions associated with the effect sizes for vocabulary and comprehension?

Regarding the first research question, we anticipated that integrated literacy and content-area instruction would yield positive impacts on vocabulary and comprehension, consistent with related extant reviews (e.g., Guo et al., 2016). With respect to the second research question, we anticipated significant relations between research design features and effects, based on a study by Cheung and Slavin (2016) that demonstrated the potential for research design features to influence study outcomes. For example, they found effects tended to be smaller when studies have higher quality research designs, and standardized measures tended to have smaller effect sizes than researcher-developed measures. For the third research question, we hypothesized that there would be significant relations between characteristics of interventions and the effects on vocabulary and comprehension. For example, students may benefit more from an intervention if it involves a variety of literacy activities (e.g., reading and writing vs. reading only; Graham et al., 2018), if comprehension strategies are taught to support students to leverage what they already know to comprehend new texts (Allen & McNamara, 2020) or if hands-on activities are provided as a component of the integrated instruction to stimulate students' interest to learn (e.g., Guthrie et al., 2006).

Method

Search strategies and criteria for inclusion/exclusion

Relevant studies were searched through a series of electronic searches, with the final search on March 31, 2020. We used three databases: ERIC and PsychInfo, to find studies published in peer-reviewed journals, as well as ProQuest Dissertations & Theses Global, to find dissertation studies. We utilized dissertation studies as gray literature to address publication bias, as they often include detailed information related to research design features and characteristics of interventions needed to answer the second and third research questions, respectively. For each electronic database, studies were searched three times with different sets of search terms (Table 1). After studies were retrieved via electronic searches, we applied inclusion and exclusion criteria (Table 2).

Table 1. List of search terms.

Search terms related to comprehension and knowledge	("reading comprehension" OR "passage comprehension" OR "reading test" OR "reading measure" OR "reading skills" OR "narrative comprehension" OR "narrative skills" OR "listening comprehension" OR "language comprehension" OR "oral comprehension" OR summar* OR recall OR retell OR "sentence verification test") AND ("prior knowledge" OR "background knowledge" OR "topic knowledge" OR "domain knowledge" OR "disciplinary knowledge" OR "cultural knowledge" OR "general knowledge" OR "world knowledge" OR "conceptual knowledge" OR "content knowledge" OR "social studies" OR science OR history OR economics OR geography OR civics OR chemistry OR biolog* OR "earth science" OR "space science" OR physics OR astronomy OR "academic knowledge")
Search terms related to vocabulary and knowledge	(vocabular* OR "semantic knowledge" OR "word knowledge") AND ("prior knowledge" OR "background knowledge" OR "topic knowledge" OR "domain knowledge" OR "disciplinary knowledge" OR "cultural knowledge" OR "general knowledge" OR "world knowledge" OR "conceptual knowledge" OR "content knowledge" OR "social studies" OR science OR history OR economics OR geography OR civics OR chemistry OR biolog* OR "earth science" OR "space science" OR physics OR astronomy OR "academic knowledge")
Search terms related to integrated instruction	"curricul* integration" OR "integrat* approach*" OR "integrat* curricul*" OR "content area literacy" OR "content literacy" OR "content area reading" OR "disciplinary literacy" OR "integrat* instruction"

Table 2. Inclusion and Exclusion Criteria.

Inclusion criteria	(1) reported in English, regardless of where the study took place, (2) appeared in peer-reviewed journals or dissertations (by the end of March, 2020 for papers retrieved from the electronic searches), (3) involved students in kindergarten through fifth grade, (4) employed an (quasi)experimental design with more than one condition, (5) included either comprehension or vocabulary outcome measures, and (6) compared integrated literacy and content-area instruction (science or social studies) with separate literacy and content-area instruction.[a]
Exclusion criteria	(1) the treatment-group and control-group students received integrated literacy and content-area instruction, while different support was added to both groups or extra support was provided to treatment-group students only, (2) there was no available information about literacy activities (reading, writing) in integrated literacy and content-area instruction, (3) necessary statistics for meta-analysis were not provided in papers, and this information could not be obtained from authors, and (4) the number of students in either treatment or control group was smaller than 10.

To be eligible for the meta-analysis, each study had to meet 6 inclusion criteria and could not meet any of 4 exclusion criteria.

[a]For the purpose of the meta-analysis, integrated literacy and content-area (i.e., science and social studies) instruction was defined as instruction in which literacy activities are used to foster content knowledge, while content teaching aims to develop literacy skills. In each study, integrated instruction was clearly operationalized by each set of authors in the description of the study. Specifically, interventions were regarded as integrated literacy and content-area instruction, if authors stated that: (1) their study examined integrated literacy and content-area instruction, (2) the intervention was designed based on the mutually enhancing relation between literacy and content-area instruction, (3) the intervention was aligned to content-area (science and/or social studies) standards (or guides) and literacy (or English Language Arts) standards (or guides), and/or (4) the intervention was aligned to content-area standards (or guides), while the intervention regularly engaged students in reading to learn the content from the standards (or guides).

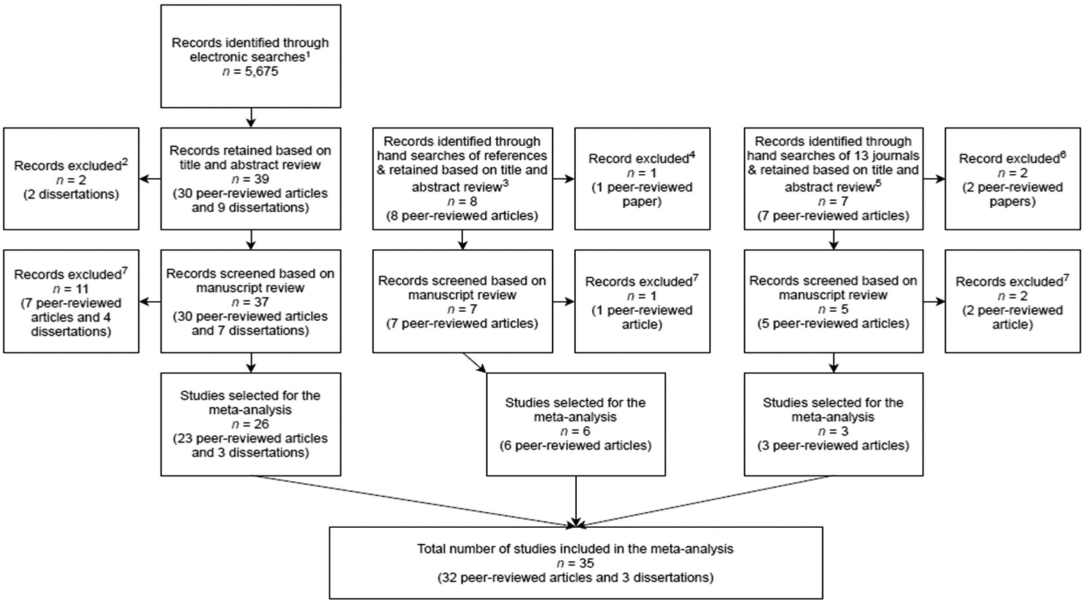

Figure 1. PRISMA flow chart of study selection process. [1]The number of records identified through ERIC and PsychInfo was 3,274, and the number of Records identified through ProQuest Dissertations & Theses Global was 2,401. The first and the third author, along with two doctoral students in the Curriculum and Instruction program (School of Teacher Education) at Florida State University read the title and abstract of each retrieved study and coded them independently by applying the inclusion criteria. The first author trained the coders (third author and two doctoral students) by explaining the inclusion criteria and modeling coding. Then the retrieved studies were double-coded by the first author and each of the three coders. The coding agreement was calculated by dividing the number of agreed coding by the number of the retrieved studies. The coding agreement for the initial selection was 98% between the first author and each of the rest of the coders. Discrepancies in coding were discussed and resolved among the two authors and the two doctoral students. When it was difficult to resolve discrepancies by reading the title and abstract only, papers were retrieved and examined more closely; [2]two dissertations were removed (Kao, 2015; Martínez, 2008) because they used the same data from the same interventions reported in two of the peer-reviewed papers already included in the set of 35 (confirmed by the authors of the dissertations); [3]we examined studies of which references were included in the articles identified from electronic searches and selected for the current review. In addition, studies reviewed by previous systematic or meta-analytic reviews (see *Past Reviews*) were examined as well; [4]Guthrie et al. (1996) did not have a control group; [5]We selected 13 journals based on publication information of previous studies examined for the meta-analysis (authors and journal titles) or the recognized importance in the field of reading research. The selected journals included *American Educational Research Journal, Educational Psychology Review, Elementary School Journal, International Journal of Science and Mathematics Education, International Research in Geographical and Environmental Education, Journal of Educational Psychology, Journal of Educational Research, Journal of Learning Disabilities, Journal of Research in Science Teaching, Reading and Writing: An Interdisciplinary Journal, Reading Research Quarterly, Scientific Studies of Reading*, and *Theory and Research in Social Education*. Publications of the 13 journals from January 1, 1990 and April 13, 2021 were examined. [6]Maerten-Rivera et al. (2016) did not include comprehension and vocabulary outcome measures. Brunner and Abd-El-Khalick (2020) compared different types of read-alouds in science instruction and did not include comprehension and vocabulary outcome measures. [7]see Table S1 for the reasoning for excluding 14 studies from the analysis.

Selection of studies for review

Figure 1 describes a PRISMA flow chart of selection process of studies for the review, and Table S1 in Supplementary Material lists the reasoning for excluding studies from the review based on the exclusion criteria. We arrived at 35 studies to include in the review (32 peer-reviewed articles and 3 dissertations).

Coding procedure

Each study was double-coded for features of research design and characteristics of interventions (see Table S2 in Supplementary Material) by different pairs among the three authors, and the interrater agreement was 95.6%.[3] Discrepancies in coding were resolved among the authors through discussion.

Table 3. Descriptive Information about Quality of Research Design and Features of Implementation of the Interventions.

Study: First author	Year	Report type	Rating of research design	Adherence	Instructors	PD	Content comparison
Aarnoutse	2003	PA	1: QE with baseline equivalence satisfied with covariate adjustment[1]	Poor	T	Y	0
Bravo	2014	PA	2: E with low attrition	Good w/R	T	N	2
Cervetti	2012	PA	2: E with low attrition	Good	T	N	2
Cervetti	2016	PA	2: E with low attrition	NI	T	N	2
Connor	2017	PA	2: E with low attrition	Good	R	Y	1
Duke	2021	PA	2: E with low attrition	Good w/R	T	Y	1
Gelzheiser	2011	PA	1: QE with baseline equivalence satisfied with covariate adjustment[1]	Good	T	Y	1
Guthrie	1998	PA	1: QE with baseline equivalence satisfied with covariate adjustment[1]	Good w/R	T	Y	2
Guthrie	1999	PA	1: QE with baseline equivalence satisfied with covariate adjustment[1]	Good w/R	T	Y	2
Guthrie	2004	PA	0: QE with a confounding factor (all control classes in one school)	Good w/R	T	Y	1
Guthrie	2009	PA	1: QE with baseline equivalence satisfied with covariate adjustment[1]	Good w/R	T	Y	0
Hanna	2008	D	0: QE with a confounding factor (one teacher for each condition)	NI	T	Y	2
Hinde	2007	PA	1: QE with baseline equivalence satisfied with covariate adjustment[1]	NI	T	N	1
Hinde	2011	PA	1: QE with baseline equivalence satisfied with covariate adjustment[1]	NI	T	N	1
Kim	2020	PA	2: E with low attrition	Good w/R	T	Y	1
Kim	2021	PA	2: E with low attrition	Good w/R	T	Y	1
Lohmann	1963	D	0: QE with baseline difference	NI	T	Y	1
Lutz	2006	PA	0: QE with a confounding factor (one teacher for all control-group students)	NI	T	Y	0
Martínez-Álvarez	2012	PA	0: QE with a confounding factor (one teacher for each condition)	Good	R	N	2
Morrow	1997	PA	2: E with low attrition	Good	T	Y	2
Neuman	2018	PA	1: E with differential attrition between conditions	Good	T	Y	1
Neuman	2021	PA	1: QE* with baseline equivalence satisfied with covariate adjustment[1]	Good	T	Y	1
Proctor	2019	PA	1: QE with baseline equivalence satisfied with covariate adjustment[1]	Good w/R	T	Y	0
Romance	1992	PA	0: QE with baseline difference	NI	T	Y	2
See	2017	PA	0: E with high attrition and unadjusted baseline difference[2]	Poor	T	Y	0
Stephens	2007	D	0: QE with baseline difference	Good	T	Y	1
Tong	2014	PA	2: E with low attrition	Good w/R	T	Y	1
Vitale	2011	PA	0: QE without information on baseline equivalence	Good w/R	T	Y	0
Vitale	2012	PA	2: E with low attrition	Good	T	Y	1
Williams	2005	PA	2: E with low attrition	Good w/R	T	Y	0
Williams	2007	PA	0: E with high attrition and unadjusted baseline difference[2]	Good w/R	T	Y	1
Williams	2009	PA	2: E with low attrition	Good w/R	T	Y	2
Williams	2014	PA	0: E with high attrition and unadjusted baseline difference[2]	Good	T	Y	1
Williams	2016	PA	1: E with differential attrition btw conditions	Good w/R	T	Y	1
Wright	2017	PA	0: QE with a confounding factor (all treatment-group teachers in one school and all control-group teachers in the other school)	Good	T	Y	1

