Journal of Educational Psychology
2007, Vol. 99, No. 3, 445-476

0022-0663/07/$12.00 DOI: 10.1037/0022-0663.99.3.445

A Meta-Analysis of Writing Instruction for Adolescent Students

Steve Graham
Vanderbilt University

Dolores Perin
Teachers College, Columbia University

There is considerable concern that the majority of adolescents do not develop the competence in writing they need to be successful in school, the workplace, or their personal lives. A common explanation for why youngsters do not write well is that schools do not do a good job of teaching this complex skill. In an effort to identify effective instructional practices for teaching writing to adolescents, the authors conducted a meta-analysis of the writing intervention literature (Grades 4–12), focusing their efforts on experimental and quasi-experimental studies. They located 123 documents that yielded 154 effect sizes for quality of writing. The authors calculated an average weighted effect size (presented in parentheses) for the following 11 interventions: strategy instruction (0.82), summarization (0.82), peer assistance (0.75), setting product goals (0.70), word processing (0.55), sentence combining (0.50), inquiry (0.32), prewriting activities (0.32), process writing approach (0.32), study of models (0.25), grammar instruction (–0.32).

Keywords: writing, composition, adolescents, meta-analysis, instruction

Writing is a flexible tool that can be used to meet a variety of goals (Diamond, 1999; Graham, 2006b). Writing allows people to communicate with others removed in both distance and time. It can promote a sense of heritage and purpose among larger groups of people. For instance, a sense of national unity was fostered when China adopted a standard system of writing 2,300 years ago (Swedlow, 1999). Writing also provides a powerful medium for persuading others, as illustrated by the impact of Thomas Paine's prerevolutionary pamphlet, *Common Sense*. People use writing to explore who they are, to combat loneliness, and to chronicle their experiences. Writing about one's feelings and experiences, for example, is beneficial psychologically and physiologically (see Smyth, 1998, for a meta-analysis of this research). Writing is also a useful tool for learning (see meta-analyses by Bangert-Drowns, Hurley, & Wilkinson, 2004; Graham & Perin, 2007). The permanence of writing makes ideas readily available for review and evaluation. Its explicitness encourages the establishment of connections between ideas, and its active nature may foster the exploration of unexamined assumptions (Applebee, 1984; Keys, 2000; Shanahan, 2004; Sperling & Freedman, 2002).

Adolescents who do not learn to write well are at a disadvantage. In school, weaker writers are less likely than their more skilled classmates to use writing to support and extend learning in content classrooms. Their grades are likely to suffer, especially in classes where writing is the primary means for assessing progress (Graham, 2006b). Their chances of attending college are reduced, because universities increasingly use writing to evaluate applicants' qualifications. At work, writing has become a gateway for employment and promotion, especially in salaried positions (see reports by the National Commission on Writing, 2004, 2005). Employees in business as well as government (local, state, and federal) are expected to produce written documentation, visual presentations, memoranda, technical reports, and electronic messages. In the community at large, as E-mail has progressively supplanted the telephone for the purpose of communication, adults who are not able to communicate in writing may be unable to participate fully in civic life.

Despite the importance of writing, too many youngsters do not learn to write well enough to meet the demands of school or the workplace. Findings from the most recent National Assessment of Educational Progress revealed that many youngsters do not develop the competence in writing needed at their respective grade levels (Persky, Daane, & Jin, 2003). Despite small improvements since the previous assessment (Greenwald, Persky, Ambell, & Mazzeo, 1999), two thirds or more of students' writing in 4th, 8th, and 12th grade, was below grade-level proficiency. In their 2003 report, the National Commission on Writing (2003) bluntly concluded that the writing of students in the United States "is not what it should be" (p. 7). Likewise, college instructors estimated that 50% of high school graduates are not prepared for college-level writing demands (Achieve, Inc. 2005), whereas American businesses spend $3.1 billion annually for writing remediation (National Commission on Writing, 2004).

Editor's Note. CarolAnne Kardash, University of Nevada, Las Vegas, was the action editor for this article.—KRH

Steve Graham, Department of Social Education, Vanderbilt University; Dolores Perin, Department of Educational Psychology, Teachers College, Columbia University.

 Andrés Henríquez, Program Officer at the Carnegie Corporation of New York, offered helpful suggestions in the conceptualization and reporting of this study. We thank Joel Levin, Chris Schatschneider, and Robert Bangert-Drowns for their advice and assistance. Regan Phillips and Beth Chase assisted in the collection and organization of material for the study. Paul Morphy served as second coder to establish reliability. Copyright the Carnegie Corporation of New York.

Correspondence concerning this article should be addressed to Steve Graham, Peabody 328, 230 Appleton Place, Vanderbilt University, Nashville, TN 37203. E-mail: steve.graham@vanderbilt.edu

One possible reason why students' writing is not what it should be is because schools are not doing an adequate job of teaching it. According to the National Commission on Writing (2003), writing is the most neglected of the three *R*s in the American classroom. The commission further noted that writing needs to be placed "squarely in the center of the school agenda" (National Commission on Writing, 2003, p. 3) and indicated that many models for effectively teaching writing exist. The instructional recommendations offered by the commission, however, are limited and relatively vague with regard to instructional practices, because they are primarily limited to increasing the amount of writing students do within and outside of school, assessing students' progress in writing, using technology to advance the learning and teaching of writing, and better preparing teachers to teach writing.

A valuable approach to more fully identifying instructional practices that have the power to transform students' writing skills is to conduct a systematic review of writing intervention research. One approach to systematically reviewing intervention studies is to conduct a meta-analysis of relevant investigations. Meta-analysis is used to summarize the magnitude and directions of the effects obtained in a set of empirical research studies examining the same basic phenomena (Lipsey & Wilson, 2001). In this article, we report a comprehensive meta-analysis, funded by the Carnegie Corporation of New York, of experimental and quasi-experimental studies that examined the effectiveness of learning-to-write interventions. The primary purpose of this review was to identify effective practices for teaching writing to adolescents (defined here as students in Grades 4–12). The research question that guided this analysis was the following: What instructional practices improve the quality of adolescent students' writing? Meta-analysis is well suited to answering this kind of question, because it provides "an estimate of the magnitude of a treatment's effect under conditions that typify studies in the literature" (Bangert-Drowns et al., 2004, p. 34). When enough studies are available, it also permits examining the relationship between specific study features and study outcomes.

The last comprehensive meta-analysis of the writing intervention literature was conducted over 20 years ago by Hillocks (1986). Identical with our review, Hillocks examined the effects of specific writing treatments on the quality of students' writing by calculating the averaged weighted effect size (weighted by number of participants) for experimental and quasi-experimental studies published in journals and other sources (e.g., dissertations, conference papers, etc.). Also, as in the present review, he included only studies in which the researchers presented evidence that the quality measure was reliably assessed. In contrast to this review, participants in the studies examined by Hillocks ranged from Grade 3 through college.

Hillocks (1986) calculated average weighted effect sizes (corrected for pretest differences) for two categories of interventions. One category involved the teachers' mode or form of instruction. The treatments in this category differed according to teacher role, activities applied, and specificity of objectives and learning tasks. The four modes of writing instruction and their average weighted effect size were as follows: (a) presentational (mainly lecture and teacher-led discussion, effect size = .02), (b) environmental (materials and problems that cooperatively engage students in a specific writing process to meet clearly defined writing goals; effect size = 0.44), (c) natural process (mainly writing and revising self-selected topics, with high levels of interactions with peers, including receiving generally positive feedback: effect size = 0.18), and (d) individualized (individual instruction, including tutoring and programmed materials; effect size = 0.17).

The other category of interventions centered on the focus of instruction (i.e., the type of content or activities the intervention was expected to improve). The six treatments in this category and their respective average weighted effect sizes were as follows: (a) grammar instruction (effect size = −0.29), (b) sentence-combining instruction (effect size = 0.35), (c) study and emulation of good models of writing (effect size = 0.22), (d) student evaluation of writing using scales (effect size = 0.36), (e) free writing (effect size = 0.16), and (f) inquiry activities (effect size = 0.56). It should be noted that the modes-of-instruction and focus-of-instruction categories were not mutually exclusive (e.g., many of the same studies were classified as both environmental and inquiry treatments), and the modes-of-instruction category was criticized in terms of how treatments were defined and their overall distinctiveness (e.g., Applebee, 1986; Fox & Suhor, 1986; Golden, 1986; Stotsky, 1988).

Since Hillocks's (1986) seminal review, four other meta-analyses of learning-to-write interventions have been published. Two of these reviews examined the effectiveness of word-processing instruction with students in Grades 1–12. Bangert-Drowns (1993) reported that word-processing studies conducted before 1992 had a small but significant average unweighted effect size of 0.27 on the quality of writing produced by school-age children. In a subsequent meta-analysis, the average weighted effect size for word-processing studies conducted after Bangert-Drowns's study, but before 2003, was 0.41 (Goldring, Russell, & Cook, 2003).

Two meta-analyses have also examined the effectiveness of teaching students strategies for planning, revising, and/or editing. Graham and Harris (2003) reported that teaching such strategies by means of the self-regulated strategy development (SRSD) model (Harris & Graham, 1996) to students in Grades 3–8 resulted in a large average unweighted effect size of 1.47 for writing quality, Similarly, a large average unweighted effect size of 1.14 for writing quality for students in Grades 2–10 was obtained by Graham (2006a) in his analysis of all writing strategy instruction research. In addition, the magnitude of effect sizes was related to type of instruction, with the SRSD model yielding a higher effect size (1.57) than all other forms of strategy instruction combined (0.89). It is important to note that in the reviews by Graham (2006a) and Graham and Harris (2003), the standard deviation for the control group was used to compute effect sizes, whereas the other meta-analyses reviewed earlier used the pooled standard deviation of the treatment and control groups.

The meta-analysis reported on in this article draws on and extends previous meta-analyses in three important ways. First, we included all pertinent studies relevant to writing instruction for adolescents (Grades 4–12) contained in the five prior meta-analyses (Bangert-Drowns, 1993; Goldring et al., 2003; Graham, 2006a; Graham & Harris, 2003; Hillocks, 1986) as well as effect sizes from new investigations that were identified through our search-and-retrieval process (almost two thirds of the effect sizes were new). All effect sizes were recalculated for this review. There was considerable variation in the methods applied in the five previous reviews, including differences in procedures for calculat-

ing effect sizes, selecting studies, coding study features, selecting an outcome variable or variables, and conducting statistical analyses. Standardization of meta-analytic procedures across new studies and studies contained in prior meta-analyses is advantageous, because it makes comparisons among treatments more viable.

Second, the treatments investigated in this review were informed by the five previous meta-analyses but were not bound by them. For example, we kept some instructional treatments included in earlier reviews (strategy instruction, word processing, grammar instruction, sentence combining, the study of models, and inquiry), reorganized categories (e.g., eliminated Hillocks's, 1986, modes and focus of instruction, but added categories for explicit instruction and scaffolding), reconceptualized treatments (Hillocks's, 1986, natural process mode was reconceptualized as the process writing approach), eliminated treatments (e.g., Hillocks's, 1986, environmental treatment was eliminated because it overlapped considerably with inquiry), and added new treatments (e.g., setting writing product goals).

Third, none of the previous meta-analyses of writing treatments examined the relationship between overall quality of studies and magnitude of effect sizes. We rectified that shortcoming here. We assessed each study along nine dimensions, ranging from procedures used to assign participants (random and matched as well as nonrandom and not matched) to treatment fidelity (evidence that the experimental treatment was delivered as intended). This allowed us to identify important strengths and weaknesses in the accumulated body of experimental writing intervention research and to determine if overall study quality predicted the quality of students' writing following instruction (for all studies and for treatments where 10 or more effect sizes were calculated). We also examined whether specific study features, such as publication source (journal vs. other type of publication) or writing genre at posttest (narrative vs. expository), differentiated between investigations with larger and smaller effect sizes (for treatments that contained 18 effect sizes or more).

In summary, the primary research question guiding this review was the following: What instructional practices improve the quality of adolescent students' writing? We also chronicled the growth of the experimental and quasi-experimental writing intervention literature since Hillocks's (1986) seminal review and assessed how much experimental evidence had accumulated beyond the five meta-analyses described earlier. Finally, we examined whether there was a relationship between overall study quality and magnitude of effect sizes. When enough effect sizes were available (18 or more), we further examined whether specific study features (e.g., writing genre) were related to the impact of the treatment.

The theoretical bases for the treatments included in this review were varied and overlapping, and in far too many instances they were unstated. Much of the intervention research reviewed in this article was influenced by cognitive (J. R. Hayes, 2000), social/contextual (Prior, 2006), or both views of writing. For example, the popular process-writing approach was shaped by the cognitive revolution in writing and sociocultural views of composing (Pritchard & Honeycutt, 2006). Other theories have also played an important role in shaping the treatments reviewed here (see Sperling & Freedman, 2002), including but not limited to self-regulation (Zimmerman & Risemberg, 1997), schema (R. Anderson & Pearson, 1984), and discourse (Chafe & Tannen, 1987) theories. These viewpoints were not equally represented in the research studies included in this analysis, and it was extremely uncommon for studies to compare a treatment based on one theoretical perspective with a treatment based on another perspective. Thus, we draw no claims about the validity of specific writing theories.

Method

Location and Selection of Studies

The strategies that we used to locate and select studies for inclusion in this meta-analysis were influenced by seven factors. First, studies were included that involved students in Grades 4 through 12. We were primarily interested in drawing conclusions and providing recommendations for adolescent writers. There is, however, no universally agreed on age or grade range for adolescence. We decided not to use the start of middle school as the beginning of adolescence, because some middle schools start at 6th grade, others start at 5th grade, and some start as early as 4th grade. Because we were interested in casting as broad a net as possible, we focused our selection on studies involving youngsters in 4th through 12th grades. Some studies ($N = 5$) that included students in Grades 4 through 12 were eliminated because they also included students in the primary grades, and it was not possible to calculate an effect size just for the older students.

Second, we included studies that were conducted with students attending regular public or private schools. Consequently, we did not include studies of writing interventions delivered in special schools for children with deafness, autism, severe emotional disturbance, and so forth. Although we believe that writing instruction is an important part of the curriculum for these students (and that they should be educated in their neighborhood schools whenever possible), the purpose of this review was to draw recommendations for the teaching of writing within regular school settings.

Third, studies that included a measure of writing quality were included in this review. Writing quality was the primary or one of several primary outcomes in all previous meta-analyses on procedures for teaching writing (Bangert-Drowns, 1993; Goldring et al., 2003; Graham, 2006a; Graham & Harris, 2003; Hillocks, 1986). Writing quality is based on readers' judgment of the overall merit of a composition, taking into account factors such as ideation, organization, vocabulary, sentence structure, and tone. These evaluations are then quantified on a numerical rating scale (Diederich, 1966). We decided to make writing quality the sole outcome measure in our analysis, because we were interested in identifying treatments that had a broad impact on writing performance. We realize that this singular focus on writing quality narrows the types of recommendations that can be drawn. For example, handwriting and spelling pose considerable challenges for some adolescents who experience difficulty with writing. Although there are a number of studies that examined whether handwriting and spelling skills can be improved by directly teaching them (see Graham, 1999, 2000), the impact of such instruction on writing quality has been examined only with primary grade children (e.g., Graham, Harris, & Fink, 2000, 2002).

Fourth, we included studies in which researchers established that the measure of writing quality was scored reliably. To be included, an investigation had to provide evidence that interrater reliability for the quality measure was .60 or higher and/or that

trainers were taught how to score compositions. Almost all contemporary measures of writing quality involve some degree of subjectivity, making it particularly important to establish reliability of scoring procedures.

Fifth, we examined investigations that used an experimental or quasi-experimental design. Each study in this meta-analysis compared at least two groups of students who received different instructional conditions. This was consistent with most previous meta-analyses in writing (Bangert-Drowns, 1993; Goldring et al., 2003; Graham, 2006a; Hillocks, 1986). As a result, we did not examine correlational, qualitative, or single-subject design studies. Additionally, we did not include studies in which students served as their own controls.

Sixth, studies were included if they provided the data needed to calculate an effect size, a weighted average effect size, and homogeneity of effect sizes. For instance, if a study did not provide information on the number of participants involved in the experiment (e.g., Kerchner & Kistinger, 1984) it was excluded, because it could not be used in the calculation of homogeneity or an average weighted effect size. Tests of homogeneity examine whether the effect sizes for an identified treatment are so disparate that a single average estimate is not appropriate, providing an impetus to explore relations between study features and study outcomes (Lipsey & Wilson, 2001). Calculating an average weighted effect size is based on the premise that studies with more participants provide better estimates of population parameter than studies with fewer participants, making larger studies more precise and reliable. With this approach, each effect size is weighted according to sample size (using the inverse of the sampling error variance), so that its contribution is proportionate to its reliability (Lipsey & Wilson, 2001).

