Writing to Read: A Meta-Analysis of the Impact of Writing and Writing Instruction on Reading

STEVE GRAHAM
MICHAEL HEBERT
Vanderbilt University

Reading is critical to students' success in and out of school. One potential means for improving students' reading is writing. In this meta-analysis of true and quasi-experiments, Graham and Herbert present evidence that writing about material read improves students' comprehension of it; that teaching students how to write improves their reading comprehension, reading fluency, and word reading; and that increasing how much students write enhances their reading comprehension. These findings provide empirical support for long-standing beliefs about the power of writing to facilitate reading.

Reading is one of the most critical skills that students must master to be successful educationally, occupationally, and socially. Students' educational success depends on their abilities to read and critically analyze information presented in textbooks and other classroom materials (Berman, 2009; Klein, 1999). Reading is essential to success in most white-collar and blue-collar jobs (Greene, 2000), with forecasters predicting an increase in the proportion of new jobs requiring strong reading skills (Carnevale & Derochers, 2004; Kirsch, Braun, Yamamoto, & Sum, 2007). Reading is part of the basic fabric of twenty-first-century life, as e-mailing, blogging, texting, Facebook, and other forms of written text are now common means for social contact and communication. Written text permeates everyday life, from messages prominently displayed on billboards and sides of buses to information provided on everyday essentials, such as cans of food and bottles of medicine.

Despite the importance of reading, many students are not skilled readers by the end of high school. The 2009 National Assessment of Educational Progress (NAEP) reported that only 38 percent of twelfth-grade students performed at

Harvard Educational Review Vol. 81 No. 4 Winter 2011

or above the "proficient" level in reading (defined as solid academic performance) (NCES, 2010). In terms of younger students, only 33 percent of fourth graders and 32 percent of eighth graders performed at these levels (NCES, 2009). In contrast, 34, 43, and 36 percent of fourth-, eighth-, and twelfth-grade students, respectively, scored at the "basic" level, denoting only partial mastery of the literacy skills needed at their grade levels. The rest of the tested students' scores were below this basic level.

Furthermore, the reading performance of students who do not speak English as their first language, students who have a disability, and students who are black, Hispanic, or Native American was significantly lower than the reading performance of students who were native English speakers, did not have a disability, or were white, respectively (NCES, 2009, 2010). The results from the NAEP clearly demonstrate that large numbers of students need help if they are to become skilled readers.

Such concerns have spurred large-scale policy actions during this decade and the last to improve children's reading, including No Child Left Behind and Reading First. Both of these approaches relied heavily on the use of scientifically based practices (involving the teaching of phonological awareness, phonics, vocabulary, fluency, and comprehension) identified by the National Reading Panel (NRP) (NICHD, 2000). Neither of these policy endeavors resulted in the types of reading gains envisioned by their advocates. While there are many possible explanations for these outcomes, one contributing factor may be that the instructional practices identified by the NRP report were too narrow and not complete.

An important policy question, then, is what else can schools do to strengthen students' reading? There are many potential actions that policy makers and educators can undertake to improve reading, ranging from enhancing children's overall language skills to providing more engaging and meaningful reading instruction. In this article, we examine the effectiveness of writing as a tool for improving students' reading. We concentrated our efforts on this often-overlooked tool for two basic reasons. One, several meta-analyses have found that writing about content classroom material can facilitate learning it (Bangert-Drowns, Hurley, & Wilkinson, 2004; Graham & Perin, 2007a). It is also possible that writing about material read enhances comprehension of it. Two, reading and writing share a close and reciprocal relationship (Fitzgerald & Shannahan, 2000), and there is evidence that reading instruction can improve students' writing skills (e.g., Graham, 2000; Krashen, 1989). Consequently, it is likely that writing instruction in turn improves students' reading skills. Three theoretical perspectives are particularly informative in understanding the possible impact of writing on reading.

According to the functional view of reading-writing connections (Fitzgerald & Shanahan, 2000), writing about text should facilitate comprehension in five ways (Applebee, 1984; Emig, 1977; Klein, 1999; Stotsky, 1982):

1. It fosters explicitness, as the writer must select which information in text is most important.
2. It is integrative, as it encourages the writer to organize ideas from text into a coherent whole, establishing explicit relationships among the ideas.
3. It facilitates reflection, as the permanence of writing makes it easier to review, reexamine, connect, critique, and construct new understandings of text ideas.
4. It can foster a personal involvement with text, as it requires active decision making about what will be written and how it will be treated.
5. It involves transforming or manipulating the language of text so that writers put ideas into their own words, making them think about what the ideas mean.

In short, writing about text should facilitate comprehending it, as it provides students with a tool for visibly and permanently recording, connecting, analyzing, personalizing, and manipulating key ideas in text.

The impact of writing on reading likely extends beyond just writing about text to the possible impact of teaching about the process of writing. According to the shared knowledge view of reading-writing connections, reading and writing are not identical skills, but both rely on common knowledge and processes (Fitzgerald & Shanahan, 2000). Consequently, instruction that improves writing skills and processes should improve reading skills and processes. We illustrate this with two examples. One, Ehri (2000) and others (e.g., Moats 2005/2006) hypothesize that teaching students how words are spelled provides them with schemata about specific connections between letters and sounds, and this should make it easier for them to identify and remember words in text containing these connections. Two, Neville and Searls (1991) hypothesize that teaching students how to construct more complex sentences by combining smaller, less complex ones should result in greater skill in understanding such units in reading.

According to the rhetorical relations view of reading-writing connections (Tierney & Shanahan, 1991), the process of composing text should enhance one's skills at comprehending text. This theoretical perspective views reading and writing as communication activities, and it assumes that writers gain insights about reading by creating text for an audience to read, even if the writer is the intended audience. Theoretically, the process of creating text should prompt students to be more thoughtful and engaged when reading text produced by others. Because writers need to make their assumptions and premises explicit as well as observe the rules of logic when composing text, this presumably makes them more aware of these same issues in the material they read.

To examine the robustness of each of these theoretical viewpoints about the impact of writing on reading, we conducted a meta-analysis to answer the following three questions for students in grades 1–12:

1. Does writing about material read enhance students' comprehension of text?
2. Does writing skills instruction strengthen students' reading skills?
3. Does increasing how much students write improve how well they read?

Meta-analysis is used to summarize the direction and magnitude of the effects obtained in a set of empirical studies examining the same basic phenomena (Lipsey & Wilson, 2001). The meta-analysis reported in this article was funded by the Carnegie Corporation of New York (Graham & Hebert, 2010), and the evidence used to answer each question was derived from either true or quasi-experiments. In both types of experiments, a writing treatment (e.g., writing a summary of text read) is compared to a control condition (e.g., reading and rereading text) to determine its impact on reading (e.g., reading comprehension). Meta-analysis is well suited to answering the three questions posed in this review, as it provides an estimate of the effectiveness of a treatment "under conditions that typify studies in the literature" (Bangert-Drowns et al., 2004, p. 34). When enough studies are available, and the variability of individual study effects is greater than the variability due to sampling error alone, meta-analysis also permits examining the relationship between study outcomes and features. In other words, conducting a meta-analysis on a large number of studies with different findings may enable us to more closely investigate the relationships between a specific set of writing procedures (e.g., writing about text) and students' grade levels, type of text read, and so forth.

This meta-analysis differs from *Writing to Read* (Graham & Hebert, 2010) in four important ways. First, we calculated effects for all quasi-experiments included in the review and adjusted these effects for possible data clustering due to hierarchical nesting of data (i.e., researchers assigned classes to treatment or control conditions but then examined student-level effects). We included both true experiments and quasi-experimental studies in our review, as a sensitivity analysis revealed that statistically similar results were obtained for both types of experiments. Second, we assessed the quality of each study, allowing us to make better judgments about the confidence that can be placed in our conclusions. Third, we updated the search for studies to include studies published after June 2008 (this involved updating the 260 original electronic searches). And fourth, we applied meta-regression to examine moderating effects of study characteristics. This procedure allowed us to examine the unique contribution of individual variables (e.g., grade) in accounting for variability in study effects, after variability due to other variables (e.g., participant training and study quality) were first controlled. While the central results and recommendations of *Writing to Read* remained the same, our use of current meta-analytic procedures allowed us to draw more nuanced conclusions and increased our confidence in the findings.

We anticipated that writing about reading would enhance students' comprehension of text, that writing instruction would improve students' reading skills, and that increasing how much students wrote would improve their reading. These predictions were based on the three theoretical views of reading-writing connections discussed earlier. We further anticipated that the variability of study effects would exceed the anticipated variability due to sampling error, at least for some of our analyses. We thought this was especially likely for Question 1 (Does writing about material read enhance students' comprehension of text?), as there are many different ways to write about text. Thus, in addition to conducting an overall analysis of the effects of writing about text, we examined the effects of specific types of writing (e.g., summary writing). When excessive variability was evident in effects for Question 1, and enough studies were available to conduct a meta-regression (for all studies in general or a particular type of writing), we were particularly interested in the unique and mediating influence of four variables: grade (as text becomes increasingly difficult with grade; Baker, Dreher, & Guthrie, 2000), subject area (as the impact of writing about text may differ by content; Bangert-Drowns et al., 2004), participant training (as writing about text may be more effective when students are taught how to do this; Graham & Perin, 2007a), and study quality (Lipsey & Wilson, 2001).

It should be noted that participant training in Question 1 involved teaching students how to use a specific writing activity as a tool for understanding material read, whereas writing instruction in Question 2 (Does writing skills instruction strengthen students' reading skills?) was not directly connected to reading or reading instruction (instruction here was focused on learning how to write). Both Questions 2 and 3 examined if the impact of writing instruction or increased writing generalized to reading.

Method

Study Inclusion and Exclusion Criteria

Studies had to meet the following six criteria to be included in this review:

1. Was a true experiment (assignment to conditions is random) or a quasi-experiment (assignment to conditions is not random)
2. Involved a treatment group that wrote about what they read, were taught to write, or increased how much they wrote
3. Included at least one reading measure that assessed the impact of the writing treatment or condition; quasi-experimental studies had to include a comparable pretest reading measure since students were not randomly assigned to conditions
4. Involved students in grades 1–12
5. Was published in English
6. Contained the statistics necessary to compute a weighted effect size (or statistics were obtained from the authors)

We excluded studies for the following four reasons. One was if the writing treatment or condition did not involve the creation of meaningful text. Consequently, studies where students copied text verbatim, practiced typing, wrote single words, and added a missing word to a sentence were excluded. The only exception involved studies where the writing treatment involved teaching spelling. Numerous literacy theorists claim that spelling instruction enhances word reading skills (e.g., Adams, 1990; Ehri, 1987), and such instruction commonly involves copying or writing individual words (Graham et al., 2008).

We also excluded studies if the control condition wrote or received writing instruction. There were two exceptions to this rule. We did include a study if treatment and control students received the same amount of writing or writing instruction as part of their typical language arts program, but the experimental manipulation for the writing condition involved additional writing or writing instruction. We also included a study if students in the control condition copied text (this never occurred) or completed written cloze activities (this occurred once).

We also excluded studies if the writing treatment was tested in a special school for students with disabilities (e.g., school for the deaf), as the moderating factors to be controlled in such settings warrant special care/attention beyond the scope of this meta-analysis. And, finally, we excluded studies if the only reading outcome measure was identical to the writing treatment, as the treatment and assessment of it could not be separated.