Report type (PA = peer-reviewed article, D = Dissertation); Rating of research design (0 = does not meet standards of research design based on What Works Clearinghouse, 1 = meeting the standards with reservation, 2 = meeting the standards without reservation); Research design (E = Experimental design, QE = Quasi-experimental design); Adherence to the program (NI = No information reported, Poor = Poor adherence reported, Good = Good adherence reported without interrater reliability of adherence, Good w/R = Good adherence with good interrater reliability reported); Instructors (T = teachers, R = research team); Professional development (PD; N = not provided, Y = provided), Content comparison (indicating whether same science and/or social studies content was taught to both treatment- and control-group students; 0 = no information available, 1 = no explicit information, but similar content was likely to be taught between conditions as content taught to treatment-group students was from state standards or district curriculum, 2 = the same content was taught to both conditions)

[1]Covariates need to be premeasures (not demographic information).

[2]Covariates to adjust baseline difference were not used in analysis.

*Two classrooms did not receive a random assignment.

Information in each study necessary for meta-analysis was obtained by two authors (primarily the first and third authors), with no discrepancies in identified information.

Statistical analyses

For each study, standardized mean differences at posttest (Cohen's *d*) were calculated between treatment- and control-group students to examine the effects of integrated literacy and content-area instruction on vocabulary and comprehension as well as any changes in content knowledge. Administering content knowledge measures at posttest was not a criterion to be eligible for the current review. However, we examined the overall effect size for content knowledge as a supplementary analysis when content knowledge measures at posttest were available. Because 23 of the studies had nested data structures, we computed effect sizes using a procedure by Hedges (2007) to adjust effect sizes with an imputed estimate of interclass correlation (ICC). We used an ICC estimate of .175 (e.g., Graham, Kiuhara, & MacKay, 2020) as it is a midpoint of ICCs between .10 and .25, reported in educational research (Hedges & Hedberg, 2007).[4] In addition, four studies (i.e, Wigfield, Guthrie, Tonks, & Perencevich, 2004; Williams et al., 2005, 2014; Williams, Stafford, Lauer, Hall, & Pollini, 2009) provided class-level summary statistics and assigned classes to conditions. Applying another procedure by Hedges (2007), we first calculated individual-level effect sizes from the class-level statistics, and then adjusted effect sizes by taking into account the ICC. All effect sizes in the form of Cohen's *d* were converted to Hedges' *g*, as Hedges' *g* can correct overestimates of effect sizes due to small sample sizes (Borenstein, Hedges, Higgins, & Rothstein, 2009).

Random-effects models were used to combine effect sizes across studies, as we hypothesized that effect sizes would vary across studies rather than assuming one true effect size (Konstantopoulos & Hedges, 2019).[5] We incorporated robust variance estimation to examine random-effects models by using the *Robumeta* statistical package in *R* (Fisher & Tipton, 2015) to account for correlated effects among multiple measures from each study (Hedges, Tipton, & Johnson, 2010). Overall effect sizes were weighted by multiplying each effect size by the inverse of its variance. Kraft's benchmark (0.05 = small, 0.2 = medium, greater than 0.2 = large) was used to interpret effect sizes as it was developed for interpreting effect sizes in K-12 field interventions in education (Kraft, 2020). In addition to overall effect sizes, estimates of heterogeneity across studies, I^2 and τ^2, were computed. Next, power analysis was conducted to examine whether there was adequate statistical power to identify moderators by using *R* package (*Metapower*; Griffin, 2021), based on formulas by Valentine, Pigott, and Rothstein (2010). Then, using the *Robumeta* package, we conducted meta-regression analyses of potential moderators regarding features of research design and characteristics of the intervention. Each moderator variable was entered into the meta-regression analysis separately.

Finally, we examined potential publication bias in effect sizes. First, we used Egger's regression approach (Egger, Smith, Schneider, & Minder, 1997), which tests whether smaller studies include systematically different effect sizes than larger studies. Then, we executed sensitivity analysis to investigate the magnitude of publication bias using the *R* package *PublicationBias* developed by Mathur and VanderWeele (2020). This test explores how strong publication bias would need to be to attenuate effect size estimates to zero (i.e., non-significant). Publication bias for each overall effect size was examined separately.

Results

Characteristics of studies

Among the 35 studies included in the meta-analysis, 33 were conducted in the United States, one took place in the Netherlands (Aarnoutse & Schellings, 2003), and one was located in the United Kingdom

(See, Gorard, & Siddiqui, 2017). The 32 studies published in peer-reviewed journals contained 231 effect sizes[6] and the three unpublished dissertations included 9 effect sizes.[7] Thirteen studies focused on the primary grades (K-2), 20 studies included the upper elementary grades (3–5), and two studies included both primary and upper grades (K-4, Connor et al., 2017; Grade 2–3; See et al., 2017). In total, the 35 studies involved 13,289 students (9,530 treatment and 7,934 control) in kindergarten through Grade 5.

Regarding implementation of the interventions, most interventions were implemented as intended (the number of the studies [k] = 26) and focused on science and/or social studies content similar or same to the content taught to control-group students ($k = 28$). More descriptions about implementation of the intervention are summarized in Table 3, and characteristics of the interventions are described in Table 4 and Table 5. In addition, among the studies that revealed demographic information about students (see Table 6), approximately 66% of students had minoritized status (i.e., students of color), 62% were eligible for lunch services, and 16% were students with reading difficulties (or disabilities) or learning difficulties (or disabilities).

Effects of integrated literacy and content-area instruction on vocabulary

Students who received integrated literacy and content-area instruction showed improvement in vocabulary (the number of the outcome measures [n] = 36, $g = 0.91$, 95% confidence interval [CI] = [.34, 1.45]). The effect size was significant ($p < .01$) and large based on Kraft's (2020) benchmarks for interpreting effect sizes. Variance in effect sizes across studies (τ^2) was 0.49, and the proportion of true variance (I^2) was 91.82, providing evidence that the variance in effect sizes was mostly due to true variance across studies, rather than stemming from sampling error. In addition, the effect size for standardized vocabulary measures was not statistically significant ($n = 5$, $g = 0.64$, CI = [−1.59, 2.88], $p = .42$). On the other hand, a large and significant overall effect size was observed for researcher-developed vocabulary measures ($n = 31$, $g = 0.86$, CI = [.33, 1.38], $p < .01$). Looking further, researcher-developed measures highly aligned to the intervention displayed a significant overall effect size, whereas less aligned researcher-developed measures did not (see Table 7).

Regarding publication bias, the Egger's regression test suggested possible publication bias, $t(34) = 4.60$, $p < .001$. However, the publication bias was not large enough to be a concern. The sensitivity analysis reported *Not Possible* for the effect size estimate and 15.72 for the lower 95% confidence interval bound of the estimate. That is, no possible level of publication bias could attenuate the effect size to zero, and only an extreme case of publication bias (i.e., positive and significant results would have to be 15.72 times as likely to be published than non-significant results; Mathur & VanderWeele, 2020) could make the confidence interval include zero.

Effects of integrated literacy and content-area instruction on comprehension

Integrated literacy and content-area instruction improved students' comprehension ($n = 149$, $g = 0.40$, CI = [.19, .62]). The overall effect size was significant ($p < .001$) and large in light of Kraft's (2020) benchmarks. Variance in effect sizes across studies (τ^2) was 0.18, and the proportion of true variance (I^2) was 81.58, suggesting some heterogeneity in variance across the studies. These values also indicate that heterogeneity existed among studies, which might be explained by moderators.

The overall effect sizes for both researcher-developed ($n = 123$, $g = 0.54$, CI = [.17, .92], $p < .01$) and standardized comprehension outcomes ($n = 26$, $g = 0.25$, CI = [.04, .46], $p < .05$) were large (Kraft, 2020) and significant. In addition, effect sizes for highly and less aligned researcher-developed measures were positive and significant (see Table 7). With respect to publication bias, the Egger's regression test indicated some evidence of potential publication bias, $t(147) = 8.22$, $p < .001$. However, the magnitude of the publication bias was not substantial. The sensitivity analysis for publication bias using *R* software determined *Not Possible* for the effect size estimate and lower 95% confidence interval

Table 4. Descriptive Information about Characteristics of Interventions.

Study							
First author	Year	Grade	Length of the intervention[1]	Content-area	Literacy activities	Multiple text genres	Writing instruction
Aarnoutse	2003	3	One academic year, 40 lessons	SS	R/W	I/N: N to introduce SS topics	0
Bravo	2014	4, 5	8 weeks, daily	science	R/W	I	0
Cervetti	2012	4	8 weeks, 40 sessions	science	R/W	I	1: students learned how to write main ideas and details, as well as claim/evidence
Cervetti	2016	4	3 days	science	R	I	n/a
Connor	2017	K-4	12 weeks, 4 days/week	both	R/W	I	0
Duke	2021	2	One academic year, approximately 66 sessions	SS	R/W	I	1: mentor texts were provided to students
Gelzheiser	2011	4	1 semester, daily[2]	SS	R/W	I/N	0
Guthrie	1998	3, 5	One academic year	science	R/W	I/N	0
Guthrie	1999	5	One academic year	science	R/W	I/N: N to introduce science topics	0
Guthrie	2004	3	12 weeks, daily	science	R/W	I/N: More emphasis on I	0
Guthrie	2009	5	12 weeks, daily	science	R/W	I/N	0
Hanna	2008	4	12 days	science	R/W	I	0
Hinde	2007	3	1–2 lessons in 3–5 months	SS	R/W	unclear	0
Hinde	2011	3	3–5 lessons in 3–5 months	SS	R/W	unclear	0
Kim	2020	1	10 days	science	R/W	I/N: More emphasis on I	1: instruction for argumentative writing was provided
Kim	2021	1–2	20 lessons in 5–10 weeks	Both	R/W	I	1: instruction for argumentative writing was provided
Lohmann	1963	4	8 months, daily	SS	R/W	I/N	0
Lutz	2006	4	12 weeks, daily	science	R/W	I/N	0
Martínez-Álvarez	2012	4	6 days, daily	science	R/W	I	0
Morrow	1997	3	30 weeks, 2 hours and 15 minute/week	science	R/W	I/N: More emphasis on N	0
Neuman	2018	K	20 weeks, daily	science	R	I/N: N to introduce science topics	n/a
Neuman	2021	K-1	21 weeks, daily	science	R	I/N: N to introduce science topics	n/a
Proctor	2019	4,5	39 days, 30–40 minutes[3]	SS	R/W	I/N: Equal emphasis on I/N	1: instruction for argumentative writing was provided, student wrote for authentic audience.
Romance	1992	4	One academic year, two hours/day	science	R/W	I	0
See	2017	2	One academic year, 2 lessons/week	both	R/W	I	0
Stephens	2007	5	12 weeks, 2–3 days/week	science	R/W	I	0
Tong	2014	5	23 weeks, daily	science	R/W	I	0
Vitale	2011	1	8 weeks, daily	science	R/W	I	0

(*Continued*)

Table 4. (Continued).