Seventh, we searched as broadly as possible to identify relevant studies. This involved locating peer-reviewed and nonreviewed studies from a variety of sources, including investigations in the previous meta-analyses (i.e., Bangert-Drowns, 1993; Goldring et al., 2003: Graham, 2006a; Graham & Harris, 2003; Hillocks, 1986), journals, theses and dissertations, conference proceeding, and books (e.g., *Science of Writing,* Levy & Ransdell, 1996; *Cognitive Processes in Writing*, Gregg & Steinberg, 1980; and the international series of books, *Studies in Writing,* published by Kluwer Academic). We decided to search broadly, going beyond published peer-reviewed articles, to reduce the possibility of bias in favor of significant findings (i.e., nonsignificant findings are rarely published in peer-reviewed journals). One drawback to this approach is that it may increase the likelihood of including studies of poor quality into the analysis, thus making interpretation of findings more difficult. To address this issue, we examined whether the effect size for studies published in journals versus other outlets differed (Lipsey & Wilson, 2001). We also examined whether study quality was related to magnitude of effect size.

A number of databases, including ERIC, PsycINFO, ProQuest, Education Abstracts (i.e., Education Full Text), and Dissertation Abstracts, were searched during May 2005 for relevant studies. We ran multiple searches in these databases, pairing the following terms with *writing* or *composition*: *adolescents, adolescents and instruction, strategy instruction, process instruction, planning, revising, peer collaboration, peer revising, peer planning, peers, collaborative learning, summary writing, summary instruction, summary strategies, motivation, motivation and instruction, achievement-motivation and instruction, technology, word processing, word processor, dictation, speech synthesis, spell checkers, sentence combining, dictation, goal setting, genre, genre and instruction, process writing approach, free-writing, writer's workshop, process approach to writing, process writing, National Writing Project, self-monitoring, self-evaluation, writing assessment, evaluative scales, usage, mechanics, grammar, inquiry, models, spelling instruction,* and *handwriting instruction*.

Once an item was located by an electronic search, the abstract was carefully read if available. If only a title was available, but the item looked promising, we obtained the document, as we also did for promising abstracts. Another source for possible studies was the reference lists in the narrative reviews, articles, book chapters, and so forth that were collected. Of 582 documents collected (this included studies from previous meta-analyses and pertinent book chapters), we found 123 documents that were suitable for inclusion.

The most common reason for excluding a study was that it was not an experimental or quasi-experimental study. Single-subject design methodology was the research method in 15% of the excluded studies. There was an experimental condition but no control condition 11% of the time. Students served as their own controls in 9% of the publications. The study was a qualitative study 3% of the time. The study involved either survey or descriptive research 2% of the time. It is important to note that there are many more surveys, descriptive studies, and qualitative investigations in the area of writing than were obtained in this search. We did not obtain a document if it was clear from the abstract that the study exclusively involved one or more of these types of research.

The next most common reason for excluding a study was that it did not include a measure of writing quality. This was the case 22% of the time. Only 2% of the studies were excluded because the researchers failed to establish reliability of the included quality measure. Another 2% of the studies were excluded because they were conducted in schools specifically for students with special needs (e.g., a school for students who are deaf), whereas 7% of studies were eliminated because the statistics needed for calculating an unweighted effect size were not provided. Additionally, 2% of the studies provided the statistics needed to calculate an unweighted effect size, but they did not provide the number of participants, making it impossible to calculate a weighted effect size. These studies were eliminated because our findings are based on weighted effect sizes. Of the excluded studies, 11% involved students who were too young (third grade or below) or too old (e.g., college students). Finally, 5% of the excluded studies were reviews, 4% described teaching procedures, and 1% did not involve writing at all.

Categorizing Studies Into Treatment Conditions

First, each study was read and placed into a treatment category identified in advance. These treatments were identical to descriptors used in the electronic searches (with the exception of writing and composition, which were too broad to use as a category) and were identified by examining previous meta-analyses and reviews of writing (e.g., Hillocks, 1986; MacArthur, Graham, & Fitzgerald, 2006) as well as by consulting a study that identified specific instructional practices included in observational studies of teachers' writing practices (Agate, 2005). Studies that did not fit neatly

within a preidentified treatment were held apart until all studies were read and sorted. At this point, the studies in each preidentified treatment were examined to determine if the intervention in each investigation represented the same general treatment. If they did not, they were placed with the studies that were not classified during the initial reading. All of the studies in this pile of unclassified investigations were read again, resulting in the construction of new treatments or, in a few instances, the placement of a study into one of the preidentified treatments. The studies in any new treatment as well as in any preidentified treatment where an additional study was placed were then reread to determine if each intervention represented the same basic treatment. As this process took place, it was necessary to refine some of our initial treatments and eliminate others (because no experimental studies of that treatment were located). It is important to note, that some studies were placed into more than one treatment, because they included two or more treatments. For example, Saddler and Graham (2005) compared the effects of sentence-combining instruction with more traditional grammar instruction. Thus, it provided an effect size for the sentence-combining treatment as well as an effect size for grammar.

Because we calculated summary effect sizes only for treatments that contained four or more effect sizes, we describe only those treatments here (see Table 1). Furthermore, all treatments with four or more effect sizes were examined and reexamined to determine if they could be grouped together in meaningful ways. Later we clarify the relationships between treatments in this meta-analysis and previous ones. We also clarify the relationships among the various categories and grouping arrangement used in this review.

Process writing approach. In the meta-analysis conducted by Hillocks (1986), he defined four treatments that represented the category *teachers' mode or form of instruction.* These modes of instruction differed in terms of the role assumed by the teacher, the kinds and order of activities applied, and the specificity and clarity of objectives and learning tasks. One of the modes defined by

Table 1
Definitions for Instructional Treatments That Contained Four or More Effect Sizes

Treatment	Definition
Process writing approach	This approach to teaching writing involves extended opportunities for writing; writing for real audiences; engaging in cycles of planning, translating, and reviewing; personal responsibility and ownership of writing projects; high levels of student interactions and creation of a supportive writing environment; self-reflection and evaluation; personalized individual assistance and instruction; and in some instances more systematic instruction.
Explicit teaching of skills, processes, or knowledge	
Grammar	This instruction involves the explicit and systematic teaching of grammar (e.g., the study of parts of speech and sentences).
Sentence combining	This instruction involves teaching students to construct more complex and sophisticated sentences through exercises in which two or more basic sentences are combined into a single sentence.
Strategy instruction	This instruction involves explicitly and systematically teaching students strategies for planning, revising, and/or editing text (Graham, 2006a). Instruction is designed to teach students to use these strategies independently. Writing strategies range from processes, such as brainstorming (which can be applied across genres), to strategies designed for specific types of writing, such as stories or persuasive essays.
Summarization	This instruction involves explicitly and systematically teaching students how to summarize texts. This can include teaching strategies for summarizing text or instructional activities designed to improve students' text summarization skills.
Text structure	This instruction involves explicitly and systematically teaching students knowledge about the structure of specific types of text, such as stories or persuasive essays.
Scaffolding students' writing	
Prewriting	This involves students engaging in activities (such as using a semantic web or brainstorming ideas) designed to help them generate or organize ideas for their composition.
Inquiry	This involves engaging students in activities that help them develop ideas and content for a particular writing task by analyzing immediate and concrete data (e.g., comparing and contrasting cases or collecting and evaluating evidence).
Procedural facilitation	This involves providing external supports (such as prompts, guides, hints, or heuristics) designed to facilitate one or more writing processes, such as planning or revising.
Peer assistance when writing	This involves students working together to plan, draft, and/or revise their compositions.
Study of models	This involves students examining examples of one or more specific types of text and attempting to emulate the patterns or forms in these examples in their own writing.
Product goals	These involve assigning students specific goals for the written product they are to complete.
Feedback	This involves students receiving input from others about the adequacy of their written product.
Alternative modes of composing: Word processing	This involves students using word processing computer programs to compose their composition.
Other: Extra writing	This involves students spending extra time writing or doing a particular kind of writing.

Hillocks was the *natural process mode*, which was characterized by general objectives, free writing with self-selected topics, writing for and receiving generally positive feedback from peers, opportunities to revise written work, and high levels of student interactions.

For our analysis, we renamed and restructured Hillocks's (1986) natural process mode. This category generated considerable controversy. It was criticized by some because it overemphasized the role of free writing in process-oriented approaches (Fox & Suhor, 1986). Others argued that it placed too much emphasis on the role of the teacher as a facilitator, noting that structured interactions between teacher and students can and do occur in this approach (Stotsky, 1988). One reason why so much controversy surrounds defining this approach is that there is no universally agreed on definition (Pritchard & Honeycutt, 2006). In addition, more structure has been added to its original conceptualization over time. For example, the National Writing Project (Nagin, 2003), which provides service and training to more than 10,000 teachers each year, has added more explicit instructional procedures to its process model over the years, including sentence-combining activities for enhancing sentence construction skills and the application of inquiry strategies for developing writing content.

In this analysis, we used the more common term, *process writing approach*, and included under its umbrella studies that examined the effectiveness of the process writing approach as described by Graves (1979), the Writers' Workshop (Calkins, 1981), the National Writing Project (Pritchard, 1987), and whole language (when a process writing approach was used). See Table 1 for a description of this approach.

Hillocks (1986) defined three other modes of instruction in his review. These included the presentational mode (lecture and teacher-led discussion, relatively clear and specific objectives, assignments that generally involve following rules or imitating models, and feedback on writing provided mostly by the teacher), the environmental mode (clear and specific objectives, materials and problems selected to engage students with each other in specifiable writing processes, and activities involving high levels of peer interaction around specific tasks), and the individualized mode (students provided with instruction on an individual basis).

Initially, we placed studies into these three modes of instruction. However, we found only one study with students in Grades 4–12 that involved the presentational mode (Caplan & Keech, 1980); thus, we did not compute an average effect size for this mode of instruction. Although we located eight studies that were classified as environmental, all but two of these (Bui, 2002; Tienken & Achilles, 2003) could be placed in other treatments: inquiry (Hillocks, 1979, 1982), study of models (A. E. Thibodeau, 1964), prewriting (Vinson, 1980), and grammar instruction (A. L. Thibodeau, 1964). The environmental mode has been criticized because it is not distinct enough from the presentational or the process writing mode. For example, both the environmental and the process writing approach involve peer interaction and activities for developing and revising a paper (Applebee, 1986), whereas Stotsky (1988) indicated that the environmental mode could be viewed as an "activity-laden version of the presentational mode" (p. 96). As a result of this lack of distinctiveness and clarity, we did not retain Hillocks's (1986) environmental treatment in this meta-analysis.

We also did not retain Hillocks's (1986) individualization treatment. We located six studies that were placed in this treatment. Three of these studies could also be placed in other treatments: peer assistance when writing (Yarrow & Topping, 2001), study of models (A. E. Thibodeau, 1964), and grammar instruction (A. L. Thibodeau, 1964). Just as important, the instructional methods in these six studies varied greatly and included individual use of programmed materials (A. E. Thibodeau, 1964; A. L. Thibodeau, 1964), older students tutoring younger students (Farrell, 1977; Utay & Utay, 1997), highly structured same-age tutoring (Yarrow & Topping, 2001), and tutors using self-initiated activities (Eagleton, 1973). Thus, these instructional procedures were too diverse to form a cogent treatment.

Explicit teaching of skills, processes, or knowledge. A considerable number of studies focused on explicitly teaching skills, processes, or knowledge. All of these studies involved sustained, direct, and systematic instruction designed to facilitate student mastery. There were five treatments that yielded at least four or more effect sizes. These were grammar, sentence combining, strategy, summarization, and text structure instruction (see Table 1 for a description of each).

Grammar and sentence combining were treatments included in Hillocks's (1986) review. Strategy instruction was the focus of two previous meta-analyses (Graham, 2006a; Graham & Harris, 2003), and we used the same criteria for identifying strategy instruction studies that were applied in these two reviews. First, students had to be taught one or more strategies for planning, revising, or editing text. Second, because the primary goal of strategy instruction is thoughtful and independent use of the target strategies, studies included in this treatment also had to meet the following criteria: (a) Students had to be shown how to use the strategy (i.e., modeling), (b) there were at least 3 or more days of instruction, and (c) instruction progressed toward students' independent use of the strategy.

Strategy instruction shared some overlap with other treatments classified as explicit teaching. Although the primary focus was always on teaching planning, revising, and/or editing strategies, some studies (especially those involving the SRSD model; Harris & Graham, 1996) also directly taught students knowledge and skills needed to use these processes. It is also important to note that some authors would classify strategy instruction as a process writing approach (see Pritchard & Honeycutt, 2006), because both instructions are concerned with promoting students' use of processes for planning, translating, and reviewing. We did not include strategy instruction under the process writing approach for two reasons: (a) Explicit and sustained instruction in writing strategies is not common in many classrooms that use a process approach model (e.g., Anthony & Anderson, 1987), and (b) such instruction is rarely included in descriptions of the components of a process writing program (see, e.g., Nagin, 2003). Even though sentence combining is included as a component of the National Writing Project, we also did not include it under the process writing approach, because it is only one of many elements included in this model.

Scaffolding students' writing. Scaffolding students' writing involves providing some form of assistance that helps the student carry out one or more processes involved in writing. These procedures include structuring how a student carries out a particular writing process, having peers help each other as they compose,

providing students with feedback on their performance, focusing students' attention on specific aspects of the task, and providing a model of what the end product should look like. There were seven scaffolding treatments that yielded four or more effect sizes each. These were prewriting, inquiry, procedural facilitation, peer assistance when composing, study of models, product goals, and feedback (see Table 1 for a description of each).

These procedures differ from the treatments included under the explicit-teaching category in an important way. The primary goal for explicit instruction treatments, such as strategy instruction and sentence combining, is to teach students specific skills, knowledge, or processes that they can use independently once instruction has ended. Thus, the goal of instruction is to place the targeted declarative and/or procedural knowledge directly under the writer's control as soon as possible. In contrast, scaffolding procedures are primarily aimed at providing the writer with some form of assistance immediately, through interactions with others and/or specific forms of structure. It is hoped that students will acquire declarative or procedural knowledge as a result of this scaffolding, but this process is less direct. For example, providing students with models of good essay provides immediate help, as it illustrates in a concrete fashion what they should try to achieve in their own writing. It is further assumed that as students repeatedly analyze these models and attempt to emulate them over time, they develop a better understanding of the criteria underlying good writing and that they increasingly apply this newly acquired knowledge without having to rely on the models for assistance.

It is also important to realize that some of these scaffolding procedures are integral components of explicit-teaching procedures. For example, with strategy instruction, students often help each other apply (and adapt) a strategy as they are learning to use it. They are encouraged to attain specific goals for their papers that can be achieved by using the strategy, and they receive feedback on how the strategy has enhanced their writing (Graham & Harris, 2003). These scaffolding procedures are not just elements of explicit-teaching procedures, however, as many of them are integral to the process approach to writing as well (e.g., prewriting and inquiry activities, feedback, and peers working together to compose a composition).

Finally, it must be noted that inquiry could be classified as a prewriting activity. We decided to keep it as a separate treatment because it uses a set of distinctive features for developing and analyzing ideas, and it was a unique category in Hillocks's (1986) review.

Alternative modes of composing. Two modes of composing that serve as an alternative to writing by hand are dictation and word processing (see Table 1 for a description). We located only 1 study in which students who used dictation were compared with a group that composed with handwriting (De La Paz & Graham, 1997). We identified 18 studies, however, in which students who used word processing were compared with students who wrote by hand. Most of these studies were included in two previous reviews (Bangert-Drowns, 1993; Goldring et al., 2003). We did not include a study by Rosenbaum (1987), reviewed by Bangert-Drowns (1993), because it did not assess overall text quality.

Other. Only one other treatment included four or more effect sizes. This treatment examined the effects of increased or extra writing on the quality of students' compositions (see Table 1). Studies in this treatment compared one group of students that did more writing or more of a particular type of writing with another group of students that did not do this extra writing.

Coding of Study Features

Each study was coded for five variables: grade, type of student, number of participants, writing genre, and publication type (see Table 2 for a description of each variable). We also developed a written description of the control condition. This provided information on who received a treatment (grade and type of student), how broadly it was applied (number of participants), what genre it was designed to impact (writing genre), and what treatment served as the comparison (control condition). These variables were also selected because we assumed that they might account for variability in effect sizes beyond participant-level sampling error (assessed by a test of homogeneity). For example, variability in effect sizes may be related to systematic differences in the control condition, type or grade level of participants, writing genre assessed, and publication outlet (journals vs. other publication outlets).