Search Strategies Used to Locate Studies

To identify possible studies for this review, we conducted 260 electronic searches (ending in January 2010) in four databases: ERIC, PsychINFO, Education Abstracts, and ProQuest (including Dissertation Abstracts International). We both read each item identified in these searches (more than fourteen thousand). If the item looked promising based on its abstract or title, we obtained it (agreement between us was 99.2 percent, with disagreements resolved by Graham). We hand-searched the following peer-reviewed journals: *Assessing Writing, Journal of Literacy, Reading and Writing: An International Journal, Reading Research and Instruction, Reading Research Quarterly, Research in the Teaching of Writing, Scientific Studies of Reading,* and *Written Communication.* Other sources for locating studies included the report from the National Reading Panel, *Teaching Children to Read* (NICHD, 2000), as well as chapters examining the relationship between writing and reading in influential books, such as the *Handbook of Reading Research* (Kamil, Mosenthall, Pearson, & Barr, 2000). Once we obtained a document, we searched the reference list to identify other studies. Of 752 documents collected, 95 experiments met the inclusion criteria. Through our independent reading, we established reliability for selected documents, with only three disagreements (reliability = 99.6%), which we resolved through discussion.

Categorizing Studies According to Questions and Methods

We read each study and placed it into a category based on the question it answered. Next, we further examined studies assigned to a specific question and placed them into pre-identified instructional subcategories. For Question 1 (impact of writing about text read), these categories were:

1. Answering questions in writing (writing short answers to questions about text before, during, or after reading it) or generating questions in writing to ask about text
2. Taking written notes about text during or after reading it; notes could be unstructured or organized via an outline, graphic organizer, column method, and so forth
3. Summarizing, or self-generating written synthesis of text read as well as summaries written using a specific format or with a set of specific rules
4. Extended writing, or written responses that extended beyond single statements in response to questions, notes, or summaries and focused on students' personal reactions to material read; analysis, interpretation, or application of the material presented in text; or explaining the text material to others
5. Writing short responses about text read (brief analogy, metaphor, and compare/contrast statement)

All Question 1 studies had a reading comprehension outcome measure.

We initially placed studies assigned to Question 2 (impact of writing instruction) in two categories: process writing (an approach that involves creating a supportive writing environment where instruction is typically personalized and students engage in cycles of planning, translating, and revising text; see Graham & Perin, 2007b, for a more complete definition) and skills instruction (writing skills taught systematically). We modified the categorical structure of Question 2, however, as few studies (k = 3; 14%) actually assessed the impact of process writing instruction on reading. Based on a subsequent reanalysis of the studies, we decided to categorize investigations by their impact on reading outcomes. This resulted in three categories that assessed the impact of: (1) process writing, text structure, and paragraph/sentence skills instruction on reading comprehension; (2) sentence/spelling instruction on reading fluency; and (3) spelling instruction on word reading skills.

Question 3 (impact of extra writing) contained a single category: studies that increase the amount of student writing. Reading comprehension was the outcome measure in all of these studies.

We used a variety of outcome measures to assess reading outcomes across the three questions, including researcher-devised and standardized norm-referenced tests. For example, researcher-devised measures of reading comprehension included answering questions about text (multiple choice and short answers), retelling what was read (orally or in writing), summarizing text read in one sentence, and identifying words systematically omitted from text (cloze procedure). No single reading measure was used by a majority of investigators for the three reading constructs studied (comprehension, fluency, and word reading), and multiple measures for the same construct were applied in many studies.

Study Feature Coding

We coded each study for study descriptors, quality indicators, and variables necessary to calculate effect sizes. Study descriptors included: grade, type of student (e.g., struggling writers, English language learners, etc.), number of participants, locale, treatment length, participant training, description of the treatment, description of the control condition, subject, genre, outcome measures, publication type, and research design.

There were eleven quality indicators:

1. Random assignment of participants to conditions
2. Total attrition of less than 10 percent of total sample
3. Total attrition of less than 10 percent and equal attrition across conditions (i.e., conditions did not differ by more than 5 percent)
4. Control condition specifically described
5. Treatment fidelity greater than .80
6. Teacher effects controlled (e.g., random assignment of teachers)
7. More than a single teacher carrying out each condition
8. Reliability of reading measure established and greater than .60
9. Posttest ceiling/floor effects for reading measure not evident (more than one standard deviation [SD] from floor and ceiling)
10. Pretest ceiling/floor effects not evident for reading posttest measure in quasi-experiments (more than one SD from floor and ceiling)
11. Pretest equivalence evident for pretest reading measure in quasi-experiments (i.e., conditions did not differ by more than 0.5 SD)

Each quality indicator was scored as 1 (met) or 0 (not met), except for quality indicators related to measures (8–11 were scored for the proportion of measures that met the criteria). A total score was calculated for each study (nine possible points for true experiments and eleven possible points for quasi-experiments). This was converted to a proportion by dividing obtained score by total possible points.

Independently, we both completed coding for study descriptors and quality indicators, as well as calculation of effect sizes. We resolved disagreements by reexamining the study. Our initial agreement was 94.8 percent for all variables.

Calculation of Effect Sizes

For true experiments, an effect size (ES) was calculated by subtracting the mean score of the treatment group at posttest ($\overline{X}_T$) from the mean score of the control group at posttest ($\overline{X}_C$) and dividing by the pooled standard deviation of the two groups (s_p). For quasi-experiments, we used the same procedure, except we first adjusted pretest differences between treatment and control conditions by subtracting mean pretest score for each group from their mean posttest score. We divided the difference between the adjusted means for treatment and control by the pooled standard deviation for posttest, as recommended by Lipsey and Wilson (2001). For some quasi-experiments, it was necessary to estimate the posttest ES from covariate-adjusted posttest means. In a few instances, we had to compute an ES separately for pretest and posttest as the pretest and posttest used different measures to assess the same construct (e.g., reading comprehension); we obtained an adjusted ES by subtracting pretest ES from posttest ES.

As a prelude to calculating some ESs, it was sometimes necessary to average the performance of two or more groups in each condition (e.g., statistics were reported separately by grade). We did this by using procedures recommended by Nouri and Greenberg (as cited in Cortina & Nouri, 2000). It was also necessary in some cases to estimate missing means and standard deviations from the statistics reported by the authors. For example, we calculated missing standard deviations by estimating residual sums of squares to compute a root mean squared error (RMSE) (e.g., Shadish, Robinson, & Congxiao, 1999). For covariate or complex factorial designs, we estimated pooled standard deviations by calculating and restoring the variance explained by covariates and other "off-factors" to the study's error term and recalculating the pooled standard deviation from the composite variance.

The quasi-experiments in this review assigned whole classes to treatment or control conditions and then examined student-level effects. Adjusting standard errors (SE) for these studies was necessary, as a portion of the total variance in such quasi-experiments was likely due to grouping or clustering within treatments, with the total variance representing a sum of group and student variances. To correct effect variance, we imputed an estimate of the ratio of between-group variance to total variance (i.e., intraclass correlation [ICC]; see Hedges, 2007) from ICC estimates provided by Hedges and Hedberg (2007). We based these estimates on the findings from national intervention studies where the reading comprehension outcomes were adjusted for pretest covariates. If an adjusted ICC was not available (this was the case for several grades), we applied an unadjusted ICC for that grade. If an unadjusted ICC for a grade was not available, we based ICC on an average of the adjusted ICCs for the grade levels adjacent to it. If a quasi-experiment in this review included more than one grade level, we averaged the ICCs for those grades. Most quasi-experiments had equal sample sizes across clusters, and we assumed equal cluster sizes when this information was not provided. We adjusted all computed effects for true and quasi-experiments for small-sample-size bias using the formula

$d_{adj} = d * \gamma$

where $\gamma = 1 - 3/4\ (n_{tx} + n_{ctrl}) - 9$ (Hedges, 1982)

Because using multiple ESs from the same study violates the assumption of independent data points fundamental to most statistical procedures (Wolf, 1986), we used three procedures to preserve the statistical independence of ESs in this review. First, if a study contained more than one ES for a single reading construct (e.g., reading comprehension), we combined these effect sizes as an unweighted average effect. Such aggregation is preferable when intercorrelations among measures are unknown (which was the case in this review), as standard error estimation is complicated when this information is missing (Gleser & Olkin, 1994).

Second, we separated ESs for different constructs (e.g., reading comprehension and word reading) for stratified analyses (i.e., they were not aggregated together in any statistical analysis). We did this also for researcher-designed and standardized norm-referenced measures of the same construct, except for analyses involving the effects of writing instruction on reading fluency and word reading (Question 2). Both of these analyses involved a small number of studies, and the researcher-devised and norm-referenced measures for each construct were similar.

Third, if a study had multiple treatment or control conditions, then we selected only one treatment comparison for inclusion in any analysis (i.e., if the study included a full version and a partial version of the treatment, then we compared only the full version to the control condition). There were some exceptions to this rule. One, in papers reporting multiple studies (Barton, 1930; Doctorow, Wittrock, & Marks, 1978; Vidal-Abarca & Gilabert, 1995), we computed an ES for each study. Two, in studies where one version of a treatment was compared to a control condition and another version of a treatment was compared to a separate control condition (Denner, 1987; Slater, 1982), we computed an ES for each treatment-control comparison. Three, in investigations where two different treatments (e.g., process writing and skills instruction; Kelley, 1984) or two different dosages of a treatment were compared to a single control (Weber & Henderson, 1989), we computed a separate ES for each comparison.

Statistical Analysis of Effect Sizes

Our meta-analysis employed a weighted random-effects model. For each of the three research questions, we calculated an average weighted ES (weighted to take into account sample size by multiplying each ES by its inverse variance). We did not calculate an average weighted ES unless there were at least four studies addressing the question or subquestion. We also calculated a confidence interval and statistical significance of the obtained weighted ES, as well as two measures of homogeneity (Q and I^2). The homogeneity measures allowed us to determine if variability in the ESs for an average weighted effect was larger than expected based on sampling error alone. When this was the case, and there were at least sixteen ESs, we conducted meta-regression moderator analysis to determine if this excess variability could be accounted for by identifiable differences between studies (e.g., training versus no training).

The meta-regression involved a mixed-effects model with maximum likelihood estimates using macros developed for SPSS. Meta-regression is similar to traditional regression analysis, where the contribution of specific variables to the prediction of the target outcome measure (in this case, ESs for individual studies) is estimated (see Konstantopoulous & Hedges, 2009). We assumed that in addition to a random effect due to sampling error, there was a systematic component to the variance among studies. The macros added a random effect variance component and recalculated the inverse variance weight before refitting the model (Lipsey & Wilson, 2001). We entered the predictor variables as a single block.

Results

Table 1 contains information about each study for Questions 1–3. Within each question, studies are arranged from early to later grades. The following information is reported: grade(s), type of student, treatment and control comparison, training for treatment, subject area, genre of writing, total quality of study score, and ES (an asterisk denotes an ES from a norm-referenced test). For studies that contain multiple treatments, we give the overall ES as well as the ES for each specific treatment (bracketed).

Quality of Research

For Questions 1–3, the proportion of studies meeting each of the eleven quality indicators is presented in table 2. While random assignment occurred in just 48 percent of all studies, most quasi-experiments evidenced pretest equivalence. We decided to retain quasi-experiments in this meta-analysis even though there was a relatively large number of true experiments, because there was no statistically significant difference in average weighted effect sizes for true and quasi-experiments ($p = .97$; ESs for true and quasi-experiments were each 0.50).