Study							
First author	Year	Grade	Length of the intervention[1]	Content-area	Literacy activities	Multiple text genres	Writing instruction
Vitale	2012	2	Full school year, daily	science	R/W	I	0
Williams	2005	2	7–8 weeks, 2 sessions/week	science	R/W	I	1: students used paragraph frames to write summaries.
Williams	2007	2	22 lessons	SS	R/W	I/N: N to introduce science topics	0
Williams	2009	2	2 months, 3 sessions/week	science	R/W	I	1: students used paragraph frames to write summaries.
Williams	2014	2	11 weeks, 2 lessons/week	SS	R/W	I/N: N to introduce science topics	0
Williams	2016	2	25 weeks, 2 lessons/week	SS	R/W	I	0
Wright	2017	K	8 weeks, daily	science	R/W	I	1: teachers provided modeling of writing to students

Content-area (SS = social studies, both = science and social studies); Literacy activities (R = Reading only, R/W = Reading and writing); Multiple text genres (I/N = both informational and narrative texts, I = informational texts, N = narrative texts); Writing instruction (0 = students engaged in writing regularly, but there was no clear information about writing instruction; 1 = students engaged in writing regularly, and writing instruction was provided to students)

[1]Interventions were implemented for an average of 66.5 days (min = 3 days, max = 180 days).

[2]The intervention was provided as one-to-one instruction (Tier 3).

[3]Some students received the intervention during pulled-out instruction by specialist teachers (Tier 2).

Table 5. Descriptive Information about Other Characteristics of Interventions.

Study First author	Year	Grade	Small group/ Pair work	Comp. Str.	Graph. Org.	Concept map	Hands-on activities	Other support
Aarnoutse	2003	3	Y	Y	N	N	Y	
Bravo	2014	4, 5	Y	Y	Y	N	Y	culturally and linguistically responsive approaches to support ELs
Cervetti	2012	4	Y	Y	N	Y	Y	
Cervetti	2016	4	Y	N	N	N	Y	
Connor	2017	K-4	Y	N	Y	N	Y	leveled texts based on students' reading levels, original sources related to social studies topics
Duke	2021	2	Y	N	Y	N	Y	literacy activities connected to authentic social problems and the local community
Gelzheiser	2011	4	n/a[1]	Y	N	N	N	phonics instruction before reading content-related texts, choice of texts provided to students to support their motivation to read
Guthrie	1998	3, 5	Y	Y	N	N	Y	choice of texts provided to students to support their motivation to read
Guthrie	1999	5	Y	Y	N	N	Y	choice of texts provided to students to support their motivation to read, texts appropriate for students with different reading levels
Guthrie	2004	3	Y	Y	Y	N	Y	fluency support embedded in the integrated instruction, choice of texts provided to student to support their motivation to read, texts appropriate for students with different reading levels
Guthrie	2009	5	Y	Y	Y	Y	Y	fluency support embedded in the integrated instruction (more fluency support for students with low reading proficiency), choice of texts provided to students to support their motivation to read, leveled texts based on students' reading levels
Hanna	2008	4	Y	N	Y	Y	N	
Hinde	2007	3	N	N	N	N	N	
Hinde	2011	3	N	Y	N	N	N	
Kim	2020	1	Y	N	N	Y	N	choice of texts provided to students to support their motivation to read
Kim	2021	1–2	Y	Y	N	Y	N	
Lohmann	1963	4	Y	N	N	N	Y[2]	watching videos related to social studies topics
Lutz	2006	4	Y	Y	N	N	N	choice of texts provided to students to support their motivation to read, texts appropriate for students with different reading levels
Martínez-Álvarez	2012	4	Y	Y	N	N	Y[2]	bilingualism (students used whatever linguistic sources available to them), web-based virtual environment
Morrow	1997	3	Y	N	N	N	N	
Neuman	2018	K	N	N	N	N	N	
Neuman	2021	K-1	N	N	N	N	N	
Proctor	2019	4,5	Y	Y	Y	Y	N	content-related texts representing bilingual people of color, bilingualism to teach vocabulary, morphology, and syntax (e.g., comparing English and Spanish or Portuguese), videos related to social studies topics
Romance	1992	4	N	Y	N	N	Y	
See	2017	2	N	N	N	N	Y	
Stephens	2007	5	Y	N	Y	N	N	
Tong	2014	5	Y	Y	Y	N	Y	fluency support embedded in the integrated instruction, language scaffolding and use of technology to support ELs
Vitale	2011	1	N	N	N	Y	Y	
Vitale	2012	2	Y	N	N	Y	Y	
Williams	2005	2	N	N/Y	Y	N	N	
Williams	2007	2	N	N/Y	Y	N	N	
Williams	2009	2	N	N/Y	Y	N	N	
Williams	2014	2	N	N/Y	Y	N	N	
Williams	2016	2	N	N/Y	Y	N	N	
Wright	2017	K	N	N	N	N	Y	dramatic play

Small group/pair work (N = no explicit statement, Y = used); Comp. Str. = Comprehension strategies (N = not taught, Y = taught, N/Y = not taught in one treatment group and taught in the other treatment group); Graph.Org. = Graphic organization (N = not used; Y = used); Concept map to show associations among words around concept (N = not used, Y = used); Hands-on activities (N = not provided, Y = provided)
[1]The treatment was provided as one-to-one instruction.
[2]Hands-on activities took place inside and outside classrooms.

bound of the estimate. The result demonstrated that the estimate of the effect size was robust as no level of publication bias could reduce the effect size to zero.

Effects of integrated literacy and content-area instruction on content knowledge

Although reporting outcome measures for content knowledge was not a criterion for inclusion, we conducted a supplementary analysis to investigate the effects of integrated literacy and content-area instruction on science and/or social studies knowledge for studies that did include this information. Overall, integrated literacy and content-area instruction had a significant and positive effect of 0.89 on knowledge measures (n = 55, CI = [.54, .1.23]). The overall effect size was significant ($p < .001$) and large according to Kraft's (2020) benchmarks. Substantial heterogeneity across the studies was observed. The between-study variance (τ^2) was 0.39, and the proportion of true variance (I^2) was 87.49.

In addition, the overall effect size for researcher-developed knowledge outcomes was large and significant (n = 49, g = 0.94, CI = [.53, 1.34], $p < .001$), whereas the effect size for standardized knowledge outcomes was not significant (n = 6, g = 0.68, CI = [−.38, 1.74], p = .13). When the researcher-developed measures were classified into the highly and less aligned measures, a significant effect was observed for the highly aligned measures only (see Table 7). Some evidence of potential publication bias for the overall effect size for knowledge outcomes was observed, based on the results of the Egger's regression, $t(53) = 2.82$, $p < .01$. The result of sensitivity analysis for publication bias, however, indicated that the publication bias was not substantial in that it would be *Not Possible* to attenuate the effect size and confidence interval to zero.

Meta-regression with potential moderators

Research design features (research question 2) and characteristics of interventions (research question 3) were examined as potential moderators to understand heterogeneity of variance in the overall effect sizes for vocabulary and comprehension. Table 8 summarizes the results regarding research design features, and Table 9 displays characteristics of interventions. We reported only the results of the analyses with adequate statistical power (see Table S3 in Supplementary Material). Meta-regression analyses with sufficient statistical power revealed that none of the research design features were significant moderators for explaining the differences in the overall effect sizes for vocabulary and comprehension. Similarly, there was no significant moderator among characteristics of the interventions to predict the variance in the overall effect size for vocabulary.

Discussion

Multiple scholars have postulated that integrated literacy and content-area instruction can enhance vocabulary and comprehension because of mutually enhancing relationships among vocabulary, comprehension, and content knowledge (Cabell & Hwang, 2020; Cervetti, Barber, Dorph, Pearson, & Goldschmidt, 2012; Duke, Pearson, Strachan, & Billman, 2011; Neuman, Newman, & Dwyer, 2011). Whereas prior reviews have synthesized some of the extant literature base, this study extends the literature by meta-analyzing the impact of (quasi) experimental studies that have integrated literacy and content-area instruction in the elementary years.

Table 6. Students' Demographic Information and Effect Sizes.

Study		Characteristics of interventions – Students' demographic				Effect sizes (*g*)	
First author	Year	Grade	%minor	%lunch service	%RD/LD	Vocabulary	Comprehension
Aarnoutse	2003	3	10	NI	NI		−0.03
Bravo	2014	4, 5	NI	71.4	5	0.09	−0.32
Cervetti	2012	4	49	55.3	NI	0.26	−0.13
Cervetti	2016	4	14	NI	NI	0.13, 0.61	−1.15– 2.35
Connor	2017	K-4	66	57	12	0.05	−0.06, 0.01
Duke	2021	2	59.7	80.4	NI		0.15
Gelzheiser	2011	4	35.5	42.5	100	0.53	0.94
Guthrie	1998	3, 5	78	NI	13		−0.29– 0.76
Guthrie	1999	5	77	NI	NI		−0.11– 0.68
Guthrie	2004	3	43	13	13.6		−0.47– 0.62
Guthrie	2009	5	40.5	NI	16.5		0.90
Hanna	2008	4	20	69	NI	0.28– 1.18	
Hinde	2007	3	NI	NI	NI		−0.13– 0.21
Hinde	2011	3	NI	NI	NI		−0.12– 0.31
Kim	2020	1	79	NI	7	−0.06, 0.53	0.04, 0.34
Kim	2021	1–2	81.5	40[1]	8.5	0.22– 0.65	−0.09
Lohmann	1963	4	NI	NI	NI		0.66
Lutz	2006	4	12	13	NI		1.12, 1.42
Martínez-Álvarez	2012	4	NI	NI	10		3.06
Morrow	1997	3	54	28	NI		0.03– 2.71
Neuman	2018	K	98.5	79	14.5	−0.08– 0.25	
Neuman	2021	K-1	99	95.5	NI	0.07– 0.48	
Proctor	2019	4, 5	85	72.2	NI		0.11, 0.15
Romance	1992	4	NI	NI	NI		0.56
See	2017	2	49.9	25.8	17.4		−0.03 – *−0.02*
Stephens	2007	5	82	84.5	10.3	0.28– 1.18	−0.31– 0.09
Tong	2014	5	100	85	NI	3.63	
Vitale	2011	1	64.8	40	NI		0.27, 0.77
Vitale	2012	2	79.5	40	NI		0.49, 0.91
Williams	2005	2	99	88	8.5		
Williams	2007	2	99.5	93	5	3.94, 4.74	−0.43– 3.09
Williams	2009	2	99	90	6	0.56, 0.91	0– 2.33
Williams	2014	2	99.6	90	10.7	1.07, 4.29	−0.18– 0.89
Williams	2016	2	97.1	84	18	1.14, 1.58	0.15– 2.02
Wright	2017	K	40	53.5	NI	1.08, 1.71	

% minor = percentage of minoritized students; % lunch service = percentage of students eligible for lunch service; % RD/LD = percentage of students with reading difficulties (or disabilities) or learning difficulties (or disabilities); NI = No Information

Impact of integrated literacy and content-area instruction on literacy outcomes

The results of our meta-analysis provide evidence that integrated literacy and content-area instruction can improve comprehension in the elementary years. The overall effect sizes for researcher-developed (ES = 0.54) and standardized outcome measures for comprehension (ES = 0.25) were statistically significant and large (Kraft, 2020). Converting the effect sizes for all measures to an U_3 metric, in order to examine the percentage of the treatment-group students who exceed the control group's mean (Valentine, Aloe, & Wilson, 2019), approximately 69% and 60% of treatment group students scored better on researcher-developed and standardized comprehension measures, respectively, than the average scores of control group students. It appears that integrated literacy and content-area instruction, when compared with traditional literacy and content-area instruction, might be as effective as other approaches to enhance comprehension. For example, Elleman, Lindo, Morphy, and Compton (2009) conducted meta-analyses on interventions for vocabulary (Pre-K – 12) and found a significant effect size for researcher-developed comprehension measure (ES = 0.50). Meta-analyses on

Table 7. Types of Outcome Measures as Potential Moderators of Effect Sizes for Vocabulary, Comprehension, and Content Knowledge.