An additional study feature was coded for investigations for two treatments. We identified studies on the process writing approach, which involved professional development of teachers. Of the 6 identified studies, 5 involved professional development through the National Writing Project (Nagin, 2003). Such training (or the lack of it) may account for variability in effect sizes.

For the strategy instruction treatment, studies were coded as using or not using the SRSD model (Harris & Graham, 1996, 1999). In a previous meta-analysis by Graham (2006a), SRSD yielded larger effect sizes than the other methods of strategy instruction combined. The SRSD model includes six stages of instruction: (a) Develop background knowledge (students are taught any background knowledge needed to use the strategy successfully), (b) describe it (the strategy as well as its purpose and benefits are described and discussed; a mnemonic for remembering the steps of the strategy may be introduced too), (c) model it (the teacher models how to use the strategy), (d) memorize it (the student memorizes the steps of the strategy and any accompanying mnemonic), (e) support it (the teacher supports or scaffolds student mastery of the strategy), and (f) independent use (students use the strategy with little or no support). SRSD instruction is also characterized by explicit teaching, individualized instruction, and criterion-based versus time-based learning. Students are treated as active collaborators in the learning process. Furthermore, they are taught a number of self-regulation skills (including goal setting, self-monitoring, self-instructions, and self-reinforcement) designed to help them manage writing strategies, the writing process, and their writing behavior. Although the teaching regimes in the non-SRSD studies varied in their inclusion of the six instructional stages described earlier, their degree of individualization, and their degree of interactive learning, they were not criterion based nor did they typically emphasize the teaching of self-regulatory skills.

For all studies in each of the treatments described in Table 1, we assessed nine quality indicators for experimental and quasi-experimental designs (see Gersten et al., 2005). These quality indicators are described in Table 2 and include assignment of participants, mortality equivalence, no ceiling or floor effects, pretest equivalence, instructor training described, type of control condition, Hawthorne effect controlled, treatment fidelity established, and teacher effects controlled. For each quality indicator, a

Table 2
Definitions for Study Characteristics and Quality Indicators

Study characteristic	Definition
Grade	Grade(s) participants attended were identified. For a few studies, it was possible to identify only high school.
Type of student	Participants were categorized as full range (representing the full range of writers found in typical classrooms), high (above average writers only), average (average writers; this category did not include the weakest and strongest writers in a classroom), basic (poor writers only), second language learners, and special needs learners (e.g., students with learning disabilities and speech and language difficulties).
Number of participants	The total number of students that participated in the treatment–control contrast was tabulated.
Writing genre	The genre of the writing sample for posttest was classified as either narrative (e.g., stories, creative writing, personal narratives, etc.) or expository (e.g., descriptive, informative, persuasive, etc.).
Publication	Each publication was categorized as a journal article, book chapter, dissertation (this included thesis), report, or conference presentation.
Study quality indicators	
Assignment of participants	Studies were classified as involving random assignment of participants to conditions, matching of participants without random assignment, and nonrandom assignment without matching.
Mortality equivalence	Mortality equivalence was met if most of the students starting the study completed it (90%) and if there was equivalent mortality across conditions.
No ceiling or floor effects at posttest	Ceiling or floor problems were not evident if the mean of the posttest quality measure for each condition was more than one standard deviation away from the lowest and highest score of the scale.
Pretest equivalence	Pretest equivalence was met if the study provided evidence that the writing quality of students in each of the conditions was equivalent prior to the start of instruction.
Instructor training	Instructor training occurred if there was a description of how teachers were prepared to administer the experimental treatment.
Type of control condition	Control conditions were classified as an alternative treatment that was clearly described or an unspecified or no-treatment control condition.
Hawthorne effect	Hawthorne effect was not evident if the researcher(s) put into place conditions to control for it (e.g., providing an alternative treatment that controlled for attention and time).
Treatment fidelity	Treatment fidelity was established if evidence was provided that the experimental treatment condition was administered as intended (fidelity over 80%).
Teacher effects controlled	Teacher effects were controlled if instructors were randomly assigned to conditions or if they taught each condition.

score of 1 was assigned if there was evidence in the publication that it was met. The only exceptions to this involved assignment of participants (a score of 1 was assigned if participants were randomly assigned to conditions, and a score of .5 was assigned if participants were not randomly assigned but were matched on one or more variables) and type of control condition (a score of 1 was assigned if participants in the control condition received an alternative treatment that was described). The scores for the nine quality indicators were summed to provide a total quality score for each study. This allowed us to assess the quality of research included in this meta-analysis as well as to determine if there was a relationship between the quality of studies and the magnitude of effect sizes. It is important to note that all studies included in this review met two additional criteria: (a) A control or comparison group was present, and (b) writing quality was scored reliably.

To establish reliability of the coding procedures used in this review, a doctoral student majoring in education was familiarized with each variable and was then asked to score 15% of the studies (randomly selected). Across studies and variables, interrater agreement between the graduate student and Steve Graham was $\kappa = .88$ ($SE = .07$).

Calculation of Effect Sizes

Effect sizes were calculated as Cohen's *d* or the standardized mean difference. This involved subtracting the mean performance of the control or the comparison group at posttest from the mean performance of the treatment group at posttest and dividing by the pooled standard deviation of the two groups.

We recalculated effect sizes for all pertinent studies included in previous meta-analyses (Bangert-Drowns, 1993; Goldring et al., 2003; Graham, 2006a; Graham & Harris, 2003; Hillocks, 1986). A variety of procedures was used in these previous reviews, ranging from dividing by the standard deviation for the control group to adjusting for preintervention treatments. Our approach allowed us to standardize the calculation of effect sizes, while including as many studies as possible (e.g., we did not exclude studies because they could not be adjusted for preintervention differences). The only exception to this rule (i.e., basing the effect size on posttest data only) involved nine studies for which we were able to calculate effect sizes only from gain scores (Benson, 1979; Caplan & Keech, 1980; Gomez, Parker, Lara-Alecio, & Gomez, 1996; B. L. Hayes, 1984; Hillocks, 1979; Howie, 1979; Kanellas, Carifio, & Dagostino, 1998; M. C. Olson & DiStefano, 1980; Reedy, 1964). We decided to include these effect sizes for two reasons. First, effect sizes from seven of these studies were included in Hillocks's (1986) meta-analysis and excluding them would have made it more difficult to compare the findings from the two reviews. Second, the two more recent studies examined the effectiveness of important but less frequently studied treatments (i.e., extra writing was studied by Gomez et al., 1996, and sentence combining was studied by Kanellas et al., 1998). The inclusion of the effect sizes from these two studies provided a broader base from which to draw a conclusion.

Effect sizes were calculated for writing quality only. As noted earlier, measures of writing quality take into account factors such as ideation, organization, vocabulary, sentence structure, and tone,

Table 3
Total Quality Score and Percentage of Studies in Which Quality Indicator Was Present by Treatment

Treatment	Total quality score			Assignment (random)	Mortality equivalence	Ceiling/ floor	Pretest equivalence	Instructor training	Control type	Hawthorne effect	Treatment fidelity	Teacher effects
	M	*SD*	*N*									
All studies	5.3	1.8	107	33	80	94	57	46	84	61	27	46
Process writing approach	4.5	2.0	21	14	62	91	48	67	57	43	33	33
Grammar	5.0	2.1	11	36	82	91	73	9	91	64	36	18
Sentence combining	5.3	2.3	5	60	80	100	60	20	100	40	20	40
Strategy instruction	6.0	2.1	20	35	95	100	65	70	75	60	55	45
Summarization	6.5	0.6	4	0	100	100	50	50	100	100	0	50
Text structure	6.2	1.6	5	80	100	100	60	20	100	100	0	60
Prewriting	5.4	0.5	5	80	100	100	60	20	100	100	0	60
Inquiry	4.8	0.5	5	20	100	100	0	40	80	80	0	60
Procedural facilitation	6.8	2.6	4	75	100	75	75	50	100	75	50	75
Peer assistance	6.4	2.0	7	43	72	100	57	43	100	100	29	86
Study of models	5.2	2.1	6	33	67	100	33	67	100	50	17	50
Product goals	7.6	1.5	5	100	100	80	80	40	100	100	80	80
Feedback	5.8	1.3	5	80	80	100	60	20	80	80	20	60
Word processing	5.0	1.2	18	17	89	89	72	39	100	39	11	33
Extra writing	5.0	1.8	6	50	67	83	50	33	83	67	33	33

Note. *Assignment* refers to the percentage of studies in which students were randomly assigned to treatment; *mortality equivalence* refers to the percentage of studies in which most of the students starting the study completed it and in which there was equivalent mortality across conditions; *ceiling/floor* refers to the percentage of studies in which no ceiling or floor problems were evident; *pretest equivalence* refers to the percentage of studies in which there was evidence that the writing quality of students in each of the conditions was equivalent prior to the start of instruction; *instructor training* refers to the percentage of studies in which there was a description of how teachers were prepared to administer the experimental treatment; *control type* refers to the percentage of studies in which control conditions were classified as an alternative treatment that was clearly described or in which an unspecified or no-treatment control condition was noted; *Hawthorne effect* refers to the percentage of studies in which the researcher put into place conditions to control for Hawthorne effects; *treatment fidelity* refers to the percentage of studies in which evidence was provided that the experimental treatment condition was administered as intended; *teacher effects* refers to the percentage of studies in which teacher effects were controlled by randomly assigning teachers to conditions or by having each teacher teach all conditions.

and quantify these evaluations on a numerical rating scale (Diederich, 1966). In some instances, this involves considering all of these factors simultaneously and assigning a composition a single score. This is referred to as the holistic approach. In other instances, a score is assigned to each factor separately. This is referred to as the analytic method. If a holistic score was available, then we calculated the effect size with this measure. If both holistic and analytic scores were available, only the holistic score was used. If only an analytic scale was available, we first calculated an effect size for each attribute separately and then averaged these separate effect sizes to obtain a general measure of writing quality (similar to a holistic score). We did not include writing mechanics in this calculation, because it is often not included as an attribute in holistic assessments. For the most part, the measures of writing quality, both holistic and analytic, focused on the attributes of writing quality (i.e., ideation, organization, vocabulary, etc.) described earlier. It is important to note that measures of quality for studies in the summary writing treatment were based on the completeness and accuracy of the written summary produced by students.

A correction for small sample size bias was made when the sample included fewer than 20 participants (Hedges & Olkin, 1985). This occurred in only three instances. When means or standard deviations were unreported, effect sizes were calculated from *t* tests, analyses of variance (ANOVAs), or regression analyses.

To avoid inflating sample size (Wolf, 1986) and violating the assumption of independence of data underlying the ANOVA analogue used to examine moderating effects of study characteristics in this review, we computed only one effect size for each study. There were two exceptions to this rule. One exception involved 11 instances where just two treatments were compared and both treatments fit one of the established treatments. For example, Curry (1997) compared strategy instruction and the process writing approach. An effect size was calculated for both treatments (effect size = 0.51 for strategy instruction and −.51 for the process writing approach). Because analyses were done for each treatment separately, this did not present a problem. The other exception involved cases where more than two different treatments were compared. When this occurred, an effect size for each treatment was calculated. For example, Knudson (1989) compared four different treatments, allowing us to calculate more than one effect size for her investigation.

The 123 documents selected yielded 154 effect sizes. One effect size was included in two different treatments (thus, there are a total of 155 effect sizes presented in Tables 3 and 4 and the Appendix). This effect size was from a study by MacArthur, Schwartz, and Graham (1991) and involved a treatment including both strategy instruction and peer collaboration.

Statistical Analysis of Effect Sizes

We conducted analyses for each treatment that included four or more effect sizes. There was precedence for this decision, as this

(*text continues on page 459*)

Table 4
Writing Instruction Treatments That Included Four or More Effect Sizes

Study	Grade	Participant type	*n*	Genre	Treatment	Quality indicator score	Publication type	Effect size
					Process writing approach			
Umbach (1990)	4	basic	60	N, E	process writing approach versus instruction in strategies for drafting paper (random assignment)	9.0	D	−0.03
Curry (1997)	4	SNL	43	N	process writing approach versus skills instruction	5.0	D	0.69
Troia and Graham (2002)	4–5	SNL	20	E	modified process writing instruction versus strategy instruction (random assignment)	8.0	J	−0.14
Gorka (1992)	4–6	full range	60	?	effects of staff training in process writing approach on students' writing versus unspecified control (random assignment)**	4.0	D	0.83
Pantier (1999)	5	full range	33	N	process writing approach versus grammar instruction	5.0	D	−0.30
Moye (1993)	5	full range	121	N	process writing approach (includes model and scales) versus teaching students to use graphic organizers	6.0	D	0.48
Robinson (1986)	5	full range	120	N	process writing approach versus traditional instruction	4.5	D	0.28
Varble (1990)	6	full range	128	N	whole language with process writing instruction versus skills instruction	3.0	J	−0.11
Gamelin (1996)	7	full range	52	E	process writing approach versus strategy instruction	5.0	D	−0.98
Hayes (1984)*	7	full range	70	N, E	process writing approach versus traditional grammar instruction	2.0	D	0.22
Yeh (1998)	7	full range	110	E	process writing approach versus strategy instruction	7.0	J	−0.14
M. C. Olson and DiStefano (1980)*	7–9	full range	390	N, E	National Writing Project training versus unspecified control**	3.0	J	0.40
Ganong (1974)	9	average & high	133	?	early form of process writing approach versus a more traditional approach in which students follow a prescribed series of writing exercises	7.0	D	−0.13
C. Roberts (2002)	6–12	full range	72	E	National Writing Project training versus unspecified control**	3.5	D	0.14
Pritchard (1987)	7–12	full range	383	?	National Writing Project training versus unspecified control**	2.0	J	0.38
Pritchard and Marshall (1994)	7–12	full range	2,635	E	National Writing Project training versus unspecified control**	3.0	J	0.50
Alloway et al. (1979)	7–12	full range	225	?	National Writing Project training versus unspecified control**	3.0	R	0.39
Gauntlett (1978)	10–12	full range	791	E	process writing approach versus skills instruction	2.0	D	0.02
Adams (1971)	12	high	56	E	early form of process writing versus skills instruction	5.0	D	0.28
Reimer (2001)	HS	full range	30	E	process writing approach versus traditional instruction	5.0	D	−1.00
Scannella (1982)	HS	full range	95	N, E	process writing approach versus traditional instruction	3.0	D	0.14
					Grammar			
Saddler and Graham (2005)	4	average & basic	44	N	grammar instruction versus sentence combining (random assignment)	8.0	J	−0.42
A. A. Anderson (1997)	5	full range & SNL	37	N	grammar instruction versus planning or revising strategy instruction (random assignment)	8.0	D	−1.40
Pantier (1999)	5	full range	33	N	grammar instruction versus process writing	5.0	D	0.30
A. L. Thibodeau (1964)	6	full range	402	N	traditional grammar instruction versus peers working on elaborative thinking and vocabulary enrichment activities	4.0	D	−0.54
A. L. Thibodeau (1964)	6	full range	363	N	traditional grammar instruction versus individual students working on self-directing, elaborative thinking vocabulary enrichment activities	4.0	D	−0.41
Howie (1979)*	9	full range	91	E	traditional grammar instruction versus sentence-combining instruction (random assignment)	6.0	D	−0.21
Hayes (1984)*	7	full range	70	N, E	grammar instruction versus process writing	2.0	D	−0.22
Kanellas et al. (1998)*	9	average	120	E	grammar instruction versus sentence combining (random assignment)	6.0	B	−0.61
Fearn and Farnan (2005)	10	full range	57	E	grammar instruction in context versus traditional grammar instruction	5.0	P	1.07
Elley et al. (1975), Comparison 1	11	average	104	?	transformational grammar versus reading and writing	3.0	J	0.00
Elley et al. (1975), Comparison 2	11	average	122	?	traditional grammar versus reading and writing	3.0	J	0.03

Table 4 (*continued*)