Pretest ceiling/floor effects were evident for 42 and 50 percent of measures for Questions 2 and 3 studies, respectively. Most posttest measures (73%) did not evidence ceiling/floor effects, and we established that acceptable reliability was evident for close to two-thirds of posttest measures. We were able to define the control conditions in most studies (82%), but rarely did we find that researchers established treatment fidelity (22%), and there was only one teacher per condition in 49 percent of studies (this was especially problematic for Question 1). For studies in Questions 1 and 2, attrition and teacher effects were not major issues, as 70 percent of studies met both attrition criteria and teacher effects were controlled in 69 percent of studies (attrition and teacher effects were major concerns for Question 3). The mean proportion of quality indicators met for studies in Questions 1 and 2 were .63 each (SDs = 0.13 and 0.19, respectively) and .55 for Question 3 (SD = 0.22).

TABLE 1 *Studies and effect sizes included in the analysis of each research question*

Study	*Grade*	*Students*	*Treatment*	*Training*	*Content*	*Genre*	*Quality*	*Measure*	*ES*
Question 1: Effect of writing about reading									
Adams-Boating, 2001	2	FR	EW vs. RI	NT	LA	N	0.55	C	1.02
Denner et al., 1989	2	FR	SW vs. NW	NT	LA	N	0.89	C	0.61
Cohen, 1983	3	A & BA	QG vs. BAU-RI	T	LA	N	0.78	C	*0.75
MacGregor,1988	3	A & AA	QGA vs. R	NT	LA	E	0.44	C	*0.34
Bayne, 1984	3 & 4	FR	NT-S vs. RI	T	LA	N	0.64	C	*0.14
Jenkins et al., 1987	3–6	LD	SUM vs. BAU-RI	T	LA	E	0.73	C	0.68
Jaekyung et al., 2008	4 & 5	FR	EW vs. R	T	LA	N & E	0.67	C	*0.36
Saunders & Goldenberg, 1999	4 & 5	FR & ELL	EW vs. R+ST	NT	LA	E	0.67	C	0.08
Chang & Sung, 2002	5	FR	NT-S vs. R	T/NT	SCI	E	0.70	C	0.49
Jennings, 1990	5	FR	MULT {SUM vs. BAU-RI} {EW vs. BAU}	T NT	SS	E	0.55	C	0.88 {0.34} {1.83}
Newlun, 1930	5	FR	SUM vs. R+ST	T	SS	E	0.64	C	*0.36
Linden & Wittrock, 1981	5	FR & ELL	SS vs. R, D, & RI	NT	LA	N	0.67	C	0.92
Amuchie, 1983	5 & 6	ELL	SUM vs. BAU-RI	T	LA	N & E	0.45	C	1.37
Leshin, 1989	5 & 6	FR	NT-U vs. U	NT	SCI	E	0.89	C	0.43
Berkowitz, 1986	6	GR & PR	MULT {QA vs. R+RR} {NT-S vs. R+RR}	T	SS	E	0.64	C	0.63 {0.35} {0.87}
Coffman, 1992	6	FR	QA vs. R	NT	NR	N	0.83	C	0.32
Taylor & Berkowitz, 1980	6	A & AA	MULT {QA vs. R+DT} {SUM vs. R+DT}	NT T	SS	E	0.61	C	0.32 {0.26} {0.43}
Copeland, 1987	6	GW & PW	EW vs. R+RR	NT	LA	E	0.67	C	1.27
Doctorow et al., 1978	6	GW	SUM vs. R	NT	NR	N	0.89	C	1.56
Doctorow et al., 1978	6	PW	SUM vs. R	NT	NR	N	0.89	C	0.98

(continued)

TABLE 1 *Studies and effect sizes included in the analysis of each research question (continued)*

Study	*Grade*	*Students*	*Treatment*	*Training*	*Content*	*Genre*	*Quality*	*Measure*	*ES*
Keown, 2008	6	FR	NT-S vs. R+D	T	SCI	E	0.73	C	0.44
Rinehart et al., 1986	6	FR	SUM vs. R+WS	T	SS	E	0.73	C	0.49
Vidal-Abarca & Gilabert, 1995	6	FR	NT-S vs. BAU-RI	T	SCI	E	0.45	C	0.21
Olsen, 1991	6 & 8	FR	EW vs. BAU-RI	NT	LA	N	0.73	C	*0.53
Ryan, 1981	6–8	FR	MULT {SUM vs. R} {NT-U vs. R}	NT	LA	N	0.55	C	0.50 {0.42} {0.52}
Barton, 1930	7	FR	NT-S vs. R	T	SS	E	0.27	C	0.63
Denner, 1987	7	FR	NT-U vs. R+RR	T	LA	N	0.73	C	0.45
Denner, 1987	7	FR	NT-S vs. R+RR	T	LA	N	0.73	C	0.70
Taylor & Beach, 1984	7	FR	MULT {QA vs. BAU-RI} {SUM vs. BAU-RI}	T	SS	E	0.55	C	0.43 {0.26} {0.75}
Bigelow, 1992	7 & 8	FR	NT-S vs. R	NT	SCI	E	0.67	C	0.78
Denner & McGinley, 1992	7 & 8	FR	SW & ShS vs. NW	NT	LA	N	0.78	C	0.60
Salisbury, 1934	7, 9, & 12	FR	SUM vs. BAU-RI	T	LA	E	0.36	C	*0.57
Armbruster & Anderson, 1980	8	FR	NT-S vs. BAU	T	SCI	E	0.36	C	0.34
Ballard, 1988	8	FR	NT-S vs. DRTA	T	LA	E	0.45	C	0.28
Chang, 1987	8	FR	NT-S vs. R	T	SCI	E	0.63	C	0.69
Denner, 1992	8	GR & PR	MULT {NT-S vs. R+RR} {NT-U vs. R+RR}	T	SS	E	0.89	C	0.54 {0.64} {0.45}
Vidal-Abarca & Gilabert, 1995	8	FR	NT-S vs. BAU-RI	T	SCI	E	0.64	C	0.22
Bates, 1981	9	GR & PR	SUM vs. R+RR	NT	LA	N	0.89	C	–0.17
Faber et al., 2000	9	GR & PR	NT-S vs. R	T	SS	E	0.78	C	0.03
Klugh, 2008	9	FR (SWD)	NT-S vs. BAU	T	SS	E	0.36	C	0.10
Slater, 1982	9	FR	NT-S vs. R+S	NT	SS	E	0.67	C	0.76
Slater, 1982	9	FR	NT-U vs. R	NT	SS	E	0.67	C	0.87
Trasborg, 2005	9 & 10	PR	SUM vs. BAU-RI	T	SS	E	0.73	C	*0.38

Langer & Applebee, 1987	9 & 11	FR	MULT {QA vs. R+S} {SUM vs. R+ST} {EW vs. R+ST}	NT	SS	E	0.63	C	0.37 {0.06} {0.51} {0.62}
Peverly & Wood, 2001	9–11	RD	QA vs. R	NT	LA	N	0.68	C	*0.44
Barton, 1930	9–12	FR	NT-S vs. R+RI	T	SS	E	0.36	C	0.37
Matthews, 1938	9–12	FR	NT-U vs. R	NT	SS	E	0.89	C	–0.15
Placke, 1987	9–12	LD	SUM vs. TT	T	SS	E	0.73	C	0.53; * –0.59
Tsai, 1995	9–12	FR	SUM vs. R+ST	T	SCI	E	0.50	C	0.28
Bean et al., 1983	10	AA	QG vs. R+D	T	SS	E	0.45	C	0.26
Graner, 2007	10	NLD & LD	SUM vs. TT	T	LA	E	0.77	C	0.20
Hayes, 1987	10	A & AA	MULT {SUM vs. R+M} {QG vs. M} {ShS vs. R+M}	NT	SCI	E	0.68	C	0.04 {–0.01} {0.14} {0.11}
Weisberg & Balajthy, 1989	10–12	PR	SUM vs. R+D	T	SS	E		C	0.43
Hare & Borchardt, 1984	11	AA	SUM vs. NW	T	SCI	E	0.45	C	0.44
Wetzel, 1990	11	FR	EW vs. R	NT	LA	N	0.55	C	*0.23
Andre & Anderson, 1979	11-12	A & AA	QG vs. R	T & NT	PSY	E	0.44	C	0.51
Bretzing & Kulhavey, 1979	11-12	FR	SUM vs. R	NT	SS	E	0.67	C	0.56
Kulhavey et al., 1975	11-12	FR	NT-U vs. R+ST	NT	LA	N	0.67	C	0.37
Schultz & Di Vesta, 1972	11-12	AA	NT-U vs. R	NT	SS	E	0.67	C	0.15
Bowman, 1989	12	FR	EW vs. R	NT	LA	N	0.55	C	0.47
Walko, 1989	12	FR	NT-U & NT-S vs. R+ST	NT	SCI	E	0.89	C	-0.11
Wong et al., 2002	12	FR	EW vs. R+D	NT	LA	N	0.55	C	0.85
Barton, 1930	HS	FR	NT-S vs. R	T	SS	E	0.36	C	0.83
Licata, 1993	HS	FR	MULT {EW- vs. R+ST} {EW- vs. R+ST}	NT	SCI	E	0.64	C	0.45 {0.56} {0.33}
Weisberg & Balajthy, 1990	NR	PR	SUM vs. R	T	SS	E	0.33	C	0.81

(continued)

TABLE 1 *Studies and effect sizes included in the analysis of each research question (continued)*

Study	*Grade*	*Students*	*Treatment*	*Training*	*Content*	*Genre*	*Quality*	*Measure*	*ES*
Question 2: Effects of writing instruction on reading comprehension									
Frey, 1993	1	FR	PrW vs. RI	T	LA	NA	0.73	C	*0.12
Fuchs et al., 2006	1	PR & PM	Sp vs. MI	T	LA	NA	0.83	W F	0.39 0.32
Uhry & Shepherd, 1993	1	FR	Sp vs. RI	T	LA	NA	0.67	C W F	*0.43 *1.78 *0.70
Conrad, 2008	2	A	Sp vs. R	T	LA	NA	0.67	W	0.62
Graham et al., 2002	2	WSP	Sp vs. MI	T	LA	NA	0.89	W	*0.51
Weber & Henderson, 1989	3–5	FR	Sp vs. RI	T	LA	NA	0.67	W F	0.54 1.17
Weber & Henderson, 1989	3–5	FR	Sp vs. RI	T	LA	NA	0.67	W F	0.34 0.79
Hunt & O'Donnell, 1970	4	FR	SC vs. RI	T	LA	NA	0.27	C	*0.26
Licari, 1990	4	FR	PrW vs. RI	T	LA	NA	0.55	C	*0.41
Crowhurst, 1991	6	FR	PE vs. R+D	T	SS	NA	0.88	C	0.36
Hamby, 2004	6	PR	WI vs. RI	T	LA	NA	0.36	C	*0.25
Kelley, 1984	6	FR (no SWD)	PrW vs. R	T	LA	NA	0.82	C	*0.40
Kelley, 1984	6	FR (no SWD)	WS vs. R	T	LA	NA	0.82	C	*0.31
Neville & Searls, 1985	6	FR	SC vs. R+Cloze	T	SS	NA NA	0.36 0.36	C C	*0.06 0.32
Hughes, 1975	7	FR	SC vs. OL	T	LA	NA	0.64	F	*0.57
Shockley, 1975	7	PR	WI vs. R+RI	T	LA	NA	0.78	C	*0.30