Vocabulary				
	B	*p*	ES	95% CI
Standardized vs. researcher-developed (n = 36)	underpowered			
Standardized (*n* = 5)			0.64	−1.59, 2.88
Researcher developed (*n* = 31)			0.86	0.33, 1.38
Less proximal vs. complete proximal (n = 31)	0.87	.09		
Less proximal researcher-developed (*n* = 6)			0.11	−0.46, 0.67
Complete proximal researcher-developed (*n* = 25)			1.05	0.44, 1.66
Comprehension				
	B	*p*	ES	95% CI
Standardized vs. researcher-developed (*n* = 149)	underpowered			
Standardized (*n* = 26)			0.25	0.04, 0.46
Researcher developed (*n* = 123)			0.54	0.17, 0.92
Less proximal vs. complete proximal (*n* = 123)	0.54	.26		
Less proximal researcher-developed (*n* = 77)			0.34	0.10, 0.58
Complete proximal researcher-developed (*n* = 46)			0.85	0.20, 1.50
Content knowledge				
	B	*p*	ES	95% CI
Standardized vs. researcher-developed (*n* = 55)	underpowered			
Standardized (*n* = 6)			0.68	−0.38, 1.74
Researcher developed (*n* = 49)			0.94	0.53, 1.34
Less proximal vs. complete proximal (*n* = 49)	underpowered			
Less proximal researcher-developed (*n* = 10)			0.70	−0.09, 1.48
Complete proximal researcher-developed (*n* = 39)			0.97	0.51, 1.42

n = the number of the outcome measures
Underpowered = due to inadequate statistical power, comparing effect sizes by types of measures was not possible

interventions for text structures (Grade 2– 12) by Hebert, Bohaty, Nelson, and Brown (2016) revealed effect sizes of 0.57 and 0.13 for researcher-developed and standardized measures, respectively.

The overall effect size for researcher-developed outcome measures for vocabulary (ES = 0.91) was statistically significant and large (Kraft, 2020). Translating to the U_3 metric, approximately 84% of treatment group students scored better than the average score of control group students for researcher-developed vocabulary measures. Similar to what we discovered for comprehension, integrated literacy and content-area instruction might be as effective as other approaches to enhance vocabulary learning of the words taught during instruction. For example, Elleman et al. (2009) reported an effect size of 0.79 and Marulis and Neuman (2010) an effect size of 1.21 for researcher-developed vocabulary measures in their meta-analysis exploring vocabulary interventions (Pre-K – 12 for Elleman et al., 2009 and Pre-K – K for; Marulis & Neuman, 2010).

However, the overall effect for standardized vocabulary measures in the present study was not statistically significant. That is, we could not detect any evidence for improving generalized vocabulary, even though integrated literacy and content-area instruction seems to support vocabulary learning of words taught in the intervention. Similarly, the integrated approach was observed to improve students' knowledge of science or social studies taught during the intervention, but not generalized content knowledge.[8]

The significant effects on researcher-developed outcomes for comprehension, vocabulary, and content knowledge are partially due to fact that researcher-developed measures more closely reflect the intervention than standardized measures (Slavin & Madden, 2011). The question remains: Why did integrated literacy and content-area instruction enhance comprehension in general, but not generalized vocabulary and content knowledge? First, students might have been able to leverage their relational understanding about topics they learned during the intervention to infer a pattern of relations among ideas they did not know well in texts (Alexander & DRLRL, 2012; Pritchard, 2019). As

Table 8. Quality of Research Design and Features of Implementation of the Interventions as Potential Moderators of Effect Sizes for Vocabulary and Comprehension.

	Vocabulary					Comprehension				
	k	*B*	*p*	ES	95% CI	*k*	*B*	*p*	ES	95% CI
Quality of research design	17	−1.59	.15			30	−0.43	.14		
Not meeting standards of research design	4			2.32	−0.26, 4.91	10			0.73	0.10, 1.36
Meeting standards of research design	13			0.54	0.14, 0.95	20			0.23	0.07, 0.40
Adherence to the program	15	0.82	.18			22	UP			
Good adherence without IRR reported	7			0.36	−0.03, 0.76	8			0.73	−0.14, 1.61
Good adherence with IRR reported	8			1.40	0.08, 2.72	14			0.31	0.09, 0.52
Professional development	17	0.79	.11			30	UP			
Not offered to teachers	3			0.30	−0.27, 0.87	6			0.44	−0.68, 1.56
Offered to teachers	14			1.09	0.36, 1.81	24			0.39	0.22, 0.56
Comparison of content taught btw conditions[1]	17	UP				30	UP			
No information available	1			0.87	n/a	7			0.37	−0.07, 0.81
Same or similar content and vocabulary	16			0.92	0.31, 1.52	23			0.42	0.15, 0.69

k = the number of the studies; UP (Underpowered) = due to inadequate statistical power, comparing effect sizes by types of measures was not possible; IRR = interrater reliability

[1]The effect sizes for vocabulary and comprehension for those studies in which students in both conditions learned same or similar content and vocabulary were significant, indicating that treatment-group students who received integrated instruction performed better on vocabulary and comprehension outcomes than control-group students who did not, even though both groups learned the same or similar content and vocabulary.

a result, students might have improved comprehending a variety of topics, resulting in better performance on standardized comprehension measures.

Second, for standardized vocabulary and content-knowledge tests, it might have been virtually impossible for students to leverage their relational understanding about topics they learned to make inferences about vocabulary and topics they did not know well in the tests. Unlike standardized comprehension tests that provided information on different topics in texts, standardized vocabulary and content-knowledge tests might rarely have provided information about vocabulary and topics, so students were unlikely to know how vocabulary and topics asked in the test were related to what they already knew. Third, the interventions might not have increased print exposure to the extent that it would have enhanced generalized vocabulary and content knowledge. Amount of print exposure is predictive of vocabulary and content knowledge as it is a crucial means to acquiring a variety of vocabulary words and information about the natural and social world (Cunningham & Stanovich, 1991; Stanovich & Cunningham, 1993). In addition, the similar results between vocabulary and content knowledge can be partially attributed to the idea that words are labels of content knowledge (Pearson & Billman, 2016). As word learning can be regarded as the surface level of content learning (Ambruster, 1992), the development of vocabulary and content knowledge might display a similar pattern. With that being said, caution must be applied in interpreting the non-significant effects on generalized vocabulary and content knowledge because the results were obtained from only three studies (Connor et al., 2017; Neuman & Kaefer, 2018; Tong, Deacon, & Cain, 2014).

To examine the variability of effect sizes for vocabulary and comprehension, features of research designs and characteristics of interventions were examined as moderators. Most moderator analyses in the present study did not have sufficient power to detect significant moderators. Among the moderation analyses with at least 80% power to reject the false null hypothesis, no statistically significant moderators were identified. The non-significant findings should be interpreted with caution. As most moderators with sufficient power had imbalanced proportions of studies for each category of the moderators, the moderators might not actually have had the 80% power (Hempel et al., 2013). Alternatively, the non-significant moderators might indicate that, regardless of features of research designs and characteristics of the interventions, the effect sizes for vocabulary and comprehension were consistent across the studies.

Table 9. Characteristics of Interventions as Potential Moderators of Overall Effect Size for Vocabulary and Comprehension.

	Vocabulary					Comprehension				
	k	*B*	*p*	ES	95% CI	*k*	*B*	*p*	ES	95% CI
Number of sessions	17	UP				30	UP			
Smaller than the medium (40)	8			1.20	0.00, 2.40	12			0.48	−0.02, 0.98
Equal to or greater than the medium	9			0.70	0.06, 1.34	18			0.35	0.12, 0.58
Hands-on activities	17	UP				30	UP			
Not used	11			0.96	0.18, 1.75	14			0.42	0.16, 0.68
Used	6			0.87	−0.37, 2.12	16			0.39	0.03, 0.76
Science or social studies[1]	15	1.32	.26			27	UP			
Science	11			0.70	0.16, 1.25	17			0.57	0.18, 0.97
Social studies	4			2.18	−0.65, 5.02	10			0.35	0.10, 0.59
Comprehension strategies[2]	17	UP				30	UP			
Not taught	12			0.96	0.25, 1.68	16			0.26	0.11, 0.42
Taught	10			1.31	0.24, 2.37	19			0.59	0.21, 0.97
Multiple genres	17	UP				28	UP			
Informational texts only	11			0.85	0.28, 1.42	15			0.38	−0.03, 0.79
Informational and narrative texts	6			1.11	−0.70, 2.92	13			0.51	0.26, 0.76
Writing instruction[3]	14	−0.90	.17			29	UP			
Not provided	8			1.62	0.22, 3.01	22			0.52	0.22, 0.81
Provided	6			0.51	0.08, 0.94	7			0.08	−0.14, 0.30
Small group/pair work	17	−0.61	.32			30	UP			
Unclear	8			1.34	0.13, 2.55	10			0.36	0.10, 0.61
Used	9			0.63	0.02, 1.24	20			0.43	0.10, 0.75
Graphic organizer	17	1.02	.08			30	UP			
Not used	8			0.40	0.12, 0.68	18			0.46	0.13, 0.79
Used	9			1.56	0.36, 2.77	12			0.32	0.06, 0.57
Concept map	17	−0.61	.20			30	UP			
Not used	13			1.14	0.32, 1.96	23			0.46	0.18, 0.75
Used	4			0.46	−0.08, 1.00	7			0.24	−0.11, 0.59
Grade[4]	16	UP				29	UP			
Primary grades (K-2)	10			1.06	0.15, 1.97	11			0.34	0.09, 0.59
Upper grades (3–5)	6			0.93	−0.26, 2.12	19			0.43	0.09, 0.77
% students with RD/LD	17	UP				30	UP			
Less than the median (10.7%)	6			1.06	−0.56, 2.68	8			0.58	−0.32, 1.49
Equal to or greater than the median	11			0.84	0.22, 1.45	22			0.32	0.15, 0.49
% minoritized students	16	UP				24	UP			
Less than 49.5%	5			0.74	0.18, 1.31	7			0.45	−0.05, 0.94
49.5% or greater than 49.5%	11			1.13	0.15, 2.11	17			0.30	0.11, 0.48
% students with lunch services[5]						19	UP			
Less than 49.5%						8			0.49	0.05, 0.92
49.5% or greater than 49.5%						11			0.20	−0.06, 0.46

k = the number of the studies; UP (Underpowered) = due to inadequate statistical power, comparing effect sizes by types of measures was not possible; % RD/LD = percentage of students with reading difficulties (or disabilities) or learning difficulties (or disabilities)

[1]Connor et al. (2017) was excluded in the analysis as both science and social studies taught in the study.

[2]Five studies had two treatment groups, one with support for comprehension strategies and the other without support for comprehension strategies.

[3]Writing instruction: not provided (students engaged in writing, but writing instruction was not provided), provided (students engaged in writing, and writing instruction was provided as well)

[4]One study (See et al., 2017) that examined the effect on comprehension included students in second and third grade.

[5]Percentage of students with lunch services was not examined as a moderator to understand the effect on vocabulary because there were only two studies in which the percentage of students with lunch services was less than 49.5%.