Study	Grade	Participant type	*n*	Genre	Treatment	Quality indicator score	Publication type	Effect size
					Sentence combining			
Saddler and Graham (2005)	4	average & basic	44	N	sentence-combining instruction versus grammar instruction (random assignment)	8.0	J	0.42
Stoddard (1982)	5–6	high	120	N	sentence-combining instruction versus unspecified control condition	3.0	D	0.66
Pedersen (1977)	7	full range	36	N, E	sentence-combining instruction added to traditional literacy instruction versus traditional literacy instruction	4.0	D	0.40
Howie (1979)*	9	full range	91	E	sentence-combining instruction versus traditional grammar instruction (random assignment)	4.0	D	0.21
Kanellas et al. (1998)*	9	average	120	E	sentence-combining instruction versus traditional grammar instruction (random assignment)	6.0	B	0.61
					Strategy instruction			
Curry (1997)	4	SNL	43	N	students taught planning strategies for story writing versus writing skills instruction***	5.0	D	0.51
Glaser (2005)	4	full range	119	N	students taught a planning strategy for story writing versus unspecified control***	8.0	D	1.24
Walser (2000)	4	full range	41	N	students taught planning and revising for story writing versus instruction in narrative writing plus journal writing (random assignment)	8.0	D	0.91
Troia and Graham (2002)	4–5	SNL	20	E	students taught a planning strategy for persuasive writing versus partial process writing model (random assignment)	8.0	J	0.14
MacArthur et al. (1991)	4–6	SNL	29	N	students taught a peer revising strategy for narrative writing in the context of a process writing approach versus process writing***	7.0	J	1.09
A. A. Anderson (1997)	5	full range & SNL	45	N	students taught a planning strategy for story writing versus grammar and written literature summarization instruction (random assignment)***	8.0	D	1.40
Sawyer et al. (1992)	5–6	SNL	22	N	students taught a planning strategy for story writing versus writing practice control condition***	6.0	J	1.86
De La Paz and Graham (1997)	5–7	SNL	32	E	students taught a planning strategy for persuasive text versus instruction on the structure of persuasive essays (random assignment)***	9.0	J	0.82
Fitzgerald and Markham (1987)	6	full range	30	N	students taught revising strategies versus reading good literature (random assignment)	8.0	J	0.32
Scardamalia et al. (1984)	6	full range	62	E	students taught strategies for being self-reflective when planning versus unspecified control	2.0	J	0.65
Welch (1992)	6	SNL	18	E	students taught a planning strategy for paragraph writing versus unspecified control	3.0	J	2.26
Welch and Jensen (1990)	6–8	basic	114	E	students taught a planning strategy for paragraph writing versus unspecified control	4.0	J	0.72
Reynolds et al. (1988)	6–8	SNL	53	E	students taught editing and sentence-level revising strategies versus students directed to plan, draft, and revise compositions	3.0	J	0.16
Gamelin (1996)	7	full range	52	E	students taught planning and revising strategies for compare–contrast essays versus process writing instruction	5.0	D	0.98
Yeh (1998)	7	full range	110	E	students taught a planning strategy for persuasive writing versus process writing approach	7.0	J	0.14
De La Paz and Graham (2002)	7–8	full range	58	E	students taught planning and revising strategies for expository writing versus traditional instruction***	8.0	J	0.95
De La Paz (2005)	8	full range	105	E	students taught a planning strategy for writing historical text versus traditional instruction***	5.0	J	1.36
Simmons et al. (1994)	8	full range	101	N	students taught planning and revising strategies for narrative text versus narrative text structure instruction in general writing strategy	5.0	B	0.40
Wong et al. (1996)	8–9	basic & SNL	38	E	students taught planning and revising strategies for expository text versus unspecified control	3.0	J	3.50
Bryson and Scardamalia (1996)	10	SNL & full range	31	E	students taught strategies for reflection when writing persuasive text versus instruction on basic elements of persuasive writing (random assignment)	7.0	J	1.27

(*table continues*)

Table 4 (*continued*)

Study	Grade	Participant type	*n*	Genre	Treatment	Quality indicator score	Publication type	Effect size
					Summarization			
Chang et al. (2002)	5	full range	66	E	students taught to summarize through progressively faded expert summaries versus unspecified control	5.0	J	0.81
Bean and Steenwyk (1984)	6	full range	41	E	students taught rule-based summary strategy versus advice to write summaries by finding main idea	5.0	J	1.09
Knight (2003)	8	high	27	N,E	students taught how to synthesize information from multiple sources versus unspecified control	5.0	D	0.18
Placke (1987)	10–12	SNL	30	E	students taught strategies for summarizing main ideas versus completing cloze tasks	6.0	D	1.12
					Text structure			
Fitzgerald and Teasley (1986)	4	basic	19	N	students taught the basic elements of a story versus reading good literature (random assignment)	8.0	J	0.70
Scardamalia and Paris (1985), Study 1	4 & 6	full range	62	E	students taught the basic elements of persuasive text versus no-treatment control (random assignment)	5.0	J	−0.45
Gordon and Braun (1986)	5	full range	54	N	students taught the basic elements of a story versus instruction in poetry and its structure (random assignment)	7.0	J	0.32
Scardamalia and Paris (1985), Study 2	6 & 10	full range	24	E	students familiarized with basic elements of persuasive text versus no-treatment control	4.0	J	0.75
Bryson and Scardamalia (1996)	10	SNL & full range	31	E	students taught basic elements of persuasive text versus strategy instruction on reflection (random assignment)	7.0	J	−1.27
					Prewriting			
Loader (1989)	4	full range	47	E	students completed a semantic web versus listing ideas for writing	6.0	D	0.32
Brodney et al. (1999)	5	full range	51	E	students read information on topic and prompted to plan versus listening to information on topic (random assignment)	4.0	J	0.95
Brodney et al. (1999)	5	full range	49	E	students prompted to plan a paper after listening to information on topic versus listening to information on topic (random assignment)	4.0	J	0.17
Reece and Cumming (1996), Study 4	5–6	full range	20	?	students encouraged to plan after brief demonstration on how to do so versus unspecified control (random assignment)	6.0	B	0.61
Vinson (1980)	9	full range	109	E	groups of students brainstormed, ideas for paper, small group discussion of which ideas to include, and students organized ideas prior to writing versus writing correct paragraphs	5.0	D	0.06
					Inquiry			
Hillocks (1982)	7–8	full range	136	N	students examined or observed various activities and collected data to write about them versus writing that was facilitated by teacher discussion	5.0	J	0.14
Hillocks (1982)	7–8	full range	139	N	students examined or observed various activities and collected data to write about them and then revised papers versus writing that was facilitated by student discussion	5.0	J	−0.05
Hillocks (1979)*	9 & 11	full range	191	N	students examined or observed various activities and collected data to write about them versus instruction in paragraph writing	4.0	J	0.75
Widvey (1971)	11	full range	72	E	students formulated hypotheses, gathered and analyzed data, and made inferences to structure discourse versus traditional instruction (random assignment)	5.0	D	0.65
Pisano (1980)	11–12	full range	60	E	students asked to respond to questions designed to engage them in critical thinking about five literature topics versus teachers' regular questioning	5.0	D	−0.07

Table 4 (*continued*)

Study	Grade	Participant type	*n*	Genre	Treatment	Quality indicator score	Publication type	Effect size
					Procedural facilitation			
Graham et al. (1995)	4–6	SNL	43	N	students provided with procedural support to facilitate goal attainment in revising versus no support (random assignment)	8.0	J	−0.02
Scardamalia and Bereiter (1985)	6	full range	38	E	students provided with procedural planning support through the use of cue cards versus no support	3.0	B	0.67
Zellermayer et al. (1991)	6 & 9	full range	40	E	students provided with procedural support through unsolicited computer prompts designed to facilitate thinking while composing versus no support (random assignment)	7.0	J	1.37
Page-Voth and Graham (1999)	7–8	SNL	20	E	students provided with procedural support to facilitate goal attainment in writing versus no support (random assignment)	9.0	J	0.25
					Peer assistance when writing			
Prater and Bermúdez (1993)	4	SLL	46	N	peers helped each other choose topics and revised/edited text versus individual work on compositions (random assignment)	6.0	J	0.19
MacArthur et al. (1991)	4–6	SNL	29	N	peers helped each other revise and edit text versus process writing instruction	7.0	J	1.09
Boscolo and Ascorti (2004)	4, 6, 8	full range	122	N	peers helped each other revise text versus teacher correction of text	5.0	B	0.96
Yarrow and Topping (2001)	5–6	full range	26	N	peers helped each other draft, revise, and edit text after receiving training in a process for composing text versus peers composing alone after this training	8.5	J	0.58
V. B. Olson (1990)	6	full range	44	N	peers provided each other with feedback on their work versus grammar instruction/individual writing	3.0	J	0.42
Hill (1990)	8	high	36	N, E	peers composed together on a word processor versus peers composing alone on a word processor (random assignment)	6.0	D	0.46
Dailey (1991)	HS	SNL & basic	60	N	peers helped each other plan, draft, and revise text versus individuals composing alone	8.0	D	1.18
					Study of models			
Knudson (1989)	4, 6, 8	high	46	E	students examined model pieces of writing to direct construction of their papers versus free writing (random assignment)	5.0	J	0.26
Knudson (1991)	4, 6, 8	full range	76	E	students examined model pieces of writing to direct construction of their papers versus free writing (random assignment)	5.0	J	0.24
A. E. Thibodeau (1964)	6	full range	408	N	students examined model pieces of writing for both narrative and expository writing versus traditional language arts instruction	5.0	D	0.44
Reedy (1964)*	9	full range	410	E	students examined model composition for six patterns used to organize expository writing versus instruction in the process of communication in writing	5.0	D	0.26
Vinson (1980)	9	full range	118	E	students examined models used to illustrate concrete detail, sensory imagery, unnecessary detail, and single impression versus writing paragraphs with emphasis on correction of first drafts	5.0	D	−0.29
Caplan and Keech (1980)*	12	full range	129	N	students examined models used to illustrate difference between showing and telling	6.0	R	0.11
					Product goals			
Graham et al. (1995)	4–6	SNL	43	N	goal to add three pieces of information while revising versus goal to make paper better (random assignment)	8.0	J	0.77
Schunk and Swartz (1993a), Experiment 1	5	full range	30	E	goal to write a certain type of paragraph versus goal to do best (random assignment)	8.0	J	1.69

(*table continues*)

Table 4 (*continued*)

Study	Grade	Participant type	*n*	Genre	Treatment	Quality indicator score	Publication type	Effect size
Schunk and Swartz (1993a), Exp. 2	5	full range	20	N, E	goal to write a certain type of paragraph versus goal to do best (random assignment)	8.0	J	1.01
Ferretti et al. (2000)	6 & 8	full range & SNL	124	E	goal to include common persuasive elements verses general goal to write a persuasive paper (random assignment)	5.0	J	0.38
Page-Voth and Graham (1999)	7–8	SNL	20	N, E	goal to include common persuasive elements versus a general goal (random assignment)	9.0	J	1.18
					Feedback			
Lumbelli et al. (1999)	6	full range	28	E	students received feedback by watching adult talk about clear and unclear sections of text written by student versus no-treatment condition (random assignment)	4.0	J	0.87
Couzijn and Rijlaarsdam (2005)	9	average	30	E	students received feedback by watching others try to execute written task directions student versus writing text (random assignment)	5.0	B	2.52
Benson (1979)*	10	full range	288	E	peers provided feedback to student on five aspects of their writing versus teacher feedback	6.0	D	0.28
Covill (1996)	10–11	high & basic	68	N	teacher provided students with feedback on content of their paper versus feedback on surface features of paper (random assignment)	7.0	D	−0.25
Duke (2003)	10–12	full range	164	E	teacher provided students with feedback on their writing versus teacher modeling how to plan (random assignment)	7.0	D	−0.61
					Word processing			
Cheever (1987)	4	full range	50	N	students composed on a computer 1 day a week throughout the school year	4.0	D	0.30
Jackiewicz (1995)	4	full range	58	E	students used a word processor during computer lab for 12 weeks	4.0	D	1.74
Moore (1987)	4–5	full range	204	N, E	students used a word processor for 10 weeks	7	D	0.44
Lichtenstein (1996)	5	full range	32	?	students used word processing for all their writing assignments for 20 weeks	4.0	D	0.75
Espinoza (1992)	6	full range	68	E	students used a word processor for 6 weeks	5.0	D	0.03
Miller (1984)	6	high	28	N	students used a word processor for 4 weeks	6.0	D	−0.09
Hagler (1993)	6	full range	76	N	students used a word processor for 1 week (random assignment)	5.0	D	0.97
Dybdahl et al. (1997)	6	full range	41	?	students used a word processor for 28 weeks	3.0	J	−0.18
Lowther et al. (2003)	6–7	full range	118	E	students had round-the-clock access to laptop computers during the school year	4.0	J	1.11
Shinn (1986)	6–8	SNL & full range	18	N	students composed on a word processor and received problem solving instruction on the computer for 12 weeks	4.0	D	1.38
Lytle (1987)	7	full range	84	N	students used a word processor for 7 weeks	7.0	D	−0.05
Dalton and Hannafin (1987)	7	basic	64	E	students completed all writing assignments for 1 year with word processing (random assignment)	5.0	J	0.28
Lerew (1997)	8	basic	150	E	students used a word processor for 20 weeks (random assignment)	5.0	D	0.89
Head (2000)	8	full range	46	E	students used a word processor for 4 weeks	6.0	D	0.01
Lam and Pennington (1995)	9	basic	17	N	students used a word processor for 1 full year	4.5	J	0.33
Philhower (1985)	9–12	SNL	22	?	students used a word processor for 16 weeks	5.5	D	0.51
Cirello (1986)	10	basic	30	E	students received 20 weeks of remedial writing instruction while using a word processor	6.5	D	1.10
Silver and Repa (1993)	HS	SLL	66	?	students used a word processor for 13 weeks	4	J	0.52
					Extra writing			
Hook (1986)	4	high	22	N	expressive writing versus transactional writing (expressive writing assessed at posttest)	3.5	D	0.18
Knudson (1989)	4, 6, 8	full range	99	E	extra writing versus three other instructional procedures (models, scales, and models/scales; random assignment)	5.0	J	0.02

Table 4 (*continued*)

Study	Grade	Participant type	*n*	Genre	Treatment	Quality indicator score	Publication type	Effect size
Knudson (1991)	4, 6, 8	full range	159	E	extra writing versus three other instructional procedures (models, scales, and models/scales; random assignment)	5.0	J	−0.04
Gomez et al. (1996)*	5	SLL	68	N	free writing versus skills instruction (random assignment)	8.0	J	−0.21
Wienke (1981)	6	full range	157	N	gradual increase in writing versus traditional instruction	3.0	E	0.11
Duin and Graves (1987)	7	full range	54	E	writing added to vocabulary instruction versus vocabulary instruction	6.0	J	0.90

Note. HS = high school. For participant type, full range = normal variation in regular classroom; high = above average writers; average = average writers; basic = struggling writers; SLL = second language learners; SNL = special needs learners (students with learning disabilities and speech and language difficulties). For genre, N = narrative; E = expository. For publication type, D = dissertation or thesis; J = journal; B = book; R = report; P = conference presentation. A single asterisk indicates that effect size was computed with gain scores. In the *Process writing approach* section, two asterisks indicate that treatment involved professional development. In the *Strategy instruction* section, three asterisks designate a self-regulated strategy development model study. A question mark indicates that the information is unknown. The quality indicator scores range from 0 to 9. When students were randomly assigned to conditions, this was noted in the *Treatment* section.

was the smallest number of effect sizes included in any treatment analyzed by Hillocks (1986) in his seminal review. Because our goal was to draw a broad set of recommendations for teaching writing, we adopted a liberal, exploratory approach, applying the principle that we should make the best of the available data (Pressley, Graham, & Harris, 2006). Nevertheless, we recognize that small sample sizes are not very reliable, and we do not report a summary statistic for a treatment with a small sample and considerable variation in effect sizes. We also based this decision on the cohesiveness of the experimental interventions in a treatment as well as the cohesiveness of conditions to which the treatment was compared.

The quality indicators for each study were summed to provide a total quality score for each study. We examined whether total study quality was related to magnitude of effect size for all studies and each treatment with 10 or more effect sizes (using an analog to regression analysis; see Lipsey & Wilson, 2001). In addition, we examined the percentage of studies overall and for each treatment that satisfied each of the nine quality indicators.

Our meta-analysis used a weighted fixed-effects model. For each treatment, we calculated the mean, standard deviation, and median for the unweighted effect sizes. We also calculated the mean and confidence interval for weighted effect sizes (when the confidence interval does not intersect 0, then the average weighted effect size is significantly greater than no effect). Weighted effect sizes were computed by multiplying each effect size by its inverse variance. Although it is best to interpret the magnitude of an effect size (e.g., peer assistance) in relation to the distribution of other mean effect sizes in the same general area (i.e., other writing treatments), a widely used rule of thumb is that an effect size of 0.20 is small, 0.50 is medium, and 0.80 is large (Lipsey & Wilson, 2001).