Jones, 1966	8	A & AA	Sp vs. NoTr	T	SCI	NA	0.73	C	*0.22
						NA		C	0.22
Phelps, 1978	8	A	SC vs. SC-NW	T	LA	NA	0.55	C	0.32
Callaghan, 1977	9 & 11	FR	SC vs. BAU	T	LA	NA	0.36	C	*0.02
Baker, 1984	10	FR	EssW vs. R	T	LA	NA	0.39	C	0.23
Gonsoulin, 1993	10–12	PR	SC vs. RI	T	LA	NA	0.55	C	*0.17
Question 3: Effects of extra writing on norm-referenced and researcher-created reading comprehension outcome measures									
Bode, 1988	1	FR	DJ vs. RI	NA	LA	NS	0.55	C	*0.36
Healy, 1991	1	FR	W-FY vs. W-HY+RI	NA	LA	NS	0.73	C	*0.56
Ramey, 1989	1	FR	W vs. R	NA	LA	SELF	0.45	C	*0.16
Sussman, 1998	1	WSP	XJW vs. RRJ	NA	LA	NS	0.67	C	0.22
Peters, 1991	2	FR	W vs. R	NA	LA	SELF	0.64	C	*0.31
Reutzel, 1985	3	FR	W vs. RI	T	LA	N	0.7	C	0.87
Soundy, 1978	3–6	FR	ExpW vs. SSR	NA	LA	SELF	0.78	C	*0.42
Roy, 1991	4 & 5	FR	RJ vs. R	NA	LA	NS	0.36	C	*0.01
Dana et al., 1991	6	FR	W vs. RI	NA	LA	PPL	0.09	C	0.24

Notes: A = average, AA = above average, BA = below average, BAU = business as usual, C = comprehension, Cloze = cloze activities, D = discussion, DJ = dialogue journals, DRTA = directed reading thinking activity, E = expository, ELL = English language learners, EW = extended writing, EssW = essay writing, ExpW = expressive writing, F = reading fluency, FR = full range, GR = good readers, GW = good writers, HS = high school, J = journals, LA = language arts, LD = learning disabled, M = math, MS = middle school, N = narrative, NA = not applicable, NLD = non–learning disabled, NoTr = no treatment, NR = not reported, NS = not specified, NT = no training, NT-S = note taking–structured, NT-U = note taking–unstructured, NW = no writing, PPL = pen pal letters, PR = poor readers, PrW = process writing, PSY = psychology, PW = poor writers, QA = question answering, QG = question generation, QGA = question generation and answering, R = reading, RD = reading disabled, RI = reading instruction, RJ = reading journals, RRJ = rereading journals, SC = sentence combining, SCI = science, SELF = self-selected topics, ShS = short statements, SI = spelling instruction, SS = social studies, SSR = sustained silent reading, ST = studying, SW = story writing, SUM = summary writing, SWD = students with disabilities, T = training, TT = test taking, U = underlining, W = word reading, W-FY = writing–full year, W-HY = writing–half year, WS = writing skills, WSP = weak spellers, XJW = extra journal writing

*Norm Referenced Measures

{ } = Specific treatment

TABLE 2 *Proportion of studies meeting requirements for each quality category by research question*

Quality Feature	*Q1: Writing about reading*	*Q2: Writing Instruction*	*Q3: Extra Writing*
Random assignment of participants	.52	.43	.33
Total attrition < 10%	.75	.76	.56
Equal attrition across conditions	.72	.62	.44
Control condition defined	.80	.81	1.00
Fidelity reported	.30	.05	.00
Teacher effects controlled	.66	.76	.44
More than 1 teacher per condition	.43	.62	.67
Reliability of the measure > .60	.62	.62	.67
No posttest ceiling/floor effects	.78	.81	.67
*No pretest ceiling/floor effects	.84	.58	.50
*Pretest equivalence	.74	.83	.67

Notes: Q1 = Question 1, Q2 = Question 2, Q3 = Question 3

*Proportion of quasi-experimental studies only.

Question 1: Does Writing About Material Read Enhance Comprehension?
— Average Weighted Effect Sizes

We found that writing about material read enhances reading comprehension, as 94 percent of studies produced a positive ES. The average weighted ES for the eleven experiments where reading comprehension was measured with a norm-referenced test was 0.37. This effect was statistically significant, and all of the variance was accounted for by sampling error (see table 3). Similarly, the average weighted ES for the fifty-five experiments applying researcher-designed reading comprehension measures was statistically greater than no effect (ES = 0.50); however, the Q test for heterogeneity was statistically significant, and I^2 indicated that 60 percent of the variance was due to between-study factors.

These findings apply to students in grades 2–12, with the majority of studies conducted with students in middle school (34%) and high school (41%). Slightly more than half of the reading material in the studies involved science and social studies text (55%), with 68 percent of the studies focusing on expository text. Students were taught how to apply the writing procedures in fewer than half of the studies (45%). While the control condition almost always involved reading or reading instruction (close to 90 percent of the time), methods for writing about text varied considerably.

The overall effects of writing about reading on reading comprehension were statistically significant and generally robust, as it was also evident when we

TABLE 3 *Average weighted effect sizes and confidence intervals for each meta-analysis conducted*

				Test of null hypothesis		*Heterogeneity*	
Question {Subset}	*Number of studies*	*Effect size*	*Confidence interval*	*Z-Score*	*p-value*	*Q-Value*	I^2
Question 1							
Comprehension—NRO	11	0.37	(0.23, 0.51)	5.15	< .001	4.79	0.00
Comprehension—RCO	55	0.50	(0.37, 0.62)	10.79	< .001	*134.66	59.90
Question 1 Subsets							
{Extended writing—RCO}	9	0.68	(0.38, 0.98)	4.41	< .001	*19.89	59.78
{Summary—RCO}	19	0.54	(0.31, 0.77)	4.59	< .001	*49.36	63.53
{Note taking—RCO}	25	0.45	(0.26, 0.63)	4.74	< .001	*59.37	59.57
{Questions—RCO}	8	0.28	(0.07, 0.48)	2.61	.008	1.66	0.00
{Poor readers/writers}	12	0.64	(0.27, 1.01)	3.43	.001	*25.92	57.56
Question 2							
Comprehension—NRO	12	0.22	(0.04, 0.41)	2.34	.019	2.01	0.00
Comprehension—RCO	5	0.27	(0.05, 0.48)	2.44	.015	0.23	0.00
Reading fluency—RCO & NRO	5	0.66	(0.27, 1.06)	3.29	.001	2.23	0.00
Word reading—RCO & NRO	6	0.62	(0.29, 0.95)	3.64	< .001	*6.29	20.51
Question 3							
Comprehension—RCO & NRO	9	0.35	(0.18, 0.51)	4.05	< .001	4.83	0.00

Notes: All of the effect sizes are significant at $p = .05$; NRO = norm-referenced outcomes; RCO = researcher-created outcomes.
*significant at $p = .05$.

TABLE 4 *Meta-regression for writing about reading (researcher-created comprehension outcomes)*

Descriptives						
	Mean ES	*R-square*	*k*			
	0.41	0.58	51			
Homogeneity analysis						
	Q	*df*	*p-value*			
Model	42.55	8	<.001			
Residual	30.63	42	.903			
Total	73.18	50	.018			
Regression coefficients						
			95% CI			
Variable	*B*	*SE*	*Lower*	*Upper*	*Z-Score*	*p-value*
Constant	0.68	0.08	0.51	0.84	8.07	<.001
ELEM vs SEC	0.01	0.12	–0.22	0.25	0.12	.901
HS vs MS	–0.33	0.08	–0.49	–0.17	–4.00	<.001
Training	–0.16	0.13	–0.41	0.09	–1.26	.207
Train x ELEM vs SEC	–0.28	0.18	–0.63	0.07	–1.58	.113
Train x HS vs MS	0.27	0.12	0.04	0.50	2.32	.020
LA vs SS/SC	–0.01	0.07	–0.14	0.12	–0.15	.877
SS vs SC	0.08	0.06	–0.03	0.19	1.43	.152
Study quality	–1.38	0.33	–2.03	–0.72	–4.13	<.001

Notes: ELEM = elementary grades, ES = average weighted effect size, HS = high school, k = number of comparisons, LA = language arts, MS = middle school, SCI = science, SEC = secondary grades (middle school and high school), SS = social studies, Train = training

examined specific types of writing activities. The findings for specific types of writing activities are reported for researcher-designed reading comprehension measures only. In table 3, we report that effect sizes for extended writing activities, summary writing, note taking, and asking or answering questions were all positive, ranging from 0.28 to 0.68. Much of the variance in ESs was due to between-study factors, not sampling error (see Q and I^2 statistics in table 3).

We also found that writing about reading had a positive impact on the comprehension of weaker readers/writers. In twelve studies with researcher-designed reading comprehension measures, a statistically significant weighted ES of 0.64 was obtained, with 83 percent of studies yielding a positive effect.

— Meta-Regression

We conducted three meta-regressions to examine if specific study features accounted for excessive variability in ESs. The first analysis involved all studies testing the effects of writing about reading. Predictor variables included grade (contrast coded to examine two orthogonal contrasts: elementary [grades 1–5] versus secondary [grades 6–12]; middle school [grades 6–8] versus high school [grades 9–12]); participant training and the interaction of training by grade (grade was contrast coded as above); subject area (contrast coded to examine two orthogonal contrasts: language arts versus social studies/science and social studies versus science); and quality of study (centered on the mean). We dropped four of the fifty-five studies in this analysis because they did not specify content area of the reading material ($k = 3$) or did not specify whether or not participants were trained in the use of the writing activity ($k = 1$).

The average weighted ES for the fifty-one studies in the meta-regression was 0.41. The analysis (see table 4) successfully explained excess variability in ESs, as it accounted for more than half of the variance (Q-value = 42.55, df[Q] = 8, $p < .001$), resulting in a nonsignificant Q-value for the residual of the model. The constant was statistically significant, indicating an average ES of 0.68 across grade levels and subject areas after accounting for variability due to study quality, training, and the interaction between grade and training. Three variables made unique and statistically significant contributions to the model. One, studies conducted in middle school settings produced an average effect that was 0.33 standard deviation units larger than studies conducted in high school. Two, training high school students to use the writing techniques produced an average effect size that was 0.27 standard deviation units larger when compared to middle school students who were trained. Three, studies of higher quality resulted in lower ESs, with a 0.10 change in study quality resulting in a decrease of 0.14 standard deviation units in the ES, as calculated from the constant (0.68), whereas studies of lower quality resulted in an increase of 0.14 standard deviation units.

The other two meta-regressions focused on summary writing and note taking, respectively. Because these analyses involved a smaller number of studies, we limited each to two predictors. For summary writing, this included grade (contrast coded as above) and participant training. The average weighted ES for the nineteen summary writing studies was 0.50. The inclusion of these two variables in the analysis (see table 5) successfully explained excess variability in ESs, as it accounted for 45 percent of the variance (Q-value = 12.27, df[Q] = 3, $p = .006$). The constant was statistically significant, indicating an average ES of 0.74 across grade when controlling for training. One variable made a unique and statistically significant contribution to the model. Studies conducted in middle school settings produced an average effect that was 0.32 standard deviation units larger than studies conducted in high school.