Limitations

Several limitations to the present study warrant note. First, we purposefully did not include different types of unpublished research beyond dissertations. We wanted to understand not only impact but also characteristics of the interventions and research design features; gray literature (e.g., conference proposals) typically does not provide such information. Second, we only considered studies written in

English, as it is the only common language spoken by all members of the author team. This may have resulted in overlooking international work. Third, we were not able to make a distinction between listening comprehension and reading comprehension because some studies were not clear in their description about comprehension measures. We were unsuccessful in our attempts to seek clarifying information from the authors of those studies. Fourth, most of our meta-regressions did not have adequate statistical power to detect moderators of the effect sizes, due to the small number of the studies included in the meta-analysis. More studies are needed to test impacts of integrated literacy and content-area instruction on vocabulary and comprehension and examine the mechanism of the impacts. Fifth, we could not capture some important instructional characteristics such as teacher efficacy and social interaction among students, as well as intensity of instructional components (e.g., frequency and time spent for writing), due to limited information.

Potential implications for research, policy, and practice

The current review of previous studies indicated that more (quasi)experimental research is needed to better understand this topic, particularly with primary-grade students. Currently, only 13 studies focused on the effects of integrated literacy and content-area instruction on vocabulary and/or comprehension for students in K-2. In addition, longitudinal investigations of this type of integrated instruction are extremely rare (Cabell & Hwang, 2020). Longitudinal studies that examine the implementation of the integrated literacy and content-area instruction over the course of multiple years can reveal long-term or cumulative impacts on vocabulary and comprehension.

With regard to practice and policy, the findings of the meta-analysis provide evidence to advocate for integrated literacy and content-area instruction in the elementary years to support vocabulary and comprehension. Most instructional and policy efforts to date, however, have focused on robust instruction of vocabulary and comprehension, with limited attention to building content knowledge (e.g., Reading First). The meta-analysis does not invalidate the importance of this instruction. Rather, our findings highlight the possibility that integrated literacy and content knowledge instruction might optimally support elementary-aged students' vocabulary and comprehension. Indeed, building students' knowledge in support for literacy development has been a salient part of the national conversation in the U.S. in recent years (e.g., Hirsch & Hansel, 2013; Wexler, 2019).

In contemporary practice in the U.S., however, compartmentalization of subject areas at school (Lammert, 2020) has led to literacy instruction that is often disconnected from building content knowledge, and content-area instruction only sparingly involves reading and/or writing instruction (Barber & Cervetti, 2019; Levstik & Barton, 2008). Many have also decried the insufficient instructional time spent in science and social studies (Teale, Paciga, & Hoffman, 2007; Tyner & Kabourek, 2020). The findings of our meta-analysis support challenging the current status quo, the separation of literacy and content-area instruction, and shifting to more fully integrate the two (e.g., Maerten-Rivera, Ahn, Lanier, Diaz, & Lee, 2016).

Author Note

We thank Thomas G. White at the University of Virginia, Xiaoning Natalie Wang and Sen Wang at the Florida State University for their contributions to this study, and Richard K. Wagner at the Florida State University for and Gabriel DellaVecchia at the University of Michigan for thoughtful feedback on the manuscript. The research reported here was supported by the Institute of Education Sciences, U.S. Department of Education, through Grant R305A170635 awarded to Florida State University (PI: Cabell). The opinions expressed are those of the authors and do not represent views of the Institute or the U.S. Department of Education.

Notes

1. *Content learning* indicates learning of science and social studies, and *content-area instruction* (or *content teaching*) means science or social studies instruction throughout the article. The current meta-analysis did not

account for other areas such as mathematics, craft, music, and Physical Education because these areas place less emphasis on learning of information about nature and society. Learning of mathematics often focuses on identifying problems in situations and executing strategies of mathematical solutions (e.g., Fuchs et al., 2021). Craft, music, and physical education often focus on skills and knowledge to perform arts or physical activities (e.g., Bresler, 1996). English language arts (ELA) is not regarded as content-area instruction, but rather as literacy instruction, because ELA traditionally emphasizes reading and writing instruction without placing much attention to learning of nature and the social world (Hwang et al., 2021). On the other hand, we use *subjects* as an inclusive term to refer to different areas taught in school (i.e., beyond science and social studies).

2. The overall effect size was not available in Romance and Vitale (2012a) and was .91 in Guthrie et al. (2007).
3. The first author developed the coding manual, and the second author, along with a literacy expert with a doctoral degree, reviewed the coding manual. After the review, the two authors coded two studies randomly selected (Connor et al., 2017; Romance & Vitale, 2017). All coding was identical between the two authors except one item about comprehension strategies. The coding item was clarified (comprehension strategies are considered taught when authors explicitly mentioned them, and discussion strategies such as think-pair-share are not regarded as comprehension strategies). Then the first author trained the third author by explaining the coding manual and modeling coding of the two studies (Connor et al., 2017; Romance & Vitale, 2017). After the training, the third author independently coded two randomly selected studies (Martínez-Álvarez, Bannan, & Peters-Burton, 2012; Stephens, 2007). The agreement rate of coding of the two studies between the first and third authors was 90.4%. All studies were double-coded, and the interrater agreement was calculated by dividing the correct coding by the total number of coding (26 items for each of the 35 studies). The authors met regularly to compare coding and resolve discrepancies of the coding.
4. Three additional studies had nested data structures (i.e., Hinde et al., 2007; Hinde, Popp, Jimenez-Silva, & Dorn, 2011; See et al., 2017), but it was not possible to calculate adjusted effect sizes for these studies because information on average cluster size (e.g., number of students per class) was not reported.
5. The effect sizes for vocabulary and comprehension were first computed by synthesizing all studies, regardless of whether the content area integrated with literacy instruction was science or social studies, because both content areas aim to attain the same instructional goal (i.e., supporting students to gain information about the world) even though the focus is different between the two areas (science mainly on information of the natural world and social studies mainly on information of the social world). Then, the effect sizes for vocabulary and comprehension were separately calculated by content area (science vs. social studies) to better understand the impact of the interventions (see Table 11 and 12).
6. 145 effect sizes for comprehension, 33 for vocabulary, and 53 for science or social studies content knowledge.
7. four effect sizes for comprehension, three for vocabulary, and two for science content knowledge.
8. The overall effect size for researcher-developed content-knowledge measures (obtained from studies that measured content knowledge at posttest in addition to vocabulary and/or comprehension) was large and statistically significant (ES = 0.94; Kraft, 2020). Translating to the U_3 metric, approximately 84% of treatment group students scored better than the average score of control group students on content knowledge measures.

Disclosure statement

The authors have no conflicts of interest to disclose.

ORCID

HyeJin Hwang http://orcid.org/0000-0002-1049-7941
Sonia Q. Cabell http://orcid.org/0000-0002-2586-5735

References

References marked with an asterisk indicate studies reviewed in the current meta-analysis.

*Aarnoutse, C., & Schellings, G. (2003). Learning reading strategies by triggering reading motivation. *Educational Studies*, *29*(4), 387–409. doi:10.1080/0305569032000159688

Alexander, P. A., & Disciplined reading and learnin, T. The Disciplined Reading and Learning Research Laboratory [DRLRL]. (2012). Reading into the future: Competence for the 21st century. *Educational Psychologist*, *47*(4), 259–280.

Alexander, P. A., Jetton, T. L., & Kulikowich, J. M. (1995). Interrelationship of knowledge, interest, and recall: Assessing a model of domain learning. *Journal of Educational Psychology*, *87*(4), 559–575. doi:10.1037/0022-0663.87.4.559

Allen, L. K., & McNamara, D. S. (2020). Defining deep reading comprehension for diverse readers. In E. B. Moje, P. P. Afflerbach, P. Enciso, & N. K. Lesaux (Eds.), *Handbook of reading research,* (pp. 261–276). Routledge

Ambruster, B. B. (1992). Vocabulary in content area lessons. *Reading Teacher*, *45*, 550–551.

August, D., Artzi, L., & Barr, C. (2016). Helping ELLs meet standards in English Language Arts and science: An intervention focused on academic vocabulary. *Reading and Writing Quarterly*, *32*(4), 373–396. doi:10.1080/10573569.2015.1039738

August, D. E., & Hakuta, K. E. (1997). *Improving schooling for language-minority children: A research agenda*. Washington, DC: National Academy Press.

Austin, C. R., Vaughn, S., & McClelland, A. M. (2017). Intensive reading interventions for inadequate responders in grades K-3: A synthesis. *Learning Disability Quarterly*, *40*(4), 191–210. doi:10.1177/0731948717714446

Barber, J., & Cervetti, G. N. (2019). *No more science kits or texts in isolation*. Portmouth, NH: Heinemann.

Biemiller, A., & Slonim, N. (2001). Estimating root word vocabulary growth in normative and advantaged populations: Evidence for a common sequence of vocabulary acquisition. *Journal of Educational Psychology*, *93*(3), 498–520. doi:10.1037/0022-0663.93.3.498

Bogaerds-Hazenberg, S. T. M., Evers-Vermeul, J., & van den Bergh, H. (2020). A meta-analysis on the effects of text structure instruction on reading comprehension in the upper elementary grades. *Reading Research Quarterly*, 1–28. doi:10.1002/rrq.311

Borenstein, M., Hedges, L. V., Higgins, J. P., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. West Sussex, UK: Wiley. doi:10.1002/9780470743386

*Bravo, M. A., & Cervetti, G. N. (2014). Attending to the language and literacy needs of English learners in science. *Equity & Excellence in Education*, *47*(2), 230–245. doi:10.1080/10665684.2014.900418

Bresler, L. (1996). Traditions and change across the arts: Case studies of arts education. *International Journal of Music Education*, *27*(1), 24–35. doi:10.1177/025576149602700104

Brunner, J. L., & Abd-El-Khalick, F. (2020). Improving nature of science instruction in elementary classes with modified science trade books and educative curriculum materials. *Journal of Research in Science Teaching,57*(2), 154–183. https://doi.org/10.1002/tea.21588

Burgoon, E. M., Henderson, M. D., & Markman, A. B. (2013). There are many ways to see the forest for the trees: A tour guide for abstraction. *Perspectives on Psychological Science*, *8*(5), 501–520. doi:10.1177/1745691613497964

Cabell, S. Q., & Hwang, H. (2020). Building content knowledge to boost comprehension in the primary grades. *Reading Research Quarterly*, *55*(S1). doi:10.1002/rrq.338

Cervetti, G. N., & Barber, J. (2008). Text in hands-on science. In E. H. Hiebert & M. Sailors (Eds.), *Finding the right texts: What works for beginning and struggling readers* (pp. 89–108). New York, NY: Guilford.

Cervetti, G. N. (2013). Integration of literacy and science. In B. M. Taylor & N. K. Duke (Eds.), *Handbook of effective literacy instruction: Research-based practice K–8* (pp. 371–393). New York: NY: Guilford.

Cervetti, G. N., & Wright, T. S. (2020). The role of knowledge in understanding and learning from text. In E. B. Moje, P. Afflerbach, P. Enciso, & N. K. Lesaux (Eds.), *Handbook of reading research: Volume V* (pp. 237–260). New York, NY: Routledge.

*Cervetti, G. N., Barber, J., Dorph, R., Pearson, P. D., & Goldschmidt, P. G. (2012). The impact of an integrated approach to science and literacy in elementary school classrooms. *Journal of Research in Science Teaching*, *49*(5), 631–658. doi:10.1002/tea.21015

*Cervetti, G. N., Wright, T. S., & Hwang, H. (2016). Conceptual coherence, comprehension, and vocabulary acquisition: A knowledge effect? *Reading and Writing: An Interdisciplinary Journal*, *29*(1), 1–19. doi:10.1007/s11145-016-9628-x

Cheung, A. C. K., & Slavin, R. E. (2016). How methodological features affect effect sizes in education. *Educational Researcher*, *45*(5), 283–292. doi:10.3102/0013189X16656615

Ciullo, S., Collins, A., Wissinger, D. R., McKenna, J. W., Lo, Y. L., & Osman, D. (2020). Students with learning disabilities in the social studies: A meta-analysis of intervention research. *Exceptional Children*, *86*(4), 393–412. doi:10.1177/0014402919893932

Ciullo, S., Lo, Y.-L. S., Wanzek, J., & Reed, D. K. (2016). A synthesis of research on informational text reading interventions for elementary students with learning disabilities. *Journal of Learning Disabilities*, *49*(3), 257–271. doi:10.1177/0022219414539566

Cohen, J. (1977). *Statistical power analysis for the behavioral sciences*. New York, NY: Academic Press.