In addition to calculating the average weighted effect size, we conducted a test of homogeneity to determine if the various effect sizes weighted and averaged together in a treatment estimated the same population effect size. When variability in effect sizes was larger than expected on the basis of sampling error alone (i.e., the homogeneity test was statistically significant), and there were at least 18 effect sizes for the treatment, we examined whether this excess variability could be accounted for by identifiable differences between studies (e.g., grade level). Using a fixed-effects model (Lipsey & Wilson, 2001), we grouped effect sizes into two mutually exclusive categories (i.e., effect sizes calculated from expository vs. narrative writing), and we tested both the homogeneity of effect sizes within each category and the difference between the levels of the two mutually exclusive categories (i.e., the independent variable). The analysis involved an analog that is similar to a one-way ANOVA (Hedges & Olkin, 1985). The independent variables that were tested were grade (4–6 vs. 7–12), type of student (struggling writers vs. full range), writing genre (expository vs. narrative), publication type (journal vs. other publication outlets combined), and student assignment (random vs. nonrandom). The statistical package for the analyses in this review was MetaWin (Rosenberg, Adams, & Gurevitch, 2000).

Results

Table 3 includes the average total quality score for studies in each of the 15 treatments that contained four or more effect sizes (see Table 1 for a description of these treatments). The total quality score was the sum of the scores for the nine quality indicators (e.g., pretest equivalence, treatment fidelity, etc.). For each treatment, we also report the percentage of studies in which each quality indicator was present. This same information is presented cumulatively for all studies included in the 15 treatments, providing a general indication of the quality of experimental and quasi-experimental writing intervention research. There were 127 effect sizes in the 15 treatments, but as noted earlier they were not all from different studies (e.g., some studies compared 3 or more different treatments). The effect sizes for these 15 treatments came from 107 different studies.

Table 4 contains information on each study for treatments that contained four or more effect sizes. Treatments are arranged in the same order that the findings are discussed in text. Studies presented under each treatment report the following information: reference, grade, type of student, number of participants, genre of posttest, a brief description of the treatment and the control condition (random assignment of students to conditions is noted as

well), total quality indicator score (the sum of the nine quality indicators), publication type, and effect size. If one of these pieces of information was unknown, then a question mark was used to indicate that this was the case. If the effect size was calculated with gain scores, an asterisk was placed after the reference. Professional development studies were signaled in the process writing approach treatment by placing two asterisks after the description of the treatment and the control condition, whereas SRSD studies were noted in the strategy instruction treatment by placing three asterisks in the same location.

Treatments that contained fewer than four effect sizes are presented in the Appendix. The only exception involves the treatments of individualized instruction and the environmental approach. As noted earlier, we decided not to retain these two treatments, even though they were included in Hillocks's (1986) earlier review. The distinctiveness and clarity of these two treatments were questionable, and most of the studies in both treatments were appropriately included in another treatment (e.g., three environmental studies involved inquiry). Three individualized instruction studies and two environmental studies, however, could not be placed in a different treatment, so they are included at the start of the Appendix. They are followed in order by treatments discussed in the text of this article. All other treatments with four or fewer studies are then presented in alphabetical order. The Appendix presents the following information on each study: reference, grade, type of student, number of participants, genre of posttest, a brief description of the treatment and the control condition, publication type, and effect size.

Table 5 contains summary statistics for all treatments with four or more effect sizes. These were arranged by type of instruction (process writing approach, explicit teaching, scaffolding, and alternative modes of composing). The summary statistics for each treatment include the following: number of effect sizes, unweighted mean effect size and standard deviation, unweighted median effect size, weighted mean effect size and confidence interval, and the *Q* statistic for homogeneity test. These are presented in the order that they are discussed in text. We did not include four treatments (text structure, procedural facilitation, feedback, and extra writing) in Table 5, because we were not able to compute summary statistics for them (because of the small number of studies, variable effect sizes, and/or disparate treatment or control conditions).

How Much Experimental Evidence on Writing Instruction for Adolescents Has Accumulated?

We first asked what evidence had accumulated on adolescent writing instruction. We were able to compute 154 different effect sizes involving writing instruction for students in Grades 4 through 12. Only 24 of these effect sizes were included in the last comprehensive review of this literature (Hillocks, 1986), whereas another 35 were included in the other four reviews conducted since then (Bangert-Drowns, 1993; Graham, 2006a; Graham & Harris, 2003; Goldring et al., 2003). Thus, 62% of the effect sizes in this meta-analysis were not included in previous reviews.

Investigations of adolescent writing instruction have covered a broad range of topics, as evidenced by the 26 separate treatments we identified (see Tables 3 and 4 and the Appendix). Although the depth of investigation is relatively shallow in all but three areas (i.e., process writing approach, strategy instruction, and word

Table 5
Summary of Experimental–Control Effect Size Statistics for Writing Intervention Treatments

Treatment	*n*	Unweighted mean effect	*SD*	Unweighted median effect	Weighted mean effect	95% confidence interval		Homogeneity
						Lower	Upper	
Process approach	21	0.09	0.46	0.14	0.32*	0.26	0.37	76.23*
Professional development	6	0.44	0.23	0.40	0.46*	0.37	0.56	5.68
No professional development	15	−0.05	0.46	−0.03	0.03	−0.07	0.13	37.03*
Grades 4–6	7	0.12	0.37	−0.03	0.27*	0.05	0.48	7.84
Grades 7–12	8	−0.20	0.56	−0.06	−0.05	−0.18	0.07	19.67*
Explicit teaching								
Grammar	11	−0.22	0.61	−0.22	−0.32*	−0.43	−0.21	47.02*
Sentence combining	5	0.46	0.18	0.42	0.50*	0.30	0.70	3.11
Strategy instruction	20	1.03	0.81	0.93	0.82*	0.69	0.95	77.66*
SRSD	8	1.15	0.41	1.17	1.14*	0.89	1.39	8.92
Non-SRSD	12	0.95	1.00	0.69	0.62*	0.44	0.80	54.98*
Summarization	4	0.83	0.38	0.95	0.82*	0.50	1.14	3.87
Scaffolding students' writing								
Prewriting	5	0.42	0.36	0.32	0.32*	0.08	0.56	7.04
Inquiry	5	0.28	0.39	0.14	0.32*	0.15	0.48	18.21*
Peer assistance	7	0.70	0.38	0.58	0.75*	0.54	0.97	9.85
Study of models	6	0.17	0.25	0.25	0.25*	0.13	0.36	12.12
Setting product goals	5	1.00	0.49	1.01	0.70*	0.43	0.96	9.99
Alternative approaches to composing: Word processing	18	0.56	0.55	0.48	0.55*	0.43	0.67	61.00*

Note. Confidence intervals that do not include zero in their range from lower to upper effect size are significantly greater than no effect. A significant homogeneity test indicates that heterogeneity among effect sizes is greater than expected if all studies shared a common underlying effect size. SRSD = self-regulated writing development; non-SRSD = strategies taught with a model other than self-regulated strategy development.
* $p < .05$.

processing), there are 15 treatments in Table 4 that contain at least four effect sizes.

Quality of Studies

All of the studies included in this meta-analysis were selected so that the following two quality indicators were present: Students in the treatment condition were compared with students in a control/comparison condition, and reliability of the outcome measure was established. We further assessed the quality of research for the 15 treatments in Table 2 by evaluating whether nine specific quality indicators were present or absent in each study. The average quality of research for specific treatments varied from a low of 4.5 for investigations studying the process approach to writing to a high of 7.6 for studies examining the impact of setting product goals (see Table 3).

Five of the nine quality indicators were present in the typical study included in Table 4; this ranged from a low of two quality indicators present to a high of nine. In the typical study, mortality of participants was limited and similar across conditions. Ceiling and floor effects were not evident for the posttest quality measure, and the control conditions were described and included something other than a no-treatment comparison condition (see Table 3). In addition, close to three out of every five studies implemented appropriate controls for the Hawthorne effect, and students in the different conditions were equivalent in terms of the quality of their writing at baseline. Nevertheless, randomization of students to conditions occurred in only one third of the studies, and teacher effects were controlled in less than one half of the investigations. Likewise, procedures for training instructors or teachers to implement the experimental treatment were reported in less than one half of the studies, and data on the fidelity of treatment implementation were provided only 27% of the time.

When all studies were considered together, there was little relation between overall quality of studies (based on our nine indicators) and obtained effect sizes; the correlation was .07. It does appear, however, that the quality of experimental and quasi-experimental research has improved over time, as more recent investigations had higher overall quality scores than earlier ones ($r = .32$, $p < .01$). Similarly, more recent investigations have produced higher effect sizes than earlier investigations ($r = .31$, $p < .01$). Surprisingly, there was no statistically significant difference between the quality of studies published in peer-reviewed journals and other outlets, such as books, dissertations, and so forth ($p > .46$). In a series of one-way ANOVAs, neither publication type nor the separate quality indicators were related to obtained effect sizes (all ps $> .15$). The only exception involved the quality indicator for mortality equivalence. The average effect size was 0.59 ($SD = 0.69$) when this criteria was met versus 0.24 ($SD = 0.28$) when it was not met, $F(1, 115) = 5.85$, $MSE = 0.39$, $p = .017$.

Process Approach to Writing Instruction

We calculated 21 effect sizes from studies that included a process writing condition or professional development in this approach (students in Grades 4 through 12; see Table 4). Five of these effect sizes were from studies included in Hillocks's (1986) meta-analysis (Adams, 1971; Alloway et al., 1979; Ganong, 1974; Gauntlett, 1978; M. C. Olson & DiStefano, 1980). In our process writing treatment, we did not include 1 effect size from a study by Wienke (1981) that was in Hillocks's review, because it mainly assessed the effects of increased writing, and not the process approach.

The average weighted effect size for the process writing approach was 0.32, and this was significantly greater than no effect, because the confidence interval did not include 0 (see Table 5). Effect sizes varied substantially, as verified by the test of homogeneity. Consequently, we examined whether specific study features moderated average weighted effect size and accounted for excess variability. Because 18 effect sizes were based on the full range of students in regular classrooms, we were not able to test whether student type moderated average weighted effect size. An identical situation existed for assignment to conditions (only 3 effect sizes were from studies involving random assignment of students).

Although we did not find that publication source (journal article vs. other publication source), grade (4–6 vs. 7–12), genre (narrative vs. expository), or overall quality of study (calculated by summing the nine quality indicators) moderated effect size (all ps $> .18$), we found that professional development was associated with larger effect sizes, Q(between) $= 55.14$, $p = .001$. The average weighted effect size for professional development (0.46) was greater than no effect, but this was not the case when professional development was not provided (0.03). In addition, variability of effect sizes for professional development did not exceed sampling error alone (see Table 5). All 6 of the effect sizes for professional development were positive, whereas 7 of the 15 effect sizes for nonprofessional development were negative (see Table 4).

The effect sizes for nonprofessional development varied substantially (see Table 5). As a result, we examined whether these effect sizes were related to specific study features. Even though we did not find that genre (narrative vs. expository) or study quality moderated these effect sizes (both ps $> .23$), students in Grades 4 through 6 had a higher average weighted effect size (0.27) than students in Grades 7 through 12 (–0.05), Q(between) $= 9.53$, $p = .012$. The average weighted effect size was statistically significant for Grades 4 through 6 but not for Grades 7 through 12. Likewise, variability of effect sizes for younger students was not greater than sampling error alone, but it was greater than sampling error alone for the older students (see Table 5).

It is important to note that there was considerable variation in the control conditions for the 15 nonprofessional development effect sizes. This treatment was contrasted with strategy instruction in four studies, with instruction in basic writing skills (such as grammar) in five studies, and with a more general traditional writing program in five studies. Nevertheless, the average weighted effect size was not moderated by these different types of control conditions ($p = .12$).

In summary, when teachers were involved in professional development to use the process writing approach, there was a moderate effect on the quality of students' writing. In the absence of such training, process writing instruction had a small effect on the writing of students in Grades 4 through 6, but did not enhance the writing of students in Grades 7 through 12.

Explicit Teaching

Grammar. We calculated 11 effect sizes for grammar instruction (students were in Grades 4 through 11; see Table 4). Five of these effect sizes were from studies included in Hillocks's (1986) earlier meta-analysis (Elley, Barham, Lamb, & Wyllie, 1975, Comparisons 1 and 2; Howie, 1979; A. E. Thibodeau, 1964; A. L. Thibodeau, 1964).

Seven of the effect sizes were negative, and the average weighted effect size was −0.32, which was statistically significant (see Table 5). There was more variability among effect sizes than could be attributed to sampling error alone, and much of this variability appeared to have been due to two outliers: an effect size of −1.40 for A. A. Anderson (1997) and 1.07 for Fearn and Farnan (2005). When we removed these outliers from the analysis, the average weighted effect dropped slightly to −0.34, and the variability among effect sizes was then small enough that it was attributable to sampling error alone, Q(total) = 17.02. This average negative effect size was still statistically significant (i.e., the confidence interval ranged from −0.46 to −0.23). Finally, magnitude of effect size was not predicted by overall quality of grammar research ($p = .18$).

Findings regarding grammar instruction must be interpreted cautiously, because the grammar treatment was the control condition for all but one of the obtained effect sizes. In addition, there was considerable variability among the comparison conditions to which grammar instruction was compared. The comparison condition for three effect sizes was sentence combining, two comparison conditions involved the process writing approach, two involved reading and writing, two involved elaborative thinking and vocabulary enrichment activities in writing, and one involved strategy instruction. It is important to note, however, that grammar instruction was not an effective treatment in any of these comparisons (as average weighted effect sizes for each cluster with two or more effect sizes ranged from −0.43 to 0.04).

Sentence-combining instruction. An alternative approach to more traditional grammar instruction is sentence combining. This involves teaching students to construct more complex and sophisticated sentences through exercises in which two or more basic sentences are combined into a single sentence. We computed five effect sizes for this treatment (students were in Grades 4–11; see Table 4); two of these effect sizes (Howie, 1979; Pedersen, 1977) were from studies included in Hillocks (1986). The control conditions for sentence combining were relatively homogeneous, primarily involving grammar instruction.

All five effect sizes were positive and yielded an average weighted effect size of 0.50, which was greater than no effect (see Table 5). Furthermore, the effect sizes appeared to be from a single population, because the test for homogeneity was not statistically significant. Thus, sentence combining had a moderate impact on the quality of students' writing.

Strategy instruction in planning, revising, and editing. Strategy instruction involves directly and explicitly teaching students how to independently use strategies for planning, revising, and/or editing text. We calculated 20 effect sizes for strategy instruction (students were in Grades 4–10; see Table 4). In some instances, these strategies involved teaching more generic processes, such as brainstorming (e.g., Troia & Graham, 2002) or a peer-revising strategy (MacArthur et al., 1991). In other instances, the strategies involved teaching strategies for accomplishing specific types of writing tasks, such as writing a story (Fitzgerald & Markham, 1987) or a persuasive essay (Yeh, 1998). All but 5 of these effect sizes (A. A. Anderson, 1997; Curry, 1997; Gamelin, 1996; Scardamalia, Bereiter, & Steinbach, 1984; Walser, 2000) were from studies included in either Graham (2006a) or Graham and Harris (2003).

All 20 of the effect sizes were positive (see Table 4), and the average weighted effect size was large (0.82) and statistically significant (see Table 5). There was considerable variability in effect sizes, however, as the test of homogeneity was significant. Consequently, we examined whether specific study characteristics moderated average effect size and accounted for excess variability.

Neither publication source (journal article vs. other publication source), grade (4–6 vs. 7–12), genre (narrative vs. expository), student assignment (random vs. nonrandom), nor overall study quality moderated the effects of strategy instruction (all ps $> .23$). Of the 20 effect sizes, 9 were based on the performance of struggling writers (basic writers or special needs learners), 9 were based on the full range of students in the regular classroom, and 2 involved both groups of writers (see Table 4). Before examining whether type of student acted as a moderator, we calculated a separate effect size for struggling writers and the full range of writers for these last 2 effect sizes (taken from A. A. Anderson, 1997, and Bryson & Scardamalia, 1996). The subsequent analysis revealed that the average weighted effect size for struggling writers (1.02) was larger than the average weighted effect size for students in regular classrooms (0.70), Q(between) = 6.06, $p = .01$. These effect sizes were significantly greater than no effect for both groups of students, but type of student did not reduce excess variability (see Table 5).

We also examined whether SRSD instruction (Harris & Graham, 1996, 1999) moderated strategy effects. In a previous meta-analysis involving students in all grades, Graham (2006a) found that this form of strategy instruction yielded a larger average effect size than all other instructional approaches combined. Our analyses confirmed this earlier finding: The average weighted effect size for SRSD (1.14) was larger than the average weighted effect size for non-SRSD interventions (0.62), Q(between) = 14.65, $p < .001$. The average weighted effect size was significantly greater than no effect for both SRSD and non-SRSD interventions, and type of instruction accounted for some of the excess variance because effect sizes for SRSD were not greater than expected sampling error alone (see Table 5).