The two predictors for the meta-analysis for note taking were grade and type of notes (structured versus unstructured). We chose to focus on types of notes instead of training, as studies employing structured note taking almost always taught students how to take notes (88 percent of the time), whereas studies employing unstructured note taking did not provide instruction. As a result of

TABLE 5 *Meta-regression for studies examining summary writing*

Descriptives						
	Mean ES	*R-square*	*k*			
	0.50	0.45	19			
Homogeneity analysis						
	Q	*df*	*p-value*			
Model	12.27	3	.006			
Residual	14.86	15	.461			
Total	27.13	18	.077			
Regression coefficients						
			95% CI			
Variable	*B*	*SE*	*Lower*	*Upper*	*Z-Score*	*p-value*
Constant	0.74	0.17	0.41	1.07	4.37	< .001
ELEM vs SEC	−0.30	0.18	−0.65	−0.06	−1.63	.103
HS vs MS	−0.32	0.10	−0.53	−0.12	−3.12	.002
Training	−0.19	0.20	−0.59	−0.96	−0.96	.340

Notes: ELEM = elementary grades, ES = average weighted effect size, HS = high school, k = number of comparisons, MS = middle school, SEC = secondary grades (middle school and high school)

TABLE 6 *Meta-regression for studies examining note taking*

Descriptives						
	Mean ES	*R-square*	*k*			
	0.42	0.28	23			
Homogeneity analysis						
	Q	*df*	*p-value*			
Model	5.29	3	.152			
Residual	13.66	20	.847			
Total	18.95	23	.704			
Regression coefficients						
			95% CI			
Variable	*B*	*SE*	*Lower*	*Upper*	*Z-Score*	*p-value*
Constant	0.32	0.15	0.03	0.61	2.16	.031
ELEM vs SEC	−0.04	0.21	−0.46	0.37	−0.18	.855
HS vs MS	−0.08	0.09	−0.26	0.11	−0.83	.409
Type of notes	0.27	0.18	−0.08	0.61	1.49	.137

Notes: ELEM = elementary grades; ES = average weighted effect size; HS = high school; k = number of comparisons; MS = middle school; SEC = secondary grades (middle school and high school); Type of notes = unstructured versus structured notes this decision, we dropped Denner (1992) from the analysis to avoid statistical dependencies, because the study included both structured and unstructured note-taking groups compared with the same control group. The average ES for the twenty-three comparisons was 0.42 (see table 6). No variables in the model were statistically significant, and the model did not account for excess variability in ESs.

Question 2: Does Writing Skill Instruction Improve Reading?

We found that writing instruction enhances students' reading, as all twenty-one experiments produced a positive ES (see table 1). Studies involving process writing, text structure, and paragraph/sentence instruction resulted in statistically significant average weighted ESs of 0.22 (k = 12) and 0.27 (k = 5) for norm-referenced and researcher-designed reading comprehension tests, respectively. Likewise, instruction involving sentence construction or spelling instruction in five studies yielded a statistically significant average weighted ES of 0.66 for a combined analysis involving both norm-referenced and researcher-designed reading fluency tests. Finally, we obtained a statistically significant average weighted ES of 0.62 in a combined analysis involving both norm-referenced and researcher-designed word reading tests in six studies where spelling was taught. Much of the variance in these estimated ESs is accounted for by sampling error (see table 3).

The findings for the impact of writing instruction on reading comprehension applied to grades 4–12. The impact of such instruction on reading fluency and word reading applied to a narrower range of students (grades 1–7 and 1–5, respectively). All but two studies (9%) were conducted in the context of language arts, and the comparison condition in 70 percent of studies was reading or reading instruction.

Question 3: Does Increasing How Much Students Write Improve Reading?

We found that increasing writing improves reading comprehension, as all nine studies produced a positive effect. We obtained a statistically significant average weighted ES of 0.35 in a combined analysis involving both norm-referenced and researcher-designed reading comprehension tests, with all variance accounted for by sampling error (see I^2 statistics in table 3).

These findings apply only to students in grades 1–6 and were all conducted in the context of language arts classes. While the control conditions involved either reading or reading instruction, the treatment varied as students wrote about self-selected topics or topics chosen in collaboration with peers, set aside fifteen extra minutes each day for sustained writing, used the Internet to write to pen pals, wrote journal entries about daily experiences, interacted with others using a dialogue journal, and wrote short passages using inference words.

Discussion

Reading is critical to success in our school, work, social, and everyday lives. An important ingredient in ensuring that students become skilled readers involves teachers' use of effective practices for promoting and teaching reading. Previous reviews, including meta-analyses, have identified a number of practices that improve students' reading in grades 1–12 (e.g., Biancarosa & Snow, 2004; NICHD, 2000; Scammacca et al., 2007). This meta-analysis extends these efforts by examining if writing about material read, instruction designed to improve writing skills and knowledge, and increased time spent writing provides additional tools for enhancing students' reading.

Caveats and Limitations

Before we summarize the findings from this review, it is important to explore the factors that may affect the interpretation of such an analysis. First, meta-analyses such as this one involve aggregating the findings from individual studies to draw general conclusions about one or more questions. The value and breadth of any conclusions drawn depend on a variety of factors, such as the quality of the investigations and who participated in the studies. For instance, it is inappropriate to draw a broad conclusion aimed at all students if the research reviewed only involves high school youth. As a result, the conclusions we draw are appropriately restricted to the grades and types of students tested. Our conclusions are also constrained by study quality. We assessed the quality of each study and used this information to indicate how much confidence can be placed in each finding.

We limited this review to true and quasi-experiments involving controlled tests, where the reading gains made by one group of students who received a writing treatment were contrasted with a comparable group of students who did not engage in writing or writing instruction. While such studies control for a number of threats to internal validity (see Campbell & Stanley, 1963), our decision to focus just on this type of research should in no way distract from the important contribution that other types of research (e.g., qualitative and single-subject) make to our understanding of writing as a tool for supporting reading and reading development.

Another concern with meta-analysis involves the similarity of the control conditions in each study. If there is considerable variability in control conditions, the conclusions drawn are clouded, as there is no common point of comparison. We do not think this is a serious concern in the current review for two reasons. One, we excluded studies where students in the control condition wrote or received writing instruction, eliminating one source of variability. Two, while control conditions were not identical across all of the studies included in this review, they mostly involved some form of reading or reading instruction (86 percent of the time). This was the case for almost 90 percent of the studies that addressed Question 1, 70 percent for Question 2, and 100 percent for Question 3.

Yet another concern involves comparability of outcome measures on which the effect sizes are based. We addressed this problem in two ways. One, we separated effect sizes for different constructs (reading comprehension, reading fluency, and word reading) for stratified analyses. We also did this for norm-referenced and researcher-designed measures for specific constructs (e.g., reading comprehension) if there were at least four effects for each type of measure. Two, if a study contained more than one measure for a single construct (e.g., reading comprehension), we computed a single overall effect size for that construct by averaging the effects for each of the individual measures. Nonetheless, the use of different measures for the same reading construct by researchers introduces some unwanted noise into the machinery of this meta-analysis.

Some writing treatments examined in this review have been the focus of more experimental and quasi-experimental research than others. For example, the impact of taking notes about material read was examined in twenty-five experiments, while we located only one study that examined the effects of writing compare and contrast statements about material read (Hayes, 1987). There is clearly a need for additional replication for understudied treatments as well as the identification and testing of new writing-to-read treatments.

Although it is not uncommon for the magnitude of an ES to be interpreted using specific benchmarks (0.20 = small effect; 0.50 = moderate effect; 0.80 = large effect), it is preferable to interpret an effect within the context of other effects obtained in the given field of study (Lipsey & Wilson, 2001). Consequently, we compared the findings from the current meta-analysis to other meta-analyses examining the effectiveness of different reading interventions (Elleman, Lindo, Morphy, & Compton, 2009; Rosenshine & Meister, 1994; Slavin, Cheung, Groff, & Lake, 2008).

Finally, like other meta-analysts before us, we had to make a host of decisions—among them, what question each study answered, which subcategory a study should be assigned to, what variables to apply in the meta-regression. On the basis of reactions to other meta-analyses (e.g., Stotsky, 1988), we have no doubt that others will question one or more of the decisions we made. In anticipation of this, we tried to make our reasoning and decision making transparent.

Does Writing About Material Read Enhance Comprehension?

The evidence from this meta-analysis shows that having students in grades 2–12 write about material read enhances their comprehension of it. This was true for students in general and students who were weaker readers or writers in particular. It also applied across expository and narrative texts as well as subject areas (language arts, science, social studies). Moreover, the study found four types of writing activities to be effective: extended writing, summary writing, note taking, and answering/generating questions. Confidence can be placed in our findings, as we replicated them repeatedly and the quality of studies was relatively high. More than half of the investigations were true experiments, and the primary weaknesses were limited to failure to include multiple teachers in each condition (occurred 57 percent of the time) and to empirically verify that treatments were implemented as intended (occurred 70 percent of the time). It is difficult to know if the second issue is critical, as researchers only began reporting fidelity data about ten years ago. In any event, the finding that writing about reading improves comprehension provides support for the functional view of reading-writing connections (Fitzgerald & Shanahan, 2000).

As a point of comparison, the ES of 0.37 for norm-referenced tests obtained for writing about reading activities in this review rivaled or exceeded the effects obtained in other meta-analyses assessing reading interventions. On norm-referenced tests, Rosenshine and Meister (1994) reported an ES of 0.32 for reciprocal teaching; Slavin and colleagues (2008) obtained an ES of 0.17 for reading programs at the middle and high school level; and Elleman and colleagues (2009) found that vocabulary instruction produced an ES of 0.10.

It is important to note that the effects of writing in response to reading varied somewhat in magnitude between middle school and high school students. We obtained higher effect sizes when middle school students wrote about their reading than when high school students did. In addition, we obtained larger effects by high school students who were taught how to write about their reading versus those who were not. There are many possible reasons for these differences, including differences in writing activities and complexity of text. For instance, text read by high school students is presumably more complex than text read by middle school students, and this may have influenced the effectiveness of writing about reading for studies involving these two groups of students. In addition, the writing activities applied by middle school and high school students differed in both type and complexity, which may also have influenced how effective writing to read was for older and younger students. In terms of the relationship between training and the impact of writing to read on high school students' comprehension, it is possible that the complexity of text read by these students as well as the complexity of the writing activities they were directed to apply enhanced the need for instruction. In any event, it is important to remember that these specific findings are correlational, and additional research is needed to determine if they can be replicated when they are directly tested through experimentation and, if replicated, to determine why such differences were obtained.

Does Teaching Writing Improve Reading?

While writing and reading are not identical skills, teaching writing has a positive carryover effect to improving reading. This finding provides support for the shared knowledge view of reading-writing connections (Fitzgerald & Shanahan, 2000). Multicomponent writing instruction (e.g., process writing, skills-based programs) resulted in improved reading comprehension for typically developing students in grades 4–12, with three studies producing positive results for weaker writers (average ES = 0.24 on norm-referenced measures of reading comprehension). Teaching spelling and sentence construction skills improved the reading fluency of typically developing students in grades 1–7, whereas spelling instruction improved the word reading skills of typically developing as well as weaker spellers in grades 1–5.

Our finding that writing instruction had an ES of 0.22 on norm-referenced measures of reading comprehension compared favorably with effects obtained in two other reviews examining the impact of a range of reading programs on norm-referenced reading tests (Elleman et al., 2009; Slavin et al., 2008). Strong confidence can be placed in the findings that writing instruction improves reading fluency and word reading, as the writing treatments in studies assessing each of these outcomes were similar and the studies were of relatively high quality (i.e., all but one study met two-thirds or more of the quality indicators). However, the confidence that can be placed in the finding that writing instruction enhances reading comprehension must be tempered somewhat. The treatments applied in studies where reading comprehension was assessed were varied, and only half of the studies met at least two-thirds of the quality indicators.

Does Increasing How Much Students Write Improve Reading?

Increasing how much students write has a positive carryover effect on how well typically developing students in grades 1–6 read. This finding provides support for the rhetorical relations view of reading-writing connections that posits that students learn about reading through the act of composing their own text (Tierney & Shanahan, 1991). The obtained ES of 0.35 on norm-referenced measures of reading comprehension rivaled or exceeded the effects obtained in three other reviews examining the impact of a range of reading programs on norm-referenced reading tests (Elleman et al., 2009; Rosenshine & Meister, 1994; Slavin et al., 2008). The confidence that can be placed in the finding that increased writing enhances reading comprehension must be tempered by concerns about the quality of the research. Only one-third of studies involved randomization. Attrition, teacher effects, and pretest ceiling/floor effects (in quasi-experiments) were a problem in almost half of the investigations. Lastly, treatment fidelity was not established in any study.