*Connor, C. M. D., Dombek, J., Crowe, E. C., Spencer, M., Tighe, E. L., Coffinger, S., . . . Petscher, Y. (2017). Acquiring science and social studies knowledge in kindergarten through fourth grade: Conceptualization, design, implementation, and efficacy testing of content-area literacy instruction (CALI). *Journal of Educational Psychology*, *109*(3), 301–320. doi:10.1037/edu0000128

Cunningham, A. E., & Stanovich, K. E. (1991). Tracking the unique effects of print exposure in children: Associations with vocabulary, general knowledge, and spelling. *Journal of Educational Psychology*, *83*(2), 264–274. doi:10.1037//0022-0663.83.2.264

Davis, A. (2018). *An integrated literacy/science intervention for English language learners in third grade*. Manhattan: Kansas State University.

de Boer, H., Donker, A. S., & van der Werf, M. P. C. (2014). Effects of the attributes of educational interventions on students' academic performance: A meta-analysis. *Review of Educational Research*, *84*(4), 509–545. doi:10.3102/0034654314540006

Dickinson, V. L., & Young, T. A. (1998). Elementary Science and Language Arts: Should We Blur the Boundaries? *School Science and Mathematics*, *98*(6), 334–339. doi:10.1111/j.1949-8594.1998.tb17429.x

Duke, N. K., Pearson, P. D., Strachan, S. L., & Billman, A. K. (2011). Essential elements of fostering and teaching reading comprehension. In S. J. Samuels & A. E. Farstrup (Eds.), *What research has to say about reading instruction* (4th ed., pp. 51–93). Newark, DE: International Reading Association.

Duke, N. K., & Carlisle, J. (2011). The development of comprehension. In M. L. Kamil, P. D. Pearson, E. Moje, & P. Afflerbach (Eds.), *Handbook of reading research* (Vol. 4, pp. 199–228). London, UK: Routledge.

*Duke, N. K., Halvorsen, A. L., Strachan, S. L., Kim, J., & Konstantopoulos, S. (2021). Putting PjBL to the test: The impact of project-based learning on second graders' social studies and literacy learning and motivation in low-SES school settings. *American Educational Research Journal*, *58*(1), 160–200. doi:10.3102/0002831220929638

Egert, F., Fukkink, R. G., & Eckhardt, A. G. (2018). Impact of in-service professional development programs for early childhood teachers on quality ratings and child outcomes: A meta-analysis. *Review of Educational Research*, *88*(3), 401–433. doi:10.3102/0034654317751918

Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *BMJ*, *315*(7109), 629 LP – 634. doi:10.1136/bmj.315.7109.629

Elleman, A. M., & Compton, D. L. (2017). Beyond comprehension strategy instruction: What's next? *Language, Speech, and Hearing Services in Schools*, *48*(2), 84–91. doi:10.1044/2017_LSHSS-16-0036

Elleman, A. M., Lindo, E. J., Morphy, P., & Compton, D. L. (2009). The impact of vocabulary instruction on passage-level comprehension of school-age children: A meta-analysis. *Journal of Research on Educational Effectiveness*, *2*(1), 1–44. doi:10.1080/19345740802539200

Eller, R. G., Pappas, C. C., & Brown, E. (1988). The lexical development of kindergarteners: Learning from written context. *Journal of Literacy Research*, *20*(1), 5–24. doi:10.1080/10862968809547621

Fisher, Z., & Tipton, E. (2015). *Robumeta: Robust variance meta-regression*. http://cran.rproject.org/web/packages/robumeta/index.html

Fitton, L., McIlraith, A. L., & Wood, C. L. (2018). Shared book reading interventions with English learners: A meta-analysis. *Review of Educational Research*, *88*(5), 712–751. doi:10.3102/0034654318790909

Foorman, B., Beyler, N., Borradaile, K., Coyne, M., Denton, C. A., Dimino, J., . . . Wissel, S. (2016). *Foundational skills to support reading for understanding in kindergarten through 3rd grade (NCEE 2016–4008)*. Washington, DC: National Center for Education Evaluation and Regional Assistance (NCEE), Institute of Education Sciences, U.S. Department of Education.

Fuchs, L. S., Seethaler, P. M., Sterba, S. K., Craddock, C., Fuchs, D., Compton, D. L., . . . Changas, P. (2021). Closing the word-problem achievement gap in first grade: Schema-based word-problem intervention with embedded language comprehension instruction. *Journal of Educational Psychology*, *113*(1), 86–103. doi:10.1037/edu0000467

*Gelzheiser, L. M., Scanlon, D., Vellutino, F., Hallgren-Flynn, L., & Schatschneider, C. (2011). Effects of the interactive strategies approach-extended: A responsive and comprehensive intervention for intermediate-grade struggling readers. *Elementary School Journal*, *112*(2), 280–306. doi:10.1371/journal.pone.0201979

Gersten, R., Haymond, K., Newman-Gonchar, R., Dimino, J., & Jayanthi, M. (2020). Meta-analysis of the impact of reading interventions for students in the primary grades. *Journal of Research on Educational Effectiveness*, *13*(2), 401–427. doi:10.1080/19345747.2019.1689591

Gillespie Rouse, A. G., & Rouse, R. (2019). Third graders' use of writing to facilitate learning of engineering concepts. *Journal of Research in Science Teaching*, *56*(10), 1406–1430. doi:10.1002/tea.21581

Glass, G. V. (1976). Primary, Secondary, and Meta-Analysis of Research. *Educational Researcher*, *5*(10), 3–8. doi:10.3102/0013189X005010003

Goodwin, A. P., & Ahn, S. (2013). A meta-analysis of morphological interventions in English: Effects on literacy outcomes for school-age children. *Scientific Studies of Reading*, *17*(4), 257–285. doi:10.1080/10888438.2012.689791

Graham, S., Kiuhara, S. A., & MacKay, M. (2020). The effects of writing on learning in science, social studies, and mathematics: A meta-analysis. *Review of Educational Research*, *90*(2), 179–226. doi:10.3102/0034654320914744

Graham, S., Liu, X., Aitken, A., Ng, C., Bartlett, B., Harris, K. R., & Holzapfel, J. (2018). Effectiveness of literacy programs balancing reading and writing instruction: A meta-analysis. *Reading Research Quarterly*, *53*(3), 279–304. doi:10.1002/rrq.194

Griffin, J. W. (2021). Calculating statistical power for meta-analysis using metapower. *The Quantitative Methods for Psychology*, *17*(1), 24–39. doi:10.20982/tqmp.17.1.p024

Guo, Y., Wang, S., Hall, A. H., Breit-Smith, A., & Busch, J. (2016). The effects of science instruction on young children's vocabulary learning: A research synthesis. *Early Childhood Education Journal*, *44*(4), 359–367. doi:10.1007/s10643-015-0721-6

Guthrie, J., Wigfield, A., & You, W. (2012). Instructional contexts for engagement and achievement in reading. In S. Christensen, C. Wylie, & A. Reschly (Eds.), *Handbook of research on student engagement* (pp. 601–634). New York, NY: Springer.

Guthrie, J. T., & Wigfield, A. (2000). Engagement and motivation in reading. In M. L. Kamil, P. B. Mosenthal, P. D. Pearson, & R. Barr (Eds.), *Handbook of reading research* (Vol. 3, pp. 403–422). Mahwah, New Jerssey: Erlbaum.

*Guthrie, J. T., Anderson, E., Alao, S., & Rinehart, J. (1999). Influences of concept-oriented reading instruction on strategy use and conceptual learning from text. *The Elementary School Journal*, *99*(4), 343–366. doi:10.1086/461929

*Guthrie, J. T., McRae, A., Coddington, C. S., Lutz Klauda, S., Wigfield, A., & Barbosa, P. (2009). Impacts of comprehensive reading instruction on diverse outcomes of low- and high-achieving readers. *Journal of Learning Disabilities*, *42*(3), 195–214. doi:10.1177/0022219408331039

Guthrie, J. T., McRae, A., & Klauda, S. L. (2007). Contributions of concept-oriented reading instruction to knowledge about interventions for motivations in reading. *Educational Psychologist*, *42*(4), 237–250. doi:10.1080/00461520701621087

Guthrie, J. T., Meter, P., McCann, A. D., Wigfield, A., Bennett, L., Poundstone, C. C., . . . Mitchell, A. M. (1996). Growth of literacy engagement: Changes in motivations and strategies during concept-oriented reading instruction. *Reading Research Quarterly*, *31*(3), 306–332. doi:10.1598/rrq.31.3.5

*Guthrie, J. T., Van Meter, P., Hancock, G. R., Alao, S., Anderson, E., & McCann, A. (1998). Does concept-oriented reading instruction increase strategy use and conceptual learning from text? *Journal of Educational Psychology*, *90*(2), 261–278. doi:10.1037/0022-0663.90.2.261

*Guthrie, J. T., Wigfield, A., Barbosa, P., Perencevich, K. C., Taboada, A., Davis, M. H., . . . Tonks, S. (2004). Increasing reading comprehension and engagement through concept-oriented reading instruction. *Journal of Educational Psychology*, *96*(3), 403–423. doi:10.1037/0022-0663.96.3.403

Guthrie, J. T., Wigfield, A., Humenick, N. M., Perencevich, K. C., Taboada, A., & Barbosa, P. (2006). Influences of stimulating tasks on reading motivation and comprehension. *The Journal of Educational Research*, *99*(4), 232–246. doi:10.3200/JOER.99.4.232-246

Hale, S. L. (2010). Efficacy of thematic units on language and literacy: A collaborative study of a shelter unit intervention with struggling first graders. Sam Houston State University, Texas.

Halvorsen, A., Alleman, J., & Brugar, K. A. (2013). Integration of literacy and social studies. In B. M. Taylor & N. K. Duke (Eds.), *Handbook of effective literacy instruction: Research-based practice K – 8* (pp. 394–416). New York, NY: Guilford Press.

*Hanna, T. T. (2008). *Effects of nonfiction guided interactive read-alouds and think-alouds on fourth grader's depth of content area science vocabulary knowledge and comprehension*. The University of Southern Mississippi. https://aquila.usm.edu/dissertations/1185

Hartzler, D. S. (2000). *A meta-analysis of studies conducted on integrated curriculum programs and their effects on student achievement*. Indiana University, IN.

Hebert, M., Bohaty, J. J., Nelson, J. R., & Brown, J. (2016). The effects of text structure instruction on expository reading comprehension: A Meta-analysis. *Journal of Educational Psychology*, *108*(5), 609–629. doi:10.1037/edu0000082

Hedges, L. V. (2007). Effect sizes in cluster-randomized designs. *Journal of Educational and Behavioral Statistics*, *32*(4), 341–370. doi:10.3102/1076998606298043

Hedges, L. V., & Hedberg, E. C. (2007). Intraclass correlation values for planning group-randomized trials in education. *Educational Evaluation and Policy Analysis*, *29*(1), 60–87. doi:10.3102/0162373707299706

Hedges, L. V., Tipton, E., & Johnson, M. C. (2010). Robust variance estimation in meta-regression with dependent effect size estimates. *Research Synthesis Methods*, *1*(1), 39–65. doi:10.1002/jrsm.5

Hempel, S., Miles, J. N. V., Booth, M. J., Wang, Z., Morton, S. C., & Shekelle, P. G. (2013). Risk of bias: A simulation study of power to detect study-level moderator effects in meta-analysis. *Systematic Reviews*, *2*(1), 107. doi:10.1186/2046-4053-2-107

Hill, H. C., & Erickson, A. (2019). Using implementation fidelity to aid in interpreting program impacts: A brief review. *Educational Researcher*, *48*(9), 590–598. doi:10.3102/0013189X19891436

*Hinde, E. R., Popp, S. E. O., Dorn, R. I., Ekiss, G. O., Mater, M., Smith, C. B., & Libbee, M. (2007). The integration of literacy and geography: The Arizona GeoLiteracy program's effect on reading comprehension. *Theory and Research in Social Education*, *35*(3), 343–365. doi:10.1080/00933104.2007.10473340

*Hinde, E. R., Popp, S. E. O., Jimenez-Silva, M., & Dorn, R. I. (2011). Linking geography to reading and English language learners' achievement in US elementary and middle school classrooms. *International Research in Geographical and Environmental Education*, *20*(1), 47–63. doi:10.1080/10382046.2011.540102

Hirsch, E. D., & Hansel, L. (2013). Why content is king. *Educational Leadership*, *71*(3), 28–33.