It is important to note that there was considerable variation in the control conditions to which strategy instruction was compared. The control conditions ranged from teaching text structure (for four effect sizes), the process writing approach (for four effect sizes), traditional instruction (for four effect sizes), unspecified control condition (for five effect sizes), practice writing (for two effect sizes), and literature study (for one effect size). The average weighted effect sizes for strategy instruction for each type of comparison condition containing four or more studies were not statistically significant ($p = .44$) and ranged from 0.50 (process approach to writing) to 1.08 (unspecified control). Thus, strategy instruction was effective across different experimental–control comparisons.

Although the available number of effect sizes was small ($N = 3$), there was tentative evidence that process goal setting can

enhance strategy instruction, at least for normally achieving and gifted students (see the Appendix). When students were explicitly told that the goal of instruction was to learn the strategies taught and when they received feedback on their progress in accomplishing this objective, effect sizes were large, ranging from 1.12 to 2.14 (Schunk & Swartz, 1993a, 1993b).

It is also interesting to note that two effect sizes provided some initial evidence that strategy instruction can enhance the power of the process writing approach with struggling writers. When Curry (1997) added strategy instruction in planning to the process writing model, the quality of text produced by students with special needs improved (effect size = 0.69; see the Appendix). Likewise, MacArthur et al. (1991) found that teaching a revising/editing strategy enhanced the writing quality of special needs learners in writing process classrooms (effect size = 1.09; see Table 4).

In summary, explicitly teaching adolescents strategies for planning, revising, and/or editing had a strong impact on the quality of their writing. This was especially the case for struggling writers and when the SRSD model was used to teach strategies. Although the effects were not as dramatic in non-SRSD studies and when instruction was evaluated with students whose writing skills provided a more normal range of variation, strategy instruction still had a moderate-to-strong impact on writing quality.

Summarization instruction. We calculated four effect sizes for summarization instruction (students were in Grades 5–12; see Table 4). Summarization instruction ranged from explicitly teaching summarization strategies (Bean & Steenwyk, 1984; Knight, 2003; Placke, 1987) to enhancing summarization by progressively fading models of a good summary (Chang, Sung, & Chen, 2002). The control conditions in summarization instruction were relatively homogeneous, primarily involving no instruction or no explicit instruction in summarization.

The effect size for each study was positive, and the average weighted effect size was large (0.82) and significantly greater than no effect size (see Table 5). In addition, variability in effect sizes could be explained by sampling error alone. Thus, teaching adolescents to summarize text had a strong impact on their ability to write more concise text.

Text structure instruction. We were able to compute five effect sizes involving teaching students the structure of specific types of text (students were in Grades 4–10; see Table 4). This treatment involved teaching the basic elements of a story (e.g., Fitzgerald & Teasley, 1986) or a persuasive essay (Scardamalia & Paris, 1985). Effect sizes diverged widely, ranging from 0.75 to –1.27 (see Table 4), $Q = 17.25$, $p = .002$, as did the control conditions (two no treatment, one strategy instruction, one literature, and one poetry). The small number of effect sizes, the disparate findings, and the variability in control conditions, made it impossible for us to draw any reliable or meaningful conclusions for this treatment.

Scaffolding Students' Writing

Prewriting activities. We calculated five effect sizes where students participated in prewriting activities before composing (students in Grades 4–9; see Table 4). This included encouraging planning before writing (Brodney, Reeves, & Kazelskis, 1999), group and individual planning before writing (Vinson, 1980, which was also included in Hillocks, 1986), reading topic-pertinent material and being encouraged to plan in advance (Brodney et al., 1999), organizing prewriting ideas by means of a semantic web (Loader, 1989), and prompting planning following a brief demonstration of how to plan (Reece & Cumming, 1996). The comparison conditions differed considerably, because they involved listening to information on the writing topic (two studies), listing ideas (one study), writing paragraphs (one study), and no treatment (one study). Despite the differences in control conditions, the effect size for each of the five studies was positive, the average weighted effect size was 0.32, which was significantly greater than no effect, and variability in effect sizes could be explained by sampling error alone (see Table 5). Collectively, these investigations show that prewriting activities had a positive and small impact on writing quality.

Inquiry. Hillocks (1986) computed five effect sizes for inquiry activities, in which students analyzed data and information before writing. We were unable to locate any other studies that applied this type of instruction. The control conditions against which inquiry was compared were relatively homogeneous, primarily involving teacher-facilitated writing.

All five effect sizes that we computed were positive (students were in Grades 7–12; see Table 4). Together, these effect sizes had a small (0.32) and statistically significant impact on writing quality (see Table 5). Variability among effect sizes, however, exceeded what would be expected by sampling error alone. It is important to note that in Hillock's (1986) original analysis, all of the effect sizes had been corrected for preintervention differences. With such an adjustment, the average weighted effect size increased to 0.52 (confidence interval ranged from 0.35 to 0.68), and the effect sizes appeared to be from the same population effect size, because the homogeneity test was not statistically significant, $Q(\text{total}) = 7.13$. Thus, inquiry had a small-to-moderate impact on writing quality.

Procedural facilitation. We calculated four effect sizes involving the impact of procedural facilitation on students' writing (students were in Grades 4–9; see Table 4). Although the control conditions against which procedural facilitation was compared were similar (i.e., no procedural facilitation) and variability in effect sizes could be explained by sampling error alone, $Q(\text{total}) = 9.82$, we did not draw any conclusions regarding the impact of procedural facilitation. We based this decision on the small number of effect sizes and the diversity in instructional procedures. For instance, instructional procedures ranged from the use of cue cards to promote planning (Scardamalia & Bereiter, 1985) to checklists reminding students to carry out specific processes (Page-Voth & Graham, 1999) to guidance from unsolicited prompts delivered while composing on the computer (Zellermayer, Salomon, Globerson, & Givon, 1991).

Peer assistance when writing. We computed seven effect sizes for peers working together to plan, draft, and/or revise their compositions (students were in Grades 4 through high school; see Table 4). The control conditions against which peer assistance was compared always involved students writing alone. All seven effect sizes were positive. The average weighted effect size was large (0.75) and greater than no effect, and variability in effect sizes could be attributed to sampling error alone (see Table 5). Collectively, these investigations show that collaborative arrangements where students help each other with one or more aspects of their writing had a strong and positive impact on writing quality.

Study of models. Six effect sizes were calculated for students' study of models (students were in Grades 4–12; see Table 4). Study of models involved students examining examples of one or more specific types of text and attempting to emulate the patterns or forms in these examples in their own writing. Four of the effect sizes (Caplan & Keech, 1980; Reedy, 1964; A. E. Thibodeau, 1964; Vinson, 1980) were from studies included in Hillocks's (1986) meta-analysis. There was considerable variability in the control conditions, as two involved free writing, one focused on paragraph writing, one included traditional language arts instruction, and another concentrated on the process of communication in writing (we were unable to determine what happened in the control condition in Caplan & Keech, 1980). Despite the variability in control conditions, five of the six effect sizes were positive, the average weighted effect size was small (0.25) and greater than no effect, and variability in effect sizes could be attributed to sampling error alone (see Table 5). Thus, the study of models resulted in small improvements in writing quality.

Product goals. Five effect sizes were computed for the product goals treatment (students were in Grades 4–8; see Table 4). This treatment included assigning a goal to add more ideas to a paper when revising (Graham, MacArthur, & Schwartz, 1995), establishing a goal to write a specific kind of paper (Schunk & Swartz, 1993a) and assigning goals to include specific structural elements in a composition (Ferretti, MacArthur, & Dowdy, 2000; Page-Voth & Graham, 1999). The control conditions all involved assigning a general goal. All five effect sizes were positive, the average weighted effect size was 0.70, which was significantly greater than no effect, and variability in effect sizes could be attributed to sampling error alone (see Table 5). Thus, assigning product goals had a strong impact on writing quality.

Feedback. We calculated five effect sizes that examined the impact of feedback on the quality of students' writing (students were in Grades 5–12; see Table 4). It is likely that the broad range of obtained effect sizes for studies in this treatment (2.52 to –0.65; $Q = 51.74$, $p < .001$) reflected the diversity of feedback procedures studied, ranging from written comments from the teacher (Covill, 1996; Duke, 2003) to feedback from peers (Benson, 1979) and feedback from watching others try to execute the task described in one's paper (Couzijn & Rijlaarsdam, 2005). There was also considerable variability in control conditions (ranging from no feedback to modeling of writing processes). Consequently, the small number of effect sizes, the diversity of instructional procedures and control conditions, and the disparate findings made it impossible for us to draw any reliable or meaningful conclusions for this treatment.

Alternative Modes of Composing: Word Processing

We computed 18 effect sizes for word processing (students were in Grades 4–12; see Table 4). Eight of these were from studies included in the meta-analysis by Bangert-Drowns (1993), and the other 10 were from studies reviewed by Goldring et al. (2003). For all of the effect sizes, the control condition involved students composing their papers by hand.

The effect size for most of these studies ($N = 16$) was positive, suggesting that word processing had a fairly consistent, positive impact on the quality of students' writing. This positive impact was reflected in an average weighted effect size of 0.55, which was significantly greater than no effect (see Table 5). Effect sizes, however, varied considerably among word-processing studies. Consequently, we examined whether specific study characteristics moderated average effect size and accounted for excess variability. We were not able to examine whether student assignment procedures were correlated with magnitude of effect sizes, because random assignment of students occurred with three effect sizes.

Neither grade (Grades 4–6 vs. 7–12), publication type (journal vs. nonjournal), genre of posttests (narrative vs. expository), type of student (struggling writers vs. full-range of regular class performance), nor overall quality of studies moderated the average weighted effect size for word processing (all ps $> .11$). Following Huber's (1977) and Hillock's (1986) advice, we also dropped the comparisons with the two highest positive effect sizes and the two highest negative effect sizes. Even with this adjustment, there was still considerable variability in effect sizes, $Q(\text{total}) = 28.90$, $p = .007$. In summary, word processing had a moderate impact on the writing of students in Grades 4–12, but there was also considerable variability from one study to the next.

Other: Extra Writing

A frequent recommendation for improving students' writing is to increase how much they write (National Commission on Writing, 2003). This recommendation has been made repeatedly through the years because of concerns about how little writing students actually do (Applebee, 2000). We were able to calculate six effect sizes that examined the impact of some form of extra writing on the quality of students' writing (students were in Grades 4–8; see Table 4).

Even though variability in effect sizes could be explained by sampling error alone, because the homogeneity test was not statistically significant, $Q(\text{total}) = 10.13$, $p = .07$, we were unable to draw any reliable or meaningful conclusion concerning the impact of extra writing on the quality of students' writing. This was due to the small number of effect sizes, the diversity of procedures for providing extra writing, and variety in control conditions. For example, in one study (Weinke, 1981), amount of writing was gradually increased and compared with traditional instruction. In another study (Gomez et al., 1996), free writing was compared with skills instruction. Hook (1986) compared groups that did different types of writing, whereas Duin and Graves (1987) examined the benefit of adding writing to vocabulary instruction. Finally, Knudson (1989, 1991) compared extra writing with the use of models and evaluative scales.

Discussion

To reap the benefits of literacy in an advanced technological society, adolescents need to develop strong writing skills. Unfortunately, a majority of adolescents in the United States do not achieve this critical goal (National Commission on Writing, 2003; Persky et al., 2003). These youngsters' future aspirations are at risk, because strong writing skills are needed to be successful at the postsecondary level, to obtain more than menial employment, and to participate fully as an adult member of the community (Perin, 2001, in press). The findings from this meta-analysis demonstrate that there are a variety of instructional procedures that improve the quality of the writing of adolescent students.

Before summarizing the primary findings from this review, it is important to determine how much empirical evidence, as defined by experimental and quasi-experimental writing intervention studies with adolescents, has accumulated since Hillocks's (1986) comprehensive review. Only 16% of the effect sizes included here were contained in Hillocks's review. Just as important, 62% of the effect sizes in this meta-analysis were not included in all of the prior meta-analyses combined (Bangert-Drowns, 1993; Graham, 2006a; Graham & Harris, 2003; Goldring et al., 2003; Hillocks, 1986). These findings provide additional justification for conducting this meta-analysis and also show that this particular area of writing research has been robust. It should be noted that the available evidence on effective writing instruction is broader than the studies examined in our review, because we did not include investigations in which students acted as their own controls, nor did we include single-subject design, qualitative, or correlational studies. An important goal for those interested in adolescent writing instruction is to synthesize the findings from these various bodies of literature.

Caveats and Limitations

The primary purpose of this review was to identify effective practices for teaching writing to adolescents. We based our analysis on quantitative studies, in which the impact of a treatment on one group was compared with a similar group who received a different treatment (or in some instances no treatment). Although there is no universally agreed on age or grade range for adolescence, the analysis focused on students in Grades 4 through 12, casting a broad net that included not only middle school and high school, but upper-elementary grades as well. We also concentrated our analyses on writing quality, because our goal was to identify instructional treatments that had a broad impact on writing performance. With these constraints in place, we calculated mean weighted effect sizes for 11 treatments containing at least four effect sizes (as was done by Hillocks, 1986). We did not calculate an average weighted effect size for 4 treatments, because of a small number of effect sizes coupled with one or more of the following problems: heterogeneity among (a) effect sizes, (b) control/comparison conditions, or (c) treatment procedures. For 3 treatments (i.e., process writing approach, strategy instruction, and word processing), there were enough comparisons available to determine if study characteristics moderated study outcomes.

Before presenting our findings and recommendations, there are a number of caveats and limitations that must be addressed. First, this review was limited to experimental and quasi-experimental studies involving controlled tests of writing interventions, in which the impact of one treatment was compared with another. Our decision to focus on these types of studies should in no way distract from the important contributions that other types of research make to our understanding of how to teach writing (see Pressley et al., 2006). This includes qualitative studies examining the practices of effective teachers of writing (e.g., Pressley, Yokoi, Rankin, Wharton-McDonald, & Mistretta, 1997), investigations that examine the correlation between writing performance and particular teaching practices (e.g., Applebee, Langer, Nystrand, & Gamoran, 2003), studies in which students act as their own controls (e.g., Graham, 1990), and single subject design studies that closely monitor the effectiveness of an intervention with a small number of students (e.g., De La Paz, 1999).

Second, because we limited the analysis to studies that assessed the quality of students' writing, we can draw no conclusions about the viability of teaching basic text transcription skills, such as spelling or handwriting, to adolescents. We were unable to find any studies that examined whether such instruction had an impact on the writing quality of students in Grades 4 through 12.

Third, the studies included in this meta-analysis did not all use the same measure of writing quality. We addressed this issue by calculating effect sizes just from holistic measures (raters assigned a single score for overall merit, taking into account factors such as ideation, organization, vocabulary, sentence structure, and tone) whenever possible or by converting the effect sizes from analytic measures (raters assign separate scores to factors such as ideation, organization, vocabulary, etc.) into a single average effect size. Although the attributes used to assess writing quality were generally similar for both the holistic and analytic scales in the studies we reviewed, there was variability. This variability included not only the attributes assessed, but the number of points on the scales and how these points were operationalized. Thus, this variability must be taken into account when interpreting our findings.

Fourth, some instructional procedures have been the focus of more research than others. For example, vocabulary instruction (e.g., Duin & Graves, 1987) may be an effective procedure for improving students' writing (see studies reported in the Appendix), but there is not enough research currently available to draw even a tentative conclusion about its impact. In addition, there were only four treatments (strategy instruction, word processing, processing writing approach, and grammar instruction) that yielded 10 or more effect sizes. Less confidence can be placed in the reliability of an average effect size when it is based on a small number of studies. Thus, the findings in this review for sentence combining, summarization instruction, prewriting, inquiry, peer assistance, study of models, and product goals must be viewed as more tentative, because they are based on a relatively small number of effect sizes.

Fifth, even for some areas that included a large number of effect sizes, not all grade levels were covered. For strategy instruction, for instance, there were no studies beyond 10th grade. The results of our meta-analysis do not permit a determination of whether the interventions are also effective at other grade levels. Interpretation of the findings from this review must be tempered by this fact.

Sixth, we were able to examine whether type of student moderated outcome effects for only two treatments, the process writing approach and strategy instruction. We were particularly interested in determining whether a treatment was effective with struggling writers but were unable to do so in most instances. It must further be noted that struggling writers in this review represented a mix of students, including those with and without special needs.