Future Research

This review examines and summarizes cumulative results from previous experimental and quasi-experimental research examining the impact of writing and writing instruction on students' reading. While a considerable body of studies has accumulated over time (ninety-five investigations), there are many weaknesses and gaps in the research base. One weakness concerns the quality of the available research. There is clearly room for improvement. Random assignment occurred in less than half of the studies reviewed (48%); pretest ceiling/

floor effects were too common in quasi-experimental designs; treatment fidelity was rarely established; and only half of the studies had multiple teachers in treatment and control conditions.

It must also be noted that we do not know what combination, or how much of each writing practice reviewed here, teachers should apply in their classrooms. There is some preliminary evidence that integrating research-based writing practices can enhance students' performance, at least in terms of improving their writing skills (e.g., Sadoski, Wilson, & Norton, 1997). Additional research is needed, however, to determine the relative effectiveness of combining writing practices to enhance students' reading.

Another concern is the number of gaps in the research base and areas where more evidence is needed. For example, most of the research focused on typical students. Across all three questions, we located just eighteen studies where an ES could be computed for students experiencing difficulty learning to read or write (three additional studies were conducted with English language learners). Almost all of these studies examined the impact of writing about reading, so we know little about whether writing instruction or increased writing positively affects the reading of the most vulnerable students in school. Likewise, findings for some questions or subquestions applied to a limited set of grades. For instance, the impact of increased writing was tested only with students in grades 1–7.

Many instructional practices that have resulted in improved writing have not been tested to determine if they enhance students' reading. (See Graham and Perin [2007a] and Rogers and Graham [2008] for reviews of research-based practices in writing.) It is likely that the impact of writing instruction on reading can be strengthened if instruction is designed so that this is intentionally promoted. Research is needed to determine if other writing practices improve reading and how writing interventions can be designed so that they maximally enhance students' reading skills.

Finally, it is important to note that research interest in the impact of writing and writing instruction on reading is declining. Forty-one percent of the studies reviewed were conducted in the 1980s, 24 percent in the 1990s, and 15 percent after the millennium. Hopefully, this meta-analysis will spur new research and demonstrate that this is a productive and urgent area for investigation.

Concluding Comments

The positive impact of writing about material read, writing instruction, and increased time spent writing reported in this review is especially notable. While writing and writing instruction should not replace reading instruction, the writing treatments we assess here provide teachers with additional proven tools for strengthening students' reading. It is important to indicate that implementing research-based treatments, such as the ones studied here, is a challenging and complex task (Graham & Perin, 2007b).

Just because a writing intervention was effective in improving students' reading in the studies included in this review does not guarantee that it will be effective in all other situations. In fact, there is rarely an exact match between the conditions in which the research was implemented and the conditions in which it is subsequently implemented by teachers. Mismatches between the conditions where a practice is implemented by a teacher and its effectiveness as established by researchers can vary widely, including differences between students (e.g., reading or writing skills, dispositions, previous school success), instructional arrangements (e.g., number of students, material resources in the classroom), and the capabilities of those implementing instruction (e.g., beliefs about teaching and learning, success in managing the classroom, and experience teaching writing and reading). As a result, the safest course of action for teachers implementing research-based practices is to directly monitor the effects of such treatments to gauge whether they are effective under these new conditions (Graham & Harris, 2005).

The effects of writing and writing instruction on reading are likely to be minimized if students write infrequently or receive little instruction in how to write. For instance, Weber and Henderson (1989) reported that more writing instruction produced greater reading gains than less writing instruction. Despite the importance of writing to reading, learning, communicating, self-expression, self-exploration, and future employment (Bangert-Drowns et al., 2004; Graham, 2006; National Commission on Writing, 2004, 2005), students write infrequently and little time is devoted to writing instruction beyond the primary grades (Applebee & Langer, 2006; Graham & Gilbert, 2010; Kiuhara, Graham, & Hawken, 2009). This is the case even though just 33 percent of eighth graders and 24 percent of twelfth graders perform at or above the "proficient" level on the NAEP writing assessment (Salahu-Din, Persky, & Miller, 2008). As the National Commission on Writing (2003) asserted: "The nation's leaders must place writing squarely in the center of the school reform agenda" (p. 3). The findings of this meta-analysis provide additional support for such a revolution.

Finally, our findings provide empirical support for teachers who currently use writing as a tool to enhance students' comprehension of the material read. This review, along with the previous *Writing Next* (Graham & Perin, 2007a) report, also provides empirical support for teaching students to write. Teachers can enhance students' understanding of the text they read by having them write about it. However, as with any new practice that teachers are applying for the first time, we recommend starting small (Graham & Harris, 2005). For example, teachers might first test whether their students' comprehension of classroom text is improved by teaching them how to apply a writing strategy for summarizing text. If this can be done, then teachers might then engage students in applying different types of extended writing activities (e.g., writing to personalize information presented in text, writing to defend a specific point of view related to material read). As new writing to read activities are tested and applied, students and teachers should discuss how these activities help them understand and remember information from text as well as when and how to apply them to new situations. Of course, these efforts are more likely to promote positive development when implemented and supported at the school level.

References

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

Adams, M. (1990). *Beginning to read: Thinking and learning about print.* Cambridge, MA: MIT Press.

*Adams-Boating, A. (2001). *Second graders' use of journal writing and the effects on reading comprehension.* Unpublished master's thesis, Keen University, Sydney, Australia.

*Amuchie, P. M. (1983). *Teaching summarization skills to bilingual elementary school children.* Unpublished doctoral dissertation, University of California, Los Angeles.

*Andre, M., & Anderson, T. H. (1979). The development and evaluation of a self-questioning study technique. *Reading Research Quarterly, 14*(4), 605–623.

Applebee, A. (1984). Writing and reasoning. *Review of Educational Research, 54*(4), 577–596.

Applebee, A., & Langer, J. (2006). *The state of writing instruction: What existing data tell us.* Albany, NY: Center on English Learning and Achievement.

*Armbruster, B. B., & Anderson, T. H. (1980). *Technical report no. 160: The effect of mapping on the free recall of expository text.* Urbana-Champaign: University of Illinois, Center for the Study of Reading.

*Baker, C. A. (1984). *Effects of comparison/contrast writing instruction on the reading comprehension of tenth-grade students.* Unpublished doctoral dissertation, Indiana State University.

Baker, L., Dreher, J., & Guthrie, J. (2000). *Engaging young readers: Promoting achievement and motivation.* New York: Guilford Press.

*Ballard, P. J. Q. (1988). *The effects of narrative instruction on the comprehension abilities of eighth grade students.* Unpublished doctoral dissertation, Purdue University.

Bangert-Drowns, R. L., Hurley, M. M., & Wilkinson, B. (2004). The effects of school-based writing-to-learn interventions on academic achievement: A meta-analysis. *Review of Educational Research, 74*(1), 29–58.

*Barton, W. A. (1930). *Contributions to education no. 411: Outlining as a study procedure.* New York: Columbia University Bureau of Publications.

*Bates, G. W. (1981). *The comparative effects of two mathemagnic activities on ninth-grade good and poor readers' comprehension, retention, and attitudes.* Unpublished doctoral dissertation, University of Wisconsin–Madison.

*Bayne, M. (1984). *A study of the use of the semantic webbing technique to improve reading comprehension of third and fourth graders.* Paper presented at the Annual Meeting of the Northern Rocky Mountain Educational Research Association, Laramie, WY.

*Bean, T. W., Singer, H., Sorter, J., & Frazee, C. (1983). Acquisition of summarization rules as a basis for question generation in learning from expository text at the high school level. In J. A. Niles (Ed.), *Searches for meaning in reading/language processing and instruction: Thirty-second yearbook of the National Reading Conference* (pp. 43–48). Rochester, NY: National Reading Conference.

*Berkowitz, S. J. (1986). Effects of instruction in text organization on sixth-grade students' memory for expository reading. *Reading Research Quarterly, 21*(2), 161–178.

Berman, I. (2009, February 25). Supporting adolescent literacy achievement. *Issue Brief,* 1–15.

Biancarosa, G., & Snow, C. (2004). *Reading Next—A vision for action and research in middle and high school literacy: A report to the Carnegie Corporation of New York.* Washington, DC: Alliance for Excellence in Education.
*Bigelow, M. L. (1992). *The effects of information processing strategies and cognitive style on achievement of selected educational outcomes.* Unpublished doctoral dissertation, Pennsylvania State University.
*Bode, B. A. (1988). *The effect of using dialogue journal writing with first graders and their parents or teachers as an approach to beginning literacy instruction.* Unpublished doctoral dissertation, University of South Florida.
*Bowman, C. (1989). *The effect of learning logs on high school literature achievement.* Unpublished master's thesis, Virginia Polytechnic Institute and State University.
*Bretzing, B., & Kulhavey, R. (1979). Note taking and depth of processing. *Contemporary Educational Psychology, 4*(2), 145–153.
*Callaghan, T. F. (1977). *The effects of sentence-combining exercises on the syntactic maturity, quality of writing, reading, and attitudes of ninth grade students.* Unpublished doctoral dissertation, University of New York at Buffalo.
Campbell, D., & Stanley, J. (1963). *Experimental and quasi-experimental designs for research.* Chicago: Rand McNally.
Carnevale, A., & Derochers, D. (2004). *Standards for what? The economic roots of K–16 reform.* Princeton, NJ: Educational Testing Service.
*Chang, K. E., & Sung, Y. T. (2002). The effect of concept mapping to enhance text comprehension and summarization. *Journal of Experimental Education, 71*(1), 5–23.
*Chang, S. J. (1987). *An application of schema theory to school learning: Learning geography with the help of a note taking schema.* Unpublished doctoral dissertation, University of Texas at Austin.
*Coffman, G. A. (1992). *The effects of basal reader questions, causal chain questions, and unaided reading on sixth graders' understanding of complex stories.* Unpublished doctoral dissertation, University of Kansas.
*Cohen, R. (1983). Self-generated questions as an aid to reading comprehension. *The Reading Teacher, 36*(8), 770–775.
*Conrad, N. (2008). From reading to spelling and spelling to reading: Transfer goes both ways. *Journal of Educational Psychology, 100*(4), 869–878.
*Copeland, K. A. (1987). *The effects of writing upon good and poor writers' learning from prose.* Unpublished doctoral dissertation, University of Texas at Austin.
Cortina, J. M., & Nouri, H. (2000). *Effect size for ANOVA designs* (Vol. 129). Thousand Oaks, CA: Sage.
*Crowhurst, M. (1991). Interrelationships between reading and writing persuasive discourse. *Research in the Teaching of English, 25*(3), 314–338.
*Dana, M. E., Scheffler, A. J., Richmond, M. G., Smith, S., & Draper, H. S. (1991). Writing to read: Pen palling for a purpose. *Reading Improvement, 28*(2), 113–118.
*Denner, P. R. (1987). *Comparison of the effects of episodic organizers and traditional note taking on story recall.* (ERIC Document Reproduction Service No. ED270731)
*Denner, P. R. (1992). *Comparison of the effects of episodic mapping and traditional note taking on the recall of historical text.* Paper presented at the Northern Rocky Mountain Educational Research Association, Rapid City, SD.
*Denner, P. R., & McGinley, W. J. (1992). Effects of prereading activities on junior high students' recall. *Journal of Educational Research, 86*(1), 11–18.
*Denner, P. R., McGinley, W. J., & Brown, E. (1989). Effects of story impressions as a prereading/writing activity on story comprehension. *Journal of Educational Research, 82*(6), 320–326.
*Doctorow, M., Wittrock, M. C., & Marks, C. (1978). Generative processes in reading comprehension. *Journal of Educational Psychology, 70*(2), 109–118.