Huck, A. (2019). Elementary Social studies content integration in CCLS: An analysis of content integration. *The Social Studies*, *110*(1), 1–16. doi:10.1080/00377996.2018.1524359

Huntley, M. A. (1998). Design and implementation of a framework for defining integrated mathematics and science education. *School Science and Mathematics*, *98*(6), 320–327. doi:10.1111/j.1949-8594.1998.tb17427.x

Hwang, H., Lupo, S. M., Cabell, S. Q., & Wang, S. (2021). What research says about leveraging the literacy block for learning. *Reading in Virginia, 42,*35–47.

Institute of Education Sciences (IES). (2017). *What Works Clearinghouse standards handbook (Version 4.0)*. Washington, DC: Author. https://ies.ed.gov/ncee/wwc/Docs/referenceresources/wwc_standards_handbook_v4.pdf

Kaldenberg, E. R., Watt, S. J., & Therrien, W. J. (2015). Reading instruction in science for students with learning disabilities: A meta-analysis. *Learning Disability Quarterly*, *38*(3), 160–173. doi:10.1177/0731948714550204

Kao, J. C. (2015). *Examining a year-long intervention program to teach expository text structure within social studies content to second-grade students*. Columbia University, New York, NY. https://academiccommons.columbia.edu/catalog/ac:187944

Kendeou, P., Butterfuss, R., Van Boekel, M., & O'Brien, E. J. (2017). Integrating relational reasoning and knowledge revision during reading. *Educational Psychology Review*, *29*(1), 27–39. doi:10.1007/s10648-016-9381-3

*Kim, J. S., Burkhauser, M. A., Mesite, L. M., Asher, C. A., Relyea, J. E., Elmore, J., . . . Relyea, J. E. (2020). Improving reading comprehension, science domain knowledge, and reading engagement through a first-grade content literacy intervention. *Journal of Educational Psychology*. doi:10.1037/edu0000465

*Kim, J. S., Relyea, J. E., Burkhauser, M. A., Scherer, E., & Rich, P. (2021). Improving elementary grade students' science and social studies vocabulary knowledge depth, reading comprehension, and argumentative writing: A conceptual replication. *Educational Psychology Review*. doi:10.1007/s10648-021-09609-6

Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. New York, NY: Cambridge University Press.

Kintsch, W. (2013). Revisiting the Construction-Integration model of text comprehension and its implications for instruction. In D. E. Alvermann, N. J. Unrau, & R. B. Ruddell (Eds.), *Theoretical models and processes of reading* (6th ed., pp. 807–839). Newark, DE: International Reading Association.

Konstantopoulos, S., & Hedges, L. V. (2019). Statistically analyzing effect sizes: Fixed- and random-effects models. In H. Cooper, L. V. Hedges, & J. C. Valentine (Eds.), *The Handbook of research synthesis and meta-analysis* (3rd ed., pp. 245–280). New York, NY: Russell Sage Foundation.

Kraft, M. A. (2020). Interpreting effect sizes of education interventions. *Educational Researcher*, *49*(4), 241–253. doi:10.3102/0013189X20912798

Lammert, C. (2020). Becoming inquirers: A review of research on inquiry methods in literacy preservice teacher preparation. *Literacy Research and Instruction*, *59*(3), 191–217. doi:10.1080/19388071.2020.1730529

Levstik, L. S., & Barton, K. C. (2008). *Researching history education: Theory, method, and context*. New York, NY: Routledge.

*Lohmann. (1963). *A comparative study of the improvement of reading skills of pupils during fourth grade in two types of curriculum organization*. The University of Tulsa, OK.

Lupo, S. M., Berry, A., Thacker, E., Sawyer, A., & Merritt, J. (2019). Rethinking text sets to support knowledge building and interdisciplinary learning. *The Reading Teacher*, 1–12. doi:10.1002/trtr.1869

*Lutz, S. L., Guthrie, J., & Davis, M. H. (2006). Scaffolding for Engagement in Elementary School Reading Instruction. *The Journal of Educational Research*, *100*(1), 3–20. doi:10.3200/JOER.100.1.3-20

Maerten-Rivera, J., Ahn, S., Lanier, K., Diaz, J., & Lee, O. (2016). Effect of a multiyear intervention on science achievement of all students including English language learners. *The Elementary School Journal*, *116*(4), 600–624. doi:10.1086/686250

Martínez, P. (2008). *Impact of an integrated science and reading intervention (INSCIREAD) on bilingual students' misconceptions, reading comprehension and transferability of strategies*. George Mason University, Fairfax, VA. http://ebot.gmu.edu/handle/1920/3179

*Martínez-Álvarez, P., Bannan, B., & Peters-Burton, E. E. (2012). Effect of strategy instruction on fourth-grade dual language learners' ability to monitor their comprehension of scientific texts. *Bilingual Research Journal*, *35*(3), 331–349. doi:10.1080/15235882.2012.734005

Marulis, L. M., & Neuman, S. B. (2010). The effects of vocabulary intervention on young Children's word learning: A meta-analysis. *Review of Educational Research*, *80*(3), 300–335. doi:10.3102/0034654310377087

Marulis, L. M., & Neuman, S. B. (2013). How vocabulary interventions affect young children at risk: A meta-analytic review. *Journal of Research on Educational Effectiveness*, *6*(3), 223–262. doi:10.1080/19345747.2012.755591

Mathur, M. B., & VanderWeele, T. J. (2020). Sensitivity analysis for publication bias in meta-analyses. *Journal of the Royal Statistical Society: Series C*, *69*(5), 1091–1119.

McCarthy, K. S., & McNamara, D. S. (2021). The Multidimensional knowledge in text comprehension framework. *Educational Psychologist*, 1–19. doi:10.1080/00461520.2021.1872379

Meece, J. L., Anderman, E. M., & Anderman, L. H. (2006). Classroom goal structure, student motivation, and academic achievement. *Annual Review of Psychology*, *57*(1), 487–503. doi:10.1146/annurev.psych.56.091103.070258

Mercer, N., Dawes, L., Wegerif, R., & Sams, C. (2004). Reasoning as a scientist: Ways of helping children to use language to learn science. *British Educational Research Journal*, *30*(3), 359–372. doi:10.1080/01411920410001689689

*Morrow, L. M., Pressley, M., Smith, J. K., & Smith, M. (1997). The effect of a literature-based program integrated into literacy and science instruction with children from diverse backgrounds. *Reading Research Quarterly*, *32*(1), 54–76. doi:10.1598/rrq.32.1.4

Moss, B. (2005). Making a case and a place for effective content area literacy instruction in the elementary grades. *The Reading Teacher*, *59*(1), 46–55. doi:10.1598/rt.59.1.5

National Assessment of Educational Progress [NAEP]. (2019). *The nation's report card: Reading*. Washington, DC: National Center for Education Statistics.

National Governors Association Center for Best Practices & Council of Chief State School Officers. (2010). *Common core state standards for english language arts and literacy in history/social studies, science, and technical subjects*. Washington, DC: Authors.

National Institute for Literacy. (2008). *Developing early literacy: Report of the National Early Literacy Panel*. Jessup, MD: Author.

National Institute of Child Health and Human Development. (2000). *Report of the National Reading Panel. Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction (NIH Publication No. 00–4769)*. Washington, DC: U.S. Government Printing Office.

Neuman, S. B., & Dwyer, J. (2011). Developing vocabulary and conceptual knowledge for low-income preschoolers: A design experiment. *Journal of Literacy Research*, *43*(2), 103–129. doi:10.1177/1086296X11403089

*Neuman, S. B., & Kaefer, T. (2018). Developing low-income children's vocabulary and content knowledge through a shared book reading program. *Contemporary Educational Psychology*, *52*, 15–24. doi:10.1016/j.cedpsych.2017.12.001

Neuman, S. B., Newman, E. H., & Dwyer, J. (2011). Educational effects of a vocabulary intervention on preschoolers' word knowledge and conceptual development: A cluster-randomized trial. *Reading Research Quarterly*, *46*(3), 249–272. http://www.jstor.com/stable/41228653

*Neuman, S. B., Samudra, P., & Danielson, K. (2021). Effectiveness of scaling up a vocabulary intervention for low-income children, pre-K through first grade. *The Elementary School Journal*, *121*(3), 385–409. doi:10.1086/712492

Neuman, S. B., & Wright, T. S. (2014). The magic of words: Teaching vocabulary in the early childhood classroom. *American Educator, 38*(2), 4–13.

Ng, J. P.-L. (1970). *The effects of bilingual science instruction on the vocabulary, comprehension achievement and conceptualization of elementary school Chinese children whose second language is English*. Los Angeles, CA: University of California.

Ozuru, Y., Dempsey, K., & McNamara, D. S. (2009). Prior knowledge, reading skill, and text cohesion in the comprehension of science texts. *Learning and Instruction*, *19*(3), 228–242. doi:10.1016/j.learninstruc.2008.04.003

Paprzycki, P., Tuttle, N., Czerniak, C. M., Molitor, S., Kadervaek, J., & Mendenhall, R. (2017). The impact of a Framework-aligned science professional development program on literacy and mathematics achievement of K-3 students. *Journal of Research in Science Teaching*, *54*(9), 1174–1196. doi:10.1002/tea.21400

Pearson, P. D., & Billman, A. K. (2016). Reading to learn science: A right that extends to every reader-expert or novice. In Z. Babaci-Wilhite (Ed.), *Human Rights in Language and STEM Education: Science, Technology, Engineering and Mathematics* (pp. 17–34). Rotterdam, Netherlands: Sense Publishers. doi:10.1007/978-94-6300-405-3.

Pladevall-Ballester, E. (2016). CLIL subject selection and young learners' listening and reading comprehension skills. *International Journal of Applied Linguistics*, *26*(1), 52–74. doi:10.1111/ijal.12079

Pritchard, T. (2019). Analogical cognition: An insight into word meaning. *Review of Philosophy and Psychology*, *10*(3), 587–607. doi:10.1007/s13164-018-0419-y

*Proctor, C. P., Silverman, R. D., Harring, J. R., Jones, R. L.,& Hartranft, A. M., (2020). Teaching bilingual learners: Effects of a language-based reading intervention on academic language and reading comprehension in grades 4 and 5. Reading Research Quarterly, *55*(1), 95–122. https://doi.org/10.1002/rrq.258

Pulido, D.,Pulido, D., Pulido, D.Pulido, D. (2007). The effects of topic familiarity and passage sight vocabulary on L2 lexical inferencing and retention through reading. *Applied Linguistics*, *28*(1), 66–86. doi:10.1093/applin/aml049

Purcell-Gates, V., Duke, N. K., & Martineau, J. A. (2007). Learning to read and write genre-specific text: Roles of authentic experience and explicit teaching. *Reading Research Quarterly*, *42*(1), 8–45. doi:10.1598/RRQ.42.1.1

Rios, J. A., Ling, G., Pugh, R., Becker, D., & Bacall, A. (2020). Identifying critical 21st-century skills for workplace success: A content analysis of job advertisements. *Educational Researcher*, *49*(2), 2. doi:10.3102/0013189X19890600

Romance, N., & Vitale, M. (2017). Implications of a cognitive science model integrating literacy in science on achievement in science and reading: Direct effects in grades 3–5 with transfer to grades 6–7. *International Journal of Science and Mathematics Education*, *15*(6), 979–995. doi:10.1007/s10763-016-9721-2

*Romance, N. R., & Vitale, M. R. (1992). A curriculum strategy that expands time for in-depth elementary science instruction by using science-based reading strategies: Effects of a year-long study in grade four. *Journal of Research in Science Teaching*, *29*(6), 545–554. doi:10.1002/tea.3660290604

Romance, N. R., & Vitale, M. R. (2012a). Expanding the role of K-5 science instruction in educational reform: Implications of an interdisciplinary model for integrating science and reading. *School Science and Mathematics*, *112*(8), 506–515. doi:10.1111/ssm.12000

Romance, N. R., & Vitale, M. R. (2012b). Interdisciplinary perspectives linking science and literacy in grades K–5: Implications for policy and practice. In B. J. Fraser, C. J. McRobbie, & K. Tobin (Eds.), In *Second international handbook of science education* (pp. 1351–1373). Dordrecht, Netherlands: Springer.