Seventh, like Hillocks (1986), we had to make to make a host of decisions about what constituted a writing treatment (e.g., process writing approach) and how to group these treatments together into coherent categories of treatments (e.g., procedures for scaffolding students' writing). On the basis of the reaction to Hillocks's review (see, e.g., Stotsky, 1988), we have no doubt that other researchers will question one or more of the decisions that we made. As a result, we tried to make our reasoning and decision making in this regard as transparent as possible.

Eighth, even though the relationship between overall study quality, as measured by nine quality indicators, and magnitude of effect size was small (0.07) and not statistically significant, the average quality of studies for specific treatments varied as did the quality of studies within a treatment. Thus, conclusions for each of the treatments must be tempered by the quality of the research. For example, conclusions concerning the setting of product goals are more valid than those for the process writing approach, because the former was more methodologically sound.

Ninth, one concern with meta-analysis involves problems of dissimilar control comparisons. For studies looking at the effectiveness of a specific treatment, students in the control conditions may be treated differently from one study to the next. This situation clouds interpretation, because there is no common point of comparison. This is also a concern in the current review. For some treatments in the current review, such as product goal setting, the control/comparison conditions were homogeneous. For other treatments, such as the process writing approach or feedback, there was considerable variation in the control/comparison conditions. In some of these instances (e.g., the process writing approach), we were able to examine whether differences in control conditions were systematically related to the obtained average weighted effect size. In other instances (e.g., feedback), diversity in control/comparison conditions along with other factors (e.g., heterogeneity in effect sizes or variability in the treatments within a category) resulted in a decision not to report an average weighted effect size.

What Instructional Practices Improve the Quality of Adolescent Students' Writing?

We were able to calculate average weighted effect sizes for 11 treatments. For all but 1 treatment, the average weighted effect size was positive and statistically greater than zero. The only exception involved the teaching of grammar, as we obtained an average weighted effect size of −0.32. There are several reasons to be cautious in interpreting this finding. One, grammar instruction was the control condition in 10 of the 11 studies in this treatment. Two, the condition to which grammar instruction was compared varied considerably, as did the 11 effect sizes. Although 7 of the effect sizes were negative and 2 were negligible (i.e., 0.00 or 0.03), additional research is needed for the reasons noted earlier as well as to investigate the possibility that the type of grammar instruction in a study by Fearn and Farnan (2005) is effective. They found that teaching students to focus on the function and practical application of grammar within the context of writing (vs. defining and describing grammar) produced strong and positive effects on students' writing. This approach merits further investigation.

In summarizing our findings later, we include a recommendation, average weighted effect size, and grade range of students tested for each treatment where the average weighted effect size was positive and statistically greater than no effect (see Table 5). We also note treatments where findings need to be interpreted more cautiously because of variability in the comparison conditions. For the process writing approach, the recommendation was not based on the average weighted effect size for all 21 studies (which was 0.32 and greater than no effect), but was based on the average weighted effect size for the 6 studies in which teachers were provided with professional development in how to implement this approach. In addition, we provided a recommendation for students in Grades 4 through 6 when such preparation was not provided (on the basis of 7 studies). For the process writing approach, the average weighted effect size was moderated by professional development, and nonprofessional development study outcomes were moderated by grade level. In addition, the average weighted effect size for professional development studies and nonprofessional development studies in Grades 4 through 6 was positive and statistically greater than no effect. This was not the case, however, for nonprofessional development studies involving students in Grades 7 through 12.

Our 10 recommendations are ordered according to their average weighted effect size. Interventions with larger effect sizes are presented before interventions with smaller effect sizes. Identical mean weighted effect sizes were found for (a) strategy and summarization instruction as well as (b) inquiry and prewriting. For the first set, we listed strategy instruction first, because more studies examined the effectiveness of this treatment. For the second set, we listed inquiry first, because the mean weighted effect size was based on students from a broader range of grades than the average effect size for prewriting (the number of studies was the same). In addition, for strategy instruction, we first presented the average weighted effect size for all 20 studies. We also presented the average weighted effect sizes for type of instruction and type of student, because both of these study characteristics moderated the mean effect size for writing quality. Each of these comparisons yielded a positive average weighted effect size that was greater than zero. Our findings and recommendations are as follows:

1. Teach adolescents strategies for planning, revising, and editing their compositions (strategy instruction, mean weighted effect size = 0.82; Grades 4–10). This was an especially powerful method for adolescents who were struggling writers (mean weighted effect size = 1.02; Grades 4–10), but it was also effective with adolescents in general (mean weighted effect size = 0.70; Grades 4–10). The self-regulated strategy development model (Harris & Graham, 1996) appears to have been a particularly potent approach for teaching writing strategies (mean weighted effect size = 1.14; Grades 4–8), but other approaches were effective too (mean weighted effect size = 0.62; Grades 4–10). Although there was considerable variation in the comparison interventions, all effect sizes were positive, and when categorized by type of comparison condition (i.e., process writing, text structure, traditional, and unspecified), average weighted effect sizes ranged from medium (0.50 when compared with the process approach) to high (1.08 when compared with unspecified interventions).

2. Teach adolescents strategies and procedures for summarizing reading material, because this improves their ability to concisely and accurately present this information in writing (for summarization, mean weighted effect size = 0.82; Grades 5–12).

3. Develop instructional arrangements in which adolescents work together to plan, draft, revise, and edit their compositions. Such collaborative activities have a strong impact on the quality of what students write (for peer assistance, mean weighted effect size = 0.75; Grade 4 through high school).

4. Set clear and specific goals for what adolescents are to accomplish with their writing product. This includes identifying the purpose of the assignment (e.g., to persuade) as well as characteristics of the final product (e.g., addresses both sides of an argument; for setting product goals, mean weighted effect size = 0.70; Grades 4–8).

5. Make it possible for adolescents to use word processing as a primary tool for writing, because it has a positive impact on the quality of their writing (for word processing, mean weighted effect size = 0.55; Grades 4–12).

6. Teach adolescents how to write increasingly complex sentences. Instruction in combining simpler sentences into more sophisticated ones enhances the quality of students' writing (for sentence combining, mean weighted effect size = 0.50; Grades 4–11).

7. Provide teachers with professional development in how to implement the process writing approach when this instructional model is used with adolescents (for the process writing approach with professional development, mean weighted effect size = 0.46; Grades 4–12). Although the process approach to writing may still be effective with students in Grades 4–6, when such training is not provided (for the process writing approach without professional development with Grades 4–6, mean weighted effect size = 0.27), this does not appear to be the case for students in Grades 7 through 12 (for the process writing approach without professional development with Grades 7–12, mean weighted effect size = −0.05). Caution must be applied when interpreting the effects for process writing, because the comparison condition in professional development studies was unspecified and comparison conditions varied considerably in nonprofessional development studies.

8. Involve adolescents in writing activities designed to sharpen their skills of inquiry. Effective inquiry activities in writing are characterized by a clearly specified goal (e.g., describe the actions of people), analysis of concrete and immediate data (e.g., observe one or more peers during specific activities), use of specific strategies to conduct the analysis (e.g., retrospectively ask the person being observed the reason for their action), and application of what was learned (e.g., write a story where the insights from the inquiry are incorporated into the composition; for inquiry, mean weighted effect size = 0.32; Grades 7–12). Caution must be exercised in interpreting the average weighted effect size for this treatment, because there was considerable variation in type of intervention with which inquiry was compared.

9. Engage adolescents in activities that help them gather and organize ideas for their compositions before they write a first draft. This includes activities such as gathering possible information for a paper through reading or developing a visual representation of their ideas before writing (for prewriting, mean weighted effect size = 0.32; Grades 4–9). Caution must be applied in interpreting the average weighted effect size for this treatment, because there was considerable variation in type of intervention with which prewriting was compared.

10. Provide adolescents with good models for each type of writing that is the focus of instruction. These examples should be analyzed, and students should be encouraged to imitate the critical elements embodied in the models (for models, mean weighted effect size = 0.25; Grades 4–12). Caution must be applied in interpreting the average weighted effect size for this treatment, because there was considerable variation in type of intervention with which prewriting was compared.

At a more general level, these findings show that it is advantageous to explicitly and systematically teach adolescents the processes and strategies involved in writing (including planning, sentence construction, summarizing, and revising). It is also advantageous for teachers to structure writing by having students work together in an organized fashion, establishing clear and reachable goals for writing assignments, providing models of what the end product should look like, and engaging students in activities that help them acquire, evaluate, and organize ideas for their writing. Adolescents are further likely to benefit from moving from composing by hand to composing by means of word processing (this change requires that they become proficient with word processing and related software).

At a more specific level, several of the findings and recommendations presented above merit additional discussion. First, the positive impact of professional development in the process approach to writing provides support for the work of the National Writing Project (Nagin, 2003), because five of the six studies assessed the impact of National Writing Project training. Additional research is needed, however, to verify these findings, because the content of National Writing Project training has changed over time, it was not always clear what teachers learned or subsequently applied in their classrooms in the five National Writing Project studies, random assignment did not occur in any of the five National Writing Project studies, and in some instances the National Writing Project teachers were volunteers. Nevertheless, it is interesting to note that many of the components included in a recent description of the National Writing Project model (e.g., peers working together, inquiry, and sentence combining; see Nagin, 2003) were found to enhance adolescents' writing in this meta-analysis.

Although the finding that the effectiveness of the process writing approach was related to participation in professional development cannot be generalized broadly (because of the methodological issues), we doubt that most of the instructional procedures described earlier, especially the more complex ones like strategy instruction, can be widely and effectively implemented without a considerable amount of teacher preparation. If these practices are to be brought to scale, they must become an integral part of both preservice and in-service teacher education for both language arts and content teachers.

Second, additional research is needed to determine why studies involving the self-regulated strategy development model yielded such large effect sizes. There are many possible reasons why this approach was so effective, including the power of the specific planning or revising strategies taught, the inclusion of instructional procedures to promote self-regulation, or the emphasis on criterion-based rather than time-based instruction. In any event, explicitly teaching students strategies for planning, revising, and editing was effective, whether this or a different model of strategy instruction was used.

Third, we were able to draw only one separate instructional recommendation for students who found writing especially challenging. Strategy instruction had a strong impact on improving the quality of these youngsters' writing. With the exception of research on word processing, however, little of the research on the other treatments focused on struggling writers. Consequently, additional research is needed to examine the effectiveness of these and other writing interventions with these students.

Although it is difficult to compare directly the findings from this meta-analysis and previous ones (because of methodological differences, such as how effect sizes were calculated and range of grade levels included in the reviews), it is worth noting that our findings were generally consistent with the previous ones

(Bangert-Drowns, 1993; Goldring et al., 2003; Graham, 2006a; Graham & Harris, 2003; Hillocks, 1986). Specifically, if an average effect size was small, medium, or large for a treatment in a prior analysis, it tended to remain so in the present review. The only notable exceptions involved inquiry (0.32 in our analysis vs. 0.56 in Hillocks, 1986) and sentence combining (0.50 in our analysis vs. 0.35 in Hillocks, 1986). In the case of inquiry, these differences appeared to be mostly due to the methods used to calculate effect sizes. The studies included in both reviews were identical with one exception (a study including third-grade students). For the investigations that were contained in both reviews, we obtained much smaller effect sizes than Hillocks (1986) for three of the five comparisons (calculated from Hillocks, 1982, and Pisano, 1980). In these investigations, students in the control condition began the study with higher scores than students in the experimental condition. Hillocks used gain scores to calculate effect sizes, correcting for these pretest differences, whereas we just used posttest scores. The procedures used to calculate effect sizes did not appear to contribute to a difference in average weighted effect sizes for sentence combining in this and Hillocks's (1986) article, because our review included only two studies (Howie, 1979; Pedersen, 1977) from Hillocks's review, and the effect sizes were almost identical.

It is also important to note that we were not able to draw any recommendation about the value of increasing the amount of writing done by adolescents. The review included too few effect sizes, too much variability in effect sizes, too much diversity in the procedures used to promote extra writing time, and too many different control/comparison conditions for us to draw any meaningful conclusions about the impact of this treatment. Clearly, additional research is needed to examine the impact of increased writing, because this is one of the most common recommendations for improving students' writing performance and a central feature of the reform platform of the National Commission on Writing (2003). Although this review does not provide any insight into the commission's recommendation to monitor student progress, it does provide support for the use of technology (at least in terms of word processing) and efforts to improve teacher preparation (at least in terms of the process writing approach).

Finally, the studies that our 10 recommendations are based on do not provide guidance on a number of important issues. For example, we do not know what combination of activities or how much of each of the recommended activities is needed to maximize writing instruction for adolescents (although there is some preliminary evidence that integrating some of these specific treatments, such as process writing and strategy instruction, can be beneficial; see Curry, 1997, and Danoff, Harris, & Graham, 1993). In addition, the recommendations are incomplete, because they do not address specific aspects of writing (such as teaching spelling, handwriting, punctuation, and vocabulary) or all forms of instruction (such as conferencing with students about their writing). Further, the recommendations do not provide clear directions for the use of technological tools other than word processing. It is possible that tools, such as speech synthesis and spell checking tools, can enhance the quality of the writing of adolescents, especially for those who struggle with this skill. Also, we can provide no directions on the use of procedures designed specifically to boost motivation (see Bruning & Horn, 2000).

Quality of Research

Although overall study quality, as measured by nine quality indicators, did not predict the magnitude of effect sizes for all studies combined or for treatments with 10 or more effect sizes, there is considerable room for improvement in the writing intervention research reviewed here. Just 33% of the studies involved random assignment of participants to conditions. In addition, procedures for controlling teacher effects were included in just 46% of studies, whereas pretest equivalence was established in only 57% of studies. Although only 46% of the studies described procedures for training teachers/instructors and just 27% measured whether the experimental treatment was implemented as intended, these data need to be interpreted cautiously. It is likely that these last two figures are not accurate. Although teachers/instructors may have received training and experimental treatments may have been implemented with fidelity, the researchers did not report the former or verify the latter. For example, reporting treatment fidelity data, at least in journal articles, has only started to be emphasized in the past 10 years or so. Nevertheless, these findings suggest that that experimental writing intervention research can be improved by increasing how often investigators use random assignment, establish pretest equivalence, control for teacher effects, provide and describe teacher/instructor training, and assess as well as report treatment fidelity.

More positively, mortality of participants as well as ceiling and floor effects for the dependent measure were not an issue in 80% or more of the studies. Likewise, the control condition was clearly specified in 84% of studies. Hawthorne effects were addressed in two out of every three studies. Most important, more recent investigations had higher overall quality scores than earlier ones, suggesting that the quality of experimental writing intervention research is improving. It is not clear why this was the case. It may be, for example, that researchers in this area are now more sophisticated. It is also possible that reporting requirements have changed, and journal editors, reviewers, and dissertation committees now require that information not typically reported in the past, such as treatment fidelity or teacher/instructor training, are included in the written document.

It was interesting to note that there was no statistically significant difference between the quality of studies published in peer-reviewed journals and other outlets, such as books, dissertations, and so forth. This was unexpected, because it is typically assumed that the peer-review process provides an assurance of quality, with only the best studies being published in professional journals. Why did this not happen here? One explanation centers on dissertations. Most of the nonjournal studies were dissertations, and although some dissertations ended up as journal articles (and were included in the journal category), this did not happen very often. Thus, it is possible that dissertation committees were critical enough that students' studies met the quality standards used by journal reviewers. In contrast, it is also possible that the peer-review process is not stringent enough in this area of research.

The observation that most dissertations were not published as journal articles, coupled with the fact that just 10 researchers (i.e., Couzijn, De La Paz, Fitzgerald, Graham, Hillocks, Knudson, MacArthur, Pritchard, Scardamalia, Schunk) had more than one publication in this review, raises serious concerns about the future of experimental writing intervention research with adolescents.

Most doctoral students who initially conduct research in this area do not continue on this path, and there is little in the way of research extending beyond an individual study or two (only 2 researchers, Graham and Scardamalia, had five or more publications, either articles or book chapters). Although it is encouraging that there have been many new studies since Hillocks's (1986) seminal review, the growth and impact of this area of research is dependent on the development of new researchers who engage in an extended program of research. Presently, neither federal (e.g., Institute of Education Sciences in the U.S. Department of Education or National Institute of Mental Health) nor private agencies devote much money to funding research in this area or the preparation of new researchers. Such investment is critical to make continued inroads into solving the writing difficulties exhibited by so many adolescents in this country.

Issues Involved in Implementing the Recommendations

Implementing research-based treatments is a challenging and complex task. Just because an intervention was effective in the studies included in this review does not guarantee that it will be effective in all other situations. For example, there is rarely, if ever, an exact match between the conditions in which the research was implemented and the conditions in which it is subsequently implemented by teachers. If the research was conducted outside a school setting, the distance between the research and implementation conditions is typically large. Even when research is conducted in schools with teachers delivering the treatment (which occurred in the majority of studies in this meta-analysis), many differences in conditions still exist. Although teachers are more likely to obtain results similar to those obtained by researchers when their classroom conditions are similar to the research conditions, the safest course is to continually monitor the effects of the treatment to gauge directly whether it is effective under these new conditions (Graham & Harris, 2005).

Another important issue in implementing evidence-based writing practices and recommendations revolves around the different organizational structures or formats for writing that exist in secondary schools. Writing instruction could occur within the context of the language arts or English classroom or even with a learning specialist, such as a special education teacher. In this situation, instruction might focus almost exclusively on developing students' writing skills. There are other options, however, that combine writing instruction with content instruction (Shanahan, 2004). With an *applied academics format,* the language arts or English teacher, as well as a learning specialist, could use subject matter, such as science or social studies, as the content of writing instruction. For example, strategies for writing persuasive essays might be taught with text read in a concurrent social studies class. In an *infused content format,* a content-area teacher could teach writing skills in the course of teaching subject matter, as encouraged by content-area literacy educators (e.g., Graham & Perin, 2007). In the *learning community format* (Perin, 2001), a content-area teacher and the English or language arts instructors could align their curricula, giving students assignments that systematically connect writing and content instruction. For instance, in a study by De La Paz (2005), a history teacher taught a historical reasoning strategy with historical documents, and the English teacher taught strategies for writing argumentation essays with the documents applied by the history teacher.

The effectiveness of these various formats has been neither tested nor compared one to another. It is also not certain how well the evidence-based practices identified in this review would fare in these different formats. Before implementing one or more of these procedures or their accompanying recommendations, careful analysis of the organizational structure or format within which they will be placed should be undertaken, with the aim of identifying factors that will facilitate and impede their effectiveness.

Concluding Comments

As Bangert-Drowns et al. (2004) noted, meta-analysis provides a useful tool for drawing "important insight from what might otherwise be a confused and disparate literature" (p. 52). The writing intervention literature certainly fits this description, because it investigates the effectiveness of a wide range of interventions. Like others before us (Hillocks, 1986), we capitalized on the strengths of meta-analysis in an attempt to identify effective writing treatments for adolescents in Grades 4–12. This was a productive strategy, because we identified a variety of effective treatments, ranging from explicitly teaching writing strategies to studying models of good writing.

References

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

Achieve, Inc. (2005). *Rising to the challenge: Are high school graduates prepared for college and work?* Washington, DC: Author.

*Adams, V. A. (1971). *A study of the effects of two methods of teaching composition to twelfth graders.* Unpublished doctoral dissertation, University of Illinois, Champaign-Urbana, IL.

Agate, L. (2005). *Investigation of primary grade writing instruction.* Unpublished doctoral dissertation, University of Maryland, College Park, MD.

*Alloway, E., Carroll, J., Emig, J., King, B., Marcotrigiano, I., Smith, J., & Spicer, W. (1979). *The New Jersey Writing Project.* New Brunswick, NJ: Rutgers University, Educational Testing Program Service, and Nineteen New Jersey Public School Districts.

*Anderson, A. A. (1997). *The effects of sociocognitive writing strategy instruction on the writing achievement and writing self-efficacy of students with disabilities and typical achievement in an urban elementary school.* Unpublished doctoral dissertation, University of Houston, Houston, TX.

Anderson, R., & Pearson, D. (1984). A schema-theoretic view of basic process in reading. In D. Pearson (Ed.), *Handbook of reading research* (pp. 378–455). New York: Longman.

*Andrade, H. G., & Boulay, B. A. (2003). Role of rubric-referenced self-assessment in learning to write. *Journal of Educational Research, 97,* 21–34.

Anthony, H., & Anderson, L. (1987). *The nature of writing instruction in regular and special education classrooms.* Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.

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

Applebee, A. (1986). Problems in process approaches: Towards reconceptualization of process instruction. In A. Petrosky & D. Bartholomae

(Eds.), *The teaching of writing.* Chicago, IL: University of Chicago Press.

Applebee, A. (2000). Alternative models of writing development. In R. Indrisano & J. Squire (Eds.), *Perspectives on writing: Research, theory, and practice* (pp. 90–110). Newark, DE: International Reading Association.

Applebee, A., Langer, J., Nystrand, M., & Gamoran, A. (2003). Discussion-based approaches to developing understanding: Classroom instruction and student performance in middle and high school English. *American Educational Research Journal, 40,* 685–730.

Bangert-Drowns, R. (1993). The word processor as an instructional tool: A meta-analysis of word processing in writing instruction. *Review of Educational Research, 63,* 69–93.

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,* 29–58.

*Bean, T. W., & Steenwyk, F. L. (1984). The effect of three forms of summarization instruction on sixth graders' summary writing and comprehension. *Journal of Reading Behavior, 16*(4), 297–306.

*Benson, N. L. (1979). *The effects of peer feedback during the writing process on writing performance, revision behavior, and attitude toward writing.* Unpublished doctoral dissertation, University of Colorado, Boulder, CO.

*Boscolo, P., & Ascorti, K. (2004). Effects of collaborative revision on children's ability to write understandable narrative texts. In L. Allal, L. Chanquoy, & P. Largy (Eds.), *Revision: Cognitive and instructional processes* (pp. 157–170). Boston: Kluwer Academic Publishers.

*Braaksma, M. A. H. (2002). *Observational learning in argumentative writing.* Amsterdam: University of Amsterdam.

*Brodney, B., Reeves, C., & Kazelskis, R. (1999). Selected prewriting treatments: Effects on expository compositions written by fifth-grade students. *Journal of Experimental Education, 68,* 5–20.

Bruning, R., & Horn, C. (2000). Developing motivation to write. *Educational Psychologist, 35,* 25–37.

*Bryson, M., & Scardamalia, M. (1996). Fostering reflectivity in the argumentative thinking of students with different learning histories. *Reading and Writing Quarterly: Overcoming Learning Difficulties, 12,* 351–384.

*Bui, Y. N. (2002). *The Demand Writing Instructional Model: Impacting the writing performance of students with learning disabilities and low-performing students from culturally and linguistically diverse backgrounds on the statewide writing assessment.* Unpublished doctoral dissertation, University of Kansas, Kansas City, KS.

Calkins, L. (1981). Case study of a nine year old writer. In D. Graves (Ed.), *A case study observing development of primary-children's composing, spelling, and motor behavior during the writing process* (pp. 239–262). Durham: University of New Hampshire.

*Caplan, R., & Keech, C. (1980). *Show-writing: A training program to help students be specific. Collaborative research study no. 2.* Berkeley, CA: Bay Area Writing Project. (ERIC Document Reproduction Service No. ED198539)

Chafe, W., & Tannen, D. (1987). The relation between written and spoken language. *Annual Review of Anthropology, 16,* 383–407.

*Chang, K. E., Sung, Y. T., & Chen, I. D. (2002). The effect of concept mapping to enhance text comprehension and summarization. *Journal of Experimental Education, 71,* 5–23.

*Cheever, M. S. (1987). *The effects of using a word processor on the acquisition of composition skills by the elementary student.* Unpublished doctoral dissertation, Northwestern University, Evanston, IL.

*Cirello, V. J. (1986). *The effect of word processing on the writing abilities of tenth grade remedial writing students.* Unpublished doctoral dissertation, New York University, New York City, NY.

*Collopy, R. M. B., & Bowman, P. (2005, April). *The impact of a Six Traits Analytic Writing Model on student achievement.* Paper presented at the annual meeting of the American Educational Research Association, Montreal, Quebec, Canada.

*Couzijn, M. (1999). Learning to write by observation of writing and reading processes: Effects on learning and transfer. *Learning and Instruction, 9,* 109–142.

*Couzijn, M., & Rijlaarsdam, G. (2005). Learning to write instructive texts by reader observation and written feedback. In G. Rijlaarsdam, H. van den Bergh, & M. Couzijn (Eds.), *Effective learning and teaching of writing: A handbook of writing in education* (2nd ed., pp. 209–240). New York: Kluwer Academic.

*Covill, A. E. (1996). *Students' revision practices and attitudes in response to surface-related feedback as compared to content-related feedback on their writing.* Unpublished doctoral dissertation, University of Washington, Seattle, WA.

*Curry, K. A. (1997). *A comparison of the writing products of students with learning disabilities in inclusive and resource room settings using different writing instruction approaches.* Unpublished doctoral dissertation, Florida Atlantic University, Boca Raton, FL.

*Dailey, E. M. (1991). *The relative efficacy of cooperative learning versus individualized learning on the written performance of adolescent students with writing problems.* Unpublished doctoral dissertation, Johns Hopkins University, Baltimore, MD.

*Dalton, D. W., & Hannafin, M. J. (1987). The effects of word processing on written composition. *Journal of Educational Research, 80,* 338–342.

Danoff, B., Harris, K. R., & Graham, S. (1993). Incorporating strategy instruction within the writing process in the regular classroom: Effects on the writing of students with and without learning disabilities. *Journal of Reading Behavior, 25,* 295–322.

De La Paz, S. (1999). Self-regulated strategy instruction in regular education settings: Improving outcomes for students with and without learning disabilities. *Learning Disabilities Research and Practice, 14,* 92–106.

*De La Paz, S. (2005). Teaching historical reasoning and argumentative writing in culturally and academically diverse middle school classrooms. *Journal of Educational Psychology, 97,* 139–158.

*De La Paz, S., & Graham, S. (1997). Effects of dictation and advanced planning instruction on the composing of students with writing and learning problems. *Journal of Educational Psychology, 89,* 203–222.

*De La Paz, S., & Graham, S. (2002). Explicitly teaching strategies, skills, and knowledge: Writing instruction in middle school classrooms. *Journal of Educational Psychology, 94,* 291–304.

Diamond, J. (1999). *Guns, germs, and steel: The fates of human societies.* New York: Norton.

Diederich, P. (1966). How to measure growth in writing ability. *English Journal, 55,* 435–449.

*Duin, A. H., & Graves, M. F. (1987). Intensive vocabulary instruction as a prewriting technique. *Reading Research Quarterly, 22,* 311–330.

*Duke, B. L. (2003). *The influence of using cognitive strategy instruction through writing rubrics on high school students' writing self-efficacy, achievement goal orientation, perceptions of classroom goal structures, self-regulation, and writing achievement.* Unpublished doctoral dissertation, University of Oklahoma, Norman, OK.

*Dybdahl, C. S., Shaw, D. G., & Blahous, E. (1997). The impact of the computer on writing: No simple answers. *Computers in the Schools, 13*(3/4), 41–53.

*Eagleton, C. J. (1973). *Reciprocal effects of eleventh-graders and twelfth-graders as tutors to sixth-graders in reading, written expression, and attitude modification.* Unpublished doctoral dissertation, American University, Washington, D. C.

*Elley, W. B., Barham, I. H., Lamb, H., & Wyllie, M. (1975). The role of grammar in a secondary school English curriculum. *New Zealand Journal of Educational Studies, 10,* 26–42.

*Espinoza, S. F. (1992). *The effects of using a word processor containing grammar and spell checkers on the composition writing of sixth graders.*

Unpublished doctoral dissertation, Texas Tech University, Lubbock, TX.

*Farrell, K. J. (1977). *A comparison of three instructional approaches for teaching written composition to high school juniors: Teacher lecture, peer evaluation, and group tutoring.* Unpublished doctoral dissertation, Boston University School of Education, Boston, MA.

*Fearn, L., & Farnan, N. (2005, April). *An investigation of the influence of teaching grammar in writing to accomplish an influence on writing.* Paper presented at the annual meeting of the American Educational Research Association, Montreal, Quebec, Canada.

*Ferretti, R. P., MacArthur, C. A., & Dowdy, N. S. (2000). The effects of an elaborated goal on the persuasive writing of students with learning disabilities and their normally achieving peers. *Journal of Educational Psychology, 92,* 694–702.

*Fitzgerald, J., & Markham, L. (1987). Teaching children about revision in writing. *Cognition and Instruction, 4,* 3–24.

*Fitzgerald, J., & Teasley, A. B. (1986). Effects of instruction in narrative structure on children's writing. *Journal of Educational Psychology, 78,* 424–432.

Fox, D., & Suhor, C. (1986). Limitations of free writing. *English Journal, 75,* 34–36.

*Gamelin, Y. M. A. (1996). *The effects of Cognitive Strategy Instruction in Writing (CSIW) on the writing skills of severely learning disabled students and their peers in an inclusive classroom.* Unpublished master's thesis, Simon Fraser University, Burnaby, British Columbia, Canada.

*Ganong, F. L. (1974). *Teaching writing through the use of a program based on the work of Donald M. Murray.* Unpublished doctoral dissertation, Boston University School of Education, Boston, MA.

*Gauntlett, J. F. (1978). *Project WRITE and its effect on the writing of high school students.* Unpublished doctoral dissertation, Northern Arizona University, Flagstaff, AZ.

Gersten, R., Fuchs, L., Compton, D., Coyne, M., Greenwood, C., & Innocenti, M. (2005). Quality indicators for group experimental and quasi-experimental research in special education. *Exceptional Children, 71,* 149–164.

*Glaser, C. (2005). *Improving the fourth-grade students' composition skills: Effects of strategy instruction and self-regulatory procedures.* Unpublished doctoral dissertation, University of Pottsburg, Germany.

Golden, P. (1986). Writing process reconsidered. *Boston Writing Project Newsletter.*

Goldring, A., Russell, M., & Cook, A. (2003). The effects of computers on student writing: A meta-analysis of studies from 1992–2002. *Journal of Technology, Learning, and Assessment, 2,* 1–51.

*Gomez, R., Jr., Parker, R., Lara-Alecio, R., & Gomez, L. (1996). Process versus product writing with limited English proficient students. *Bilingual Research Journal, 20,* 209–233.

*Gordon, C. J., & Braun, C. (1986). Mental processes in reading and writing: A critical look at self-reports as supportive data. *Journal of Educational Research, 79*(5), 292–301.

*Gorka, D. M. (1992). *The effects of a staff development program in writing process on learners' writing skills and attitudes toward writing.* Unpublished doctoral dissertation, Pennsylvania State University, University Park, PA.

Graham, S. (1990). The role of production factors in Learning Disabled students' compositions. *Journal of Educational Psychology, 82,* 781–791.

Graham, S. (1999). Handwriting and spelling instruction for students with learning disabilities. *Learning Disabilities Quarterly, 22,* 78–98.

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

Graham, S. (2006a). Strategy instruction and the teaching of writing: A meta-analysis. In C. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (pp. 187–207). New York: Guilford.

Graham, S. (2006b). Writing. In P. Alexander & P. Wiine (Eds.), *Handbook of educational psychology* (pp. 457–477). Mahwah, NJ: Erlbaum.

Graham, S., & Harris, K. R. (2003). Students with learning disabilities and the process of writing: A meta-analysis of SRSD studies. In L. Swanson, K. R. Harris, & S. Graham (Eds.), *Handbook of research on learning disabilities* (pp. 383–402). New York: Guilford.

Graham, S., & Harris, K. (2005). *Writing better.* Baltimore, MD: Brookes.

Graham, S., Harris, K. R., & Fink, B. (2000). Is handwriting causally related to learning to write? Treatment of handwriting problems in beginning writers. *Journal of Educational Psychology, 92,* 620–633.

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

*Graham, S., MacArthur, C., & Schwartz, S. (1995). Effects of goal setting and procedural facilitation on the revising behavior and writing performance of students with writing and learning problems. *Journal of Educational Psychology, 87,* 230–240.

Graham, S., & Perin, D. (2007). *Writing Next: Effective strategies to improve writing of adolescents in middle and high schools.* New York: Carnegie Corporation of New York.

Graves, D. (1979). What children show us about revision. *Language Arts, 56,* 312–319.

Greenwald, E., Persky, H., Ambell, J., & Mazzeo, J. (1999). *National assessment of educational progress: 1998 report card for the nation and the states.* Washington, D. C.: US Department of Education.

Gregg, L., & Steinberg, E. (Eds.) (1980). *Cognitive processes in writing.* Hillsdale, NJ: Erlbaum.

*Guastello, E. F. (2001). Parents as partners: Improving children's writing. In W. M. Linek, E. G. Sturtevant, J. A. R. Dugan, & P. E. Linder (Eds.), *Celebrating the voices of literacy: Yearbook of the College Reading Association* (pp. 279–295). Readyville, TN: College Reading Association.

*Hagler, W. J. (1993). *The effects of the word processor on the revision behaviors of sixth grade students.* Unpublished doctoral dissertation, Auburn University