Ehri, L. (1987). Learning to read and spell words. *Journal of Reading Behavior, 19*(1), 5–31.

Ehri, L. (2000). Learning to read and learning to spell: Two sides of a coin. *Topics in Language Disorders, 20*(3), 19–49.

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

Emig, J. (1977). Writing as a mode of learning. *College Composition and Communication, 28*(2), 122–128.

*Faber, J. E., Morris, J. D., & Lieberman, M. G. (2000). The effect of note taking on ninth grade students' comprehension. *Reading Psychology, 21*(3), 257–270.

Fitzgerald, J., & Shanahan, T. (2000). Reading and writing relations and their development. *Educational Psychologist, 35*(1), 39–50.

*Frey, J. L. (1993). *The process writing approach and its effects on the reading comprehension of first-grade students in the Mississippi delta.* Unpublished doctoral dissertation, Delta State University.

*Fuchs, L. S., Fuchs, D., Hamlet, C. L., Powell, S. R., Capizzi, A. M., & Seethaler, P. (2006). The effects of computer-assisted instruction on number combination skill in at-risk first graders. *Journal of Learning Disabilities, 39*(5), 467–475.

Gleser, L. J., & Olkin, I. (1994). Stochastically dependent effect sizes. In H. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis* (pp. 339–355). New York: Russell Sage Foundation.

*Gonsoulin, K. W. (1993). *The effects of sentence combining curricula on descriptive writing performance, reading achievement, and writing apprehension: A study with high-risk secondary students.* Unpublished doctoral dissertation, University of Florida.

Graham, S. (2000). Should the natural learning approach replace traditional spelling instruction? *Journal of Educational Psychology, 92,* (2), 235–247.

Graham, S. (2006). Writing. In P. Alexander & P. Winne (Eds.), *Handbook of educational psychology* (pp. 457–478). Mahwah, NJ: Lawrence Erlbaum.

Graham, S., & Gilbert, J. (2010). Teaching writing to elementary students in grades 4 to 6: A national survey. *Elementary School Journal, 110*(4), 494–518.

Graham, S., & Harris, K. R. (2005). *Writing better: Teaching writing processes and self-regulation to students with learning problems.* Baltimore: Brookes.

*Graham, S., Harris, K. R., & Fink-Chorzempa, B. (2002). Contributions of spelling instruction to the spelling, writing, and reading of poor spellers. *Journal of Educational Psychology, 94*(4), 669–686.

Graham, S., & Hebert, M. (2010). *Writing to read: Evidence for how writing can improve reading.* Washington, DC: Alliance for Excellence in Education.

Graham, S., Morphy, P., Harris, K., Fink-Chorzempa, B., Saddler, B., Moran, S., & Mason, L. (2008). Teaching spelling in the primary grades: A national survey of instructional practices and adaptations. *American Educational Research Journal, 45*(3), 796–825.

Graham, S., & Perin, D. (2007a). *Writing next: Effective strategies to improve writing of adolescent middle and high school.* Washington, DC: Alliance for Excellence in Education.

Graham, S., & Perin, D. (2007b). A meta-analysis of writing instruction for adolescent students. *Journal of Educational Psychology, 99*(3), 445–476.

*Graner, P. G. (2007). *The effects of strategic summarization instruction in the performance of students with and without disabilities in secondary inclusive classes.* Unpublished doctoral dissertation, University of Kansas.

Greene, J. (2000). *The cost of remedial education: How much Michigan pays when students fail to learn basic skills.* Midland, MI: Mackinac Center for Public Policy.

*Hamby, J. (2004). *Explicit writing instruction: Effects on sixth grade students' writing and reading achievement.* Unpublished doctoral dissertation, University of San Diego.

*Hare, V. C., & Borchardt, J. (1984). Direct instruction of summarization skills. *Reading Research Quarterly, 20*(1), 62–78.

*Hayes, D. A. (1987). The potential for directing study in combined reading and writing activity. *Journal of Reading Behavior, 19*(4), 333–352.

*Healy, N. A. (1991). *First-graders writing with invented or traditional spelling: Effects on the development of decoding ability and writing skill.* Unpublished doctoral dissertation, University of Minnesota.

Hedges, L. V. (1982). Estimation of effect size from a series of independent experiments. *Psychological Bulletin, 92*(2), 490–499.

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

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.

*Hughes, T. O. (1975). *Sentence combining: A means of increasing reading comprehension.* (ERIC Document Reproduction Service No. ED112421)

*Hunt, K. W., & O'Donnell, R. (1970). *An elementary school curriculum to develop better writing skills.* (ERIC Document Reproduction Service No. ED050108)

*Jaekyung, L., Collins, J., & Fox, J. (2008). *When writing serves reading: Randomized trials or writing intensive reading comprehension (WIRC) in low-performing urban elementary schools.* Manuscript submitted for publication.

*Jenkins, J. R., Heliotis, J. D., & Stein, M. L. (1987). Improving reading comprehension by using paragraph restatements. *Exceptional Children, 54*(1), 54–63.

*Jennings, J. H. (1990). A comparison of summary and journal writing as components of an interactive comprehension model. In J. Zuttell & S. McCormick (Eds.), *Learner factors/Teacher Factors: Issues in literacy research and instruction: Fortieth yearbook of the National Reading Conference* (pp. 67–82). Chicago: National Reading Conference.

*Jones, J. L. (1966). *Effects of spelling instruction in eighth-grade biological science upon scientific spelling, vocabulary, and reading comprehension; general spelling, vocabulary, and reading comprehension; science progress; and science achievement.* Unpublished doctoral dissertation, University of Maryland.

Kamil, M., Mosenthal, P., Pearson, D., & Barr, R. (2000). *Handbook of reading research* (Vol. 3). London: Lawrence Erlbaum.

*Kelley, K. (1984). *The effects of writing instruction on reading comprehension and story writing ability.* Unpublished doctoral dissertation, University of Pittsburgh.

*Keown, S. L. (2008). *Effects of the use of thematic organizers in conjunction with concept mapping on learning, misconceptions, and retention in middle school science class.* Unpublished doctoral dissertation, Tennessee State University.

Kirsch, I., Braun, H., Yamamoto, K., & Sum, A. (2007). *America's perfect storm: Three forces changing our nation's future.* Princeton, NJ: Educational Testing Service.

Kiuhara, S., Graham, S., & Hawken, L. (2009). Teaching writing to high school students: A national survey. *Journal of Educational Psychology, 101*(1), 136–160.

Klein, P. (1999). Reopening inquiry into cognitive processes in writing-to-learn. *Educational Psychology Review, 11*(3), 203–270.

*Klugh, J. R. (2008). *An analysis of the effectiveness of graphic organizers in a social studies classroom.* Unpublished doctoral dissertation, California University of Pennsylvania.

Konstantopoulous, S., & Hedges, L. V. (2009). Analyzing effect sizes: Fixed-effects models. In H. C. Cooper, L. V. Hedges, & J. C. Valentine (Eds.), *The handbook of research synthesis and meta-analysis* (pp. 337–356). New York: Russell Sage Foundation.

Krashen, S. (1989). We acquire vocabulary and spelling by reading: Additional evidence for the input hypothesis. *Modern Language Journal, 73*, 440–464.

*Kulhavey, R. W., Dyer, J. W., & Silver, L. (1975). The effects of notetaking and test expectancy on the learning of text material. *Journal of Educational Research, 68*(10), 363–365.

*Langer, J., & Applebee, A. (1987). *How writing shapes thinking: A study of teaching and learning.* Urbana, IL: National Council of Teachers of English.

*Leshin, C. B. (1989). *Spatial representation and reading comprehension in 5th and 6th grade students.* Unpublished doctoral dissertation, Arizona State University.

*Licari, R. (1990). *The effects of directed writing activities on reading comprehension performance.* Paper presented at the Annual Meeting of the Northeastern Educational Research Association, Ellenville, NY.

*Licata, K. P. (1993). *Writing about mathematical relations in science: Effects on achievement.* Buffalo: State University of New York at Buffalo.

*Linden, M., & Wittrock, M. C. (1981). The teaching of reading comprehension according to the model of generative learning. *Reading Research Quarterly, 17*(1), 44–57.

Lipsey, M., & Wilson, D. (2001). *Practical meta-analysis.* Thousand Oaks, CA: Sage.

*MacGregor, S. K. (1988). Use of self-questioning with a computer-mediated text system and measures of reading performance. *Journal of Reading Behavior, 20*(2), 131–148.

*Matthews, C. O. (1938). Comparison of methods of study for immediate and delayed recall. *Journal of Educational Psychology, 29*(2), 101–106.

Moats, L. (2005/2006). How spelling supports reading: And why it is more regular and predictable than you may think. *American Educator, 29*(4), 12–22, 42–43.

National Center for Educational Statistics [NCES]. (2009). *Reading 2009: National assessment of educational progress at grades 4 and 8.* Washington, DC: Institute of Educational Sciences.

National Center for Educational Statistics [NCES]. (2010). *Grade 12 reading and math 2009 national and pilot state results.* Washington, DC: Institute of Educational Sciences.

National Commission on Writing. (2003). *The neglected R: The need for a writing revolution.* Retrieved from www.collegeboard.com

National Commission on Writing. (2004). *Writing: A ticket to work or a ticket out; A survey of business leaders.* Retrieved from www.collegeboard.com

National Commission on Writing. (2005). *Writing: A powerful message from state government.* Retrieved from www.collegeboard.com

National Institutes of Children's Health and Development [NICHD]. (2000). *Report of the national reading panel: Teaching students to read; An evidenced-based assessment of the scientific research literature on reading and its implications for reading instruction; Reports of the subgroups.* Bethesda, MD: National Institute of Child Health and Human Development.

*Neville, D. D., & Searls, E. F. (1985). The effect of sentence-combining and kernel-identification training on the syntactic component of reading comprehension. *Research in the Teaching of English, 19*(1), 37–61.

Neville, D., & Searls, E. (1991). A meta-analytic review of the effect of sentence-combining on reading comprehension. *Reading Research and Instruction, 31*(1), 63–76.

*Newlun, C. (1930). *Teaching children to summarize in fifth grade history* (Teachers College Contributions to Education No. 404). New York: Teachers College, Columbia University.

*Olsen, M. O. (1991). *The effects of the reader response journal technique on reading comprehension, attitude toward reading, and writing ability of sixth and eighth-graders.* Unpublished doctoral dissertation, University of Connecticut.

*Peters, P. A. (1991). *A self-generated writing program and its effects on the writing and reading growth in second-grade children.* Unpublished doctoral dissertation, Boston University.

*Peverly, S. T., & Wood, R. (2001). The effects of adjunct questions and feedback on improving the reading comprehension skills of learning-disabled adolescents. *Contemporary Educational Psychology, 26*(1), 25–43.

*Phelps, S. F. (1978). *The effects of integrating sentence-combining activities and guided reading procedures on the reading and writing performance of eighth-grade students.* Unpublished doctoral dissertation, Syracuse University.

*Placke, E. (1987). *The effects of cognitive strategy instruction on learning disabled adolescents' reading comprehension and summary writing.* Unpublished doctoral dissertation, State University of New York.

*Ramey, E. K. (1989). *The effect of shared journal writing on the development of reading comprehension of first-grade students.* Unpublished doctoral dissertation, Auburn University.

*Reutzel, D. R. (1985). Story maps improve comprehension. *The Reading Teacher, 38*(4), 400–404.

*Rinehart, S. D., Stahl, S. A., & Erickson, L. G. (1986). Some effects of summarization training on reading and studying. *Reading Research Quarterly, 21*(4), 422–438.

Rogers, L., & Graham, S. (2008). A meta-analysis of single subject design writing intervention research. *Journal of Educational Psychology, 100*(4), 879–906.

Rosenshine, B., & Meister, C. (1994). Reciprocal teaching: A review of the research. *Review of Educational Research, 64*(4), 479–530.

*Roy, C. P. (1991). *The effects of a related reading-writing activity upon intermediate students' reading comprehension, vocabulary, and essay writing development.* Unpublished doctoral dissertation, Rutgers University.

*Ryan, M. T. (1981). *Effects of paraphrase note taking on prose learning.* Unpublished doctoral dissertation, University of Connecticut.

Sadoski, M., Wilson, V., & Norton, D. (1997). The relative contributions of research-based composition activities to writing improvement in lower and middle grades. *Research in the Teaching of English, 31*(1), 120–150.

Salahu-Din, D., Persky, H., and Miller, J. (2008). *The nation's report card: Writing 2007* (NCES 2008–468). Washington, DC: National Center for Education Statistics.

*Salisbury, R. (1934). A study of the transfer effects of training in logical organization. *Journal of Educational Research, 28*(4), 241–254.

*Saunders, W. M., & Goldenberg, C. (1999). Effects of instructional conversations and literature logs on limited and fluent English proficient students' story comprehension and thematic understanding. *Elementary School Journal, 99*(4), 277–301.

Scammacca, N., Roberts, G., Vaughn, S., Edmonds, M., Wexler, J., Reutebuch, C., & Torgesen, J. (2007). *Intervention for struggling readers: A meta-analysis with implications for practice.* Portsmouth, NH: RMC Research.

*Schultz, C. B., & Di Vesta, F. J. (1972). Effects of passage organization and note taking on the selection of clustering strategies and on recall of textual materials. *Journal of Educational Psychology, 63*(3), 244–252.

Shadish, W. R., Robinson, L., & Congxiao, L. (1999). *ES: A computer program for effect size calculation.* Memphis, TN: University of Memphis.

*Shockley, S. J. (1975). *An investigation into the effects of training in syntax on reading comprehension.* Unpublished doctoral dissertation, University of Georgia.

*Slater, W. (1982). *The effects of structural organizers and rhetorical predicates on the recall of expository text.* Unpublished master's thesis, University of Minnesota.

Slavin, R., Cheung, A., Groff. C., & Lake, C. (2008). Effective reading programs for middle and high schools: A best evidence synthesis. *Reading Research Quarterly, 43*(3), 290–322.

*Soundy, C. S. (1978). *Effects of writing experiences in the expressive mode on children's reading comprehension and writing ability.* Unpublished doctoral dissertation, Rutgers University.

Stotsky, S. (1982). The role of writing in developmental reading. *Journal of Reading, 25*(4), 320–340.

Stotsky, S. G. (1988). Commentary. *Research in the Teaching of English, 22*(1), 89–99.

*Sussman, G. L. (1998). *The effects of phonologically constructed spelling on first graders' literacy development.* Unpublished doctoral dissertation, Fordham University.

*Taylor, B. M., & Beach, R. W. (1984). The effects of text structure instruction on middle-grade students' comprehension and production of expository text. *Reading Research Quarterly, 19*(2), 134–146.

*Taylor, B., & Berkowitz, S. (1980). Facilitating children's comprehension of content material. In M. L. Kamil & A. J. Moe (Eds.), *Perspectives in reading research and instruction. Twenty-ninth yearbook of the national reading conference* (pp. 64–68). Clemson, SC: National Reading Conference.

Tierney, R., & Shanahan, T. (1991). Research on the reading-writing relationship: Interactions, transactions, and outcomes. In R. Barr, M. Kamil, P. Mosenthal, & D. Pearson (Eds.), *The handbook of reading research* (Vol. 2; pp. 246–280). New York: Longman.

*Trasborg, C. F. (2005). *Comparing the effectiveness of direct explanation and note-taking training on the reading comprehension of secondary students.* Unpublished doctoral dissertation, City University of New York.

*Tsai, B. R. (1995). *Effects of student-generated summaries, instructor-provided summaries, and frequency of summarization during computer-based instruction.* Unpublished doctoral dissertation, University of Minnesota.

*Uhry, J. K., & Shepherd, M. J. (1993). Segmentation/spelling instruction as part of a first-grade reading program: Effects on several measures of reading. *Reading Research Quarterly, 28*(3), 218–233.

*Vidal-Abarca, E., & Gilabert, R. (1995). Teaching strategies to create visual representations of key ideas in content area text materials: A long-term intervention inserted in school curriculum. *European Journal of Psychology of Education, 10*(4), 433–447.

*Walko, J. K. (1989). *The effect of varied note-taking strategies on students' ability to profit from visualized instruction as measured by tests assessing different educational objectives.* Unpublished doctoral dissertation, Pennsylvania State University.

*Weber, W. R., & Henderson, E. H. (1989). A computer-based program of word study: Effects on reading and spelling. *Reading Psychology, 10*(2), 157–171.

*Weisberg, R., & Balajthy, E. (1989). Transfer effects of instructing poor readers to recognize expository text structure. In S. McCormick & J. Zutell (Eds.), *Cognitive and social perspectives for literacy research and instruction: Thirty-eighth yearbook of the National Reading Conference* (pp. 279–286). Chicago: National Reading Conference.

*Weisberg, R., & Balajthy, E. (1990). Development of disabled readers' metacomprehension ability through summarization training using expository text: Results of three studies. *Journal of Reading, Writing, and Learning Disabilities International, 6*(2), 117–136.

*Wetzel, G. H. (1990). *The effects of writing-to-learn on literature comprehension on English literature.* Unpublished doctoral dissertation, Temple University.

Wolf, I. (1986). *Meta-analysis: Quantitative methods for research synthesis.* Beverly Hills, CA: Sage.

*Wong, B. Y. L., Kuperis, S., Jamieson, D., Keller, L., & Cull-Hewitt, R. (2002). Effects of guided journal writing on students' story understanding. *Journal of Educational Research, 95*(3), 179–193.

Notes on Contributors

SUSAN ENGEL is a senior lecturer in psychology and founding director of the Program in Teaching at Williams College. Her research interests include children's narratives, play, the development of curiosity, and, more generally, teaching and learning. Her work has appeared in journals such as *Cognitive Development* and the *American Education Research Journal,* and she has also published a number of columns and op-ed pieces on education in the *New York Times.* She is the author of four books: *The Stories Children Tell: Making Sense of the Narratives of Childhood* (1995), *Context is Everything: The Nature of Memory* (1999), *Real Kids: Making Sense in Everyday Life* (2005), and *Red Flags or Red Herrings: Predicting Who Your Child Will Become* (2011). Engel is a founder of an experimental school in New York State, where for fourteen years she served as the educational adviser. She is currently working with the Spencer Foundation in trying to develop new ways to measure what children learn in school.

STEVE GRAHAM is the Curry Ingram Professor of Literacy at Vanderbilt University. His research focuses on identifying the factors that contribute to writing development and writing difficulties, developing and validating effective instructional procedures for teaching writing, and the use of technology to enhance writing performance. He is the past editor of *Exceptional Children and Contemporary Educational Psychology.* He serves as an editor on the upcoming three-volume series, *American Psychological Association Educational Psychology Handbook.* In addition, he is a senior editor of the *What Works for Special Needs Learners* series published by Guilford Press. Finally, he is a member of the Adolescent and Adult Literacy Panel formed by the National Research Council, he is the chair of the practice guide on writing instruction for elementary grade students being developed by What's Works Clearing House, and he is the member of the research advisory panel and early history panel for the National Writing Project. Steve has received numerous awards, including Career Research Award from the International Council for Exceptional Children, Samuel A. Kirk Award from the Division of Learning Disabilities, Distinguished Researcher Award from the special education interest group of the American Educational Research Association, Don Johnston Lectureship Award for Career Contributions to Literacy, and Distinguished Alumni 100th Anniversary Valdosta State University. From 1991 to 2002, he was the third most productive scholar in journals in educational psychology.

STEPHEN D. HANCOCK is an associate professor of multicultural education in the College of Education at the University of North Carolina–Charlotte, where his primary research interests are sociocultural perspectives in the urban elementary school context and the development of healthy academic relationships as it relates to the perceptions and psychology of self and others. In addition, his research interest focuses on intercultural identity in domestic and foreign spaces. His most recent work includes "Dysfunctional Policies and Relationships: Grooming Inmates through Curricula and Pedagogical Malpractice," *Journal of Curriculum and Pedagogy* (forthcoming), and "Deconstructing Teacher Bias: Exploring Attribution Bias in Urban Elementary Schools," in L. Howell, C. Lewis, & N. Carter (Eds.), *Yes We Can! Improving Urban Schools Through Innovative Educational Reform* (forthcoming), which investigates the impact of academic relationships and perceptions of self and others. Hancock is past chair of the Committee on Scholars of Color in the American Educational Research Association and currently serves as president of the North Carolina chapter of the National Association for Multicultural Education.

MICHAEL HEBERT is a doctoral candidate in special education at Vanderbilt University's Peabody College of Education. He is in the Experimental Education Research Training Program (ExpERT) at Vanderbilt, supported by the U.S. Department of Education's Institute for Education Sciences (IES). Mr. Hebert's research interests include writing development, reading development, and how writing may influence reading development, especially for students with reading and writing difficulties. He is an author on two influential meta-analyses, "Writing to Read" and "Informing Writing," both of which were funded by Carnegie Corporation of New York. He has several years of classroom teaching experience at the elementary level, including a year teaching on a Navajo reservation in Arizona. Mr. Hebert also has several years of experience as a reading specialist in El Segundo, California, where he taught students with reading difficulties, and he is a National Writing Project Fellow through the California Writing Project at UCLA.

HANA KAWAI graduated in May 2011 from Brown University, where she double-majored in education policy and ethnic studies. She is currently co-coordinating and facilitating the All Girl Everything Ultimate Program, a girls' sports and leadership program in Seattle, as well as working full-time in the software field. She hopes to return to school either to pursue a master's in teaching or to further her study of education policy. She plans to work with and alongside young people for justice in schools and beyond.

WENDY LUTTRELL is a professor of Urban Education and Social-Personality Psychology at the Graduate Center, City University of New York. She studies the process of self and identity formation and transformation in school settings and how gender, race, class, and sexuality systems of inequality take root in young people's self-evaluations and actions. She is the author of two award-winning books on this topic, *Schoolsmart and Motherwise: Working-Class Women's Identity and Schooling* (1997) and *Pregnant Bodies, Fertile Minds: Gender, Race and the Schooling of Pregnant Teens* (2003), and is also the editor of *Qualitative Educational Research: Readings on Reflexive Methodology and Transformative Practice* (2010). Her current longitudinal projects, *Children Framing Childhoods* and *Looking Back*, examine the role that gender, race, and immigrant status play in how diverse young people growing up in working-class communities portray their social and emotional worlds. Throughout her career, Luttrell has directed community-based, university, and teacher inquiry projects dedicated to advancing affective, moral, and social justice in and around schools and that promote innovative research and teaching practices.