*See, B. H., Gorard, S., & Siddiqui, N. (2017). Can explicit teaching of knowledge improve reading attainment? An evaluation of the Core Knowledge curriculum. *British Educational Research Journal*, *43*(2), 372–393. doi:10.1002/berj.3278

Shanahan, T., Callison, K., Carriere, C., Duke, N. K., Pearson, P. D., Schatschneider, C., & Torgesen, J. (2010). *Improving reading comprehension in kindergarten through 3rd grade: IES practice guide.* (NCEE 2010–4038).

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. doi:10.1037/a0014217

Silverman, R. D., Johnson, E., Keane, K., & Khanna, S. (2020). Beyond decoding: A meta-analysis of the effects of language comprehension interventions on K–5 students' language and literacy outcomes. *Reading Research Quarterly*, *55*(S1), S207–S233. doi:10.1002/rrq.346

Simmons, D., Hairrell, A., Edmonds, M., Vaughn, S., Larsen, R., Willson, V., . . . Byrns, G. (2010). A comparison of multiple-strategy methods: Effects on fourth-grade students' general and content-specific reading comprehension and vocabulary development. *Journal of Research on Educational Effectiveness*, *3*(2), 121–156. doi:10.1080/19345741003596890

Slavin, R., & Madden, N. A. (2011). Measures inherent to treatments in program effectiveness reviews. *Journal of Research on Educational Effectiveness*, *4*(4), 370–380. doi:10.1080/19345747.2011.558986

Snow, C.E. (2002). *Reading for understanding: Toward an R&D program in reading comprehension*. Santa Monica, CA: Rand Corporation.

Stahl, S. A., Hare, V. C., Sinatra, R., & Gregory, J. F. (1991). Defining the role of prior knowledge and vocabulary in reading comprehension: The retiring of number 41. *Journal of Reading Behavior*, *23*(4), 487–508.

Stanovich, K. E., & Cunningham, A. E. (1993). Where does knowledge come from? Specific associations between print exposure and information acquisition. *Journal of Educational Psychology*, *85*(2), 211–229. doi:10.1037/0022-0663.85.2.211

*Stephens, K. E. (2007). *The impact of a science-based integrated instructional protocol on the motivation, reading comprehension, and science achievement of fourth and fifth graders*. Texas A&M University-Commerce, TX.

Steyvers, M., & Tenenbaum, J. B. (2005). The large-scale structure of semantic networks: Statistical analyses and a model of semantic growth. *Cognitive Science*, *29*(1), 41–78. doi:10.1207/s15516709cog2901_3

Stoddart, T., Pinal, A., Latzke, M., & Canaday, D. (2002). Integrating Inquiry Science and Language Development for English Language Learners. *Journal of Research in Science Teaching*, *39*(8), 664–687. doi:10.1002/tea.10040

Swanborn, M. S. L., & De Glopper, K. (1999). Incidental word learning while reading: A meta-analysis. *Review of Educational Research*, *69*(3), 261–285. doi:10.3102/00346543069003261

Swanson, E., Hairrell, A., Kent, S., Ciullo, S., Wanzek, J. A., & Vaughn, S. (2014). A synthesis and meta-analysis of reading interventions using social studies content for students with learning disabilities. *Journal of Learning Disabilities*, *47*(2), 178–195. doi:10.1177/0022219412451131

Swanson, E., Vaughn, S., Wanzek, J., Petscher, Y., Heckert, J., Cavanaugh, C., . . . Tackett, K. (2011). A synthesis of read-aloud interventions on early reading outcomes among preschool through third graders at risk for reading difficulties. *Journal of Learning Disabilities*, *44*(3), 258–275. doi:10.1177/0022219410378444

Teale, W. H., Paciga, K. A., & Hoffman, J. L. (2007). Beginning reading instruction in urban schools: The curriculum gap ensures a continuing achievement gap. *The Reading Teacher*, *61*(4), 344–348. doi:10.1598/rt.61.4.8

Thibaut, L., Knipprath, H., Dehaene, W., & Depaepe, F. (2018). The influence of teachers' attitudes and school context on instructional practices in integrated STEM education. *Teaching and Teacher Education*, *71*, 190–205. doi:10.1016/j.tate.2017.12.014

*Tong, X., Deacon, S. H., & Cain, K. (2014). Morphological and syntactic awareness in poor comprehenders: Another piece of the puzzle. *Journal of Learning Disabilities*, *47*(1), 22–33. doi:10.1177/0022219413509971

Truckenmiller, A. J., Park, J., Dabo, A., & Wu Newton, Y. C. (2019). Academic language instruction for students in grades 4 through 8: A literature synthesis. *Journal of Research on Educational Effectiveness*, *12*(1), 135–159. doi:10.1080/19345747.2018.1536773

Tyner, A., & Kabourek, S. (2020). *Social studies instruction and reading comprehension: Evidence from the Early Childhood Longitudinal Study*. Washington, DC: Thomas B. Fordham Institute. https://fordhaminstitute.org/national/resources/social-studies-instruction-andreading-comprehension

Valentine, J. C., Aloe, A., & Wilson, S. (2019). Interpreting effect sizes. In H. Cooper, L. Hedges, & J. C. Valentine (Eds.), *The Handbook of research synthesis and meta-analysis* (3rd ed., pp. 433–452). New York, NY: Russell Sage Foundation.

Valentine, J. C., Pigott, T. D., & Rothstein, H. R. (2010). How many studies do you need? A primer on statistical power for meta-analysis. *Journal of Educational and Behavioral Statistics*, *35*(2), 215–247. doi:10.3102/1076998609346961

van den Broek, P., Kendeous, P., Kremer, K., Lynch, J., Butler, J., White, M. J., & Lorch, E. P. (2005). Assessment of comprehension abilities in young children. In S. G. Paris & S. A. Stahl (Eds.), *Children's reading comprehension and assessment* (pp. 107–130). Mahwah, NJ: Lawrence Erlbaum Associates.

*Vitale, M. R., & Romance, N. R. (2011). Adaptation of a knowledge-based instructional intervention to accelerate student learning in science and early literacy in grades 1 and 2. *Journal of Curriculum and Instruction*, *5*(2), 79–93. doi:10.3776/joci.2011.v5n2p79-93

*Vitale, M. R., & Romance, N. R. (2012). Using in-depth science instruction to accelerate student achievement in science and reading comprehension in grades 1–2. *International Journal of Science and Mathematics Education*, *10*(2), 457–472. doi:10.1007/s10763-011-9326-8

Wang, J. H.-Y., & Guthrie, J. T. (2004). Modeling the effects of intrinsic motivation, extrinsic motivation, amount of reading, and past reading achievement on text comprehension between U.S. and Chinese students. *Reading Research Quarterly*, *39*(2), 162–186. doi:10.1598/RRQ.39.2.2

Wanzek, J., Vaughn, S., Scammacca, N., Gatlin, B., Walker, M. A., & Capin, P. (2016). Meta-analyses of the effects of tier 2 type reading interventions in grades K-3. *Educational Psychology Review*, *28*(3), 551–576. doi:10.1007/s10648-015-9321-7

Wanzek, J., Vaughn, S., Scammacca, N. K., Metz, K., Murray, C. S., Roberts, G., & Danielson, L. (2013). Extensive reading interventions for students with reading difficulties after grade 3. *Review of Educational Research*, *83*(2), 163–195. doi:10.3102/0034654313477212

Wanzek, J., Wexler, J., Vaughn, S., & Ciullo, S. (2010). Reading interventions for struggling readers in the upper elementary grades: A synthesis of 20 years of research. *Reading and Writing: An Interdisciplinary Journal*, *23*(8), 889–912. doi:10.1007/s11145-009-9179-5

Weiser, B., & Mathes, P. (2011). Using encoding instruction to improve the reading and spelling performances of elementary students at risk for literacy difficulties: A best-evidence synthesis. *Review of Educational Research*, *81*(2), 170–200. doi:10.3102/0034654310396719

Wexler, N. (2019). *The knowledge gap: The hidden cause of America's broken education system–and how to fix it*. New York: NY: Avery.

Wigfield, A., Guthrie, J. T., Tonks, S., & Perencevich, K. C. (2004). Children's motivation for reading: Domain specificity and instructional influences. *The Journal of Educational Research*, *97*(6), 299–310. doi:10.3200/JOER.97.6.299-310

*Williams, J. P., Hall, K. M., Lauer, K. D., Stafford, K. B., DeSisto, L. A., & DeCani, J. S. (2005). Expository text comprehension in the primary grade classroom. *Journal of Educational Psychology*, *97*(4), 538–550. doi:10.1037/0022-0663.97.4.538

*Williams, J. P., Kao, J. C., Pao, L. S., Ordynans, J. G., Atkins, J. G., Cheng, R., & DeBonis, D. (2016). Close analysis of texts with structure (CATS): An intervention to teach reading comprehension to at-risk second graders. *Journal of Educational Psychology*, *108*(8), 1061–1077. doi:10.1037/edu0000117

*Williams, J. P., Nubla-Kung, A. M., Pollini, S., Stafford, K. B., Garcia, A., & Snyder, A. E. (2007). Teaching Cause – Effect Text Structure Through Social Studies Content to At-Risk Second Graders. *Journal of Learning Disabilities*, *40* (2), 111–120. doi:10.1177/00222194070400020201

*Williams, J. P., Pollini, S., Nubla-Kung, A. M., Snyder, A. E., Garcia, A., Ordynans, J. G., & Grant Atkins, J. (2014). An intervention to improve comprehension of cause/effect through expository text structure instruction. *Journal of Educational Psychology*, *106*(1), 1–17. doi:10.1037/a0033215

*Williams, J. P., Stafford, K. B., Lauer, K. D., Hall, K. M., & Pollini, S. (2009). Embedding reading comprehension training in content-area Instruction. *Journal of Educational Psychology*, *101*(1), 1–20. doi:10.1037/a0013152

Woodall, S. R. (2003). *Does inquiry based social studies instruction make a difference for fifth and sixth grade students: A study of affect, critical thinking, and reading comprehension*. Texas A&M University-Commerce, TX.

Wright, T. S., & Cervetti, G. N. (2016). A systematic review of the research on vocabulary instruction that impacts text comprehension. *Reading Research Quarterly*, 1–24. doi:10.1002/rrq.163

*Wright, T. S., & Gotwals, A. W. (2017). Supporting kindergartners' science talk in the context of an integrated science and disciplinary literacy curriculum. *The Elementary School Journal*, *117*(3), 513–537. doi:10.1086/690273

Xanthou, M. (2011). The impact of CLIL on L2 vocabulary development and content knowledge. *English Teaching*, *10* (4), 116–126.

Zahorik, J. A. (1996). Elementary and Secondary Teachers' Reports of How They Make Learning Interesting. *The Elementary School Journal*, *96*(5), 562–564. doi:10.1086/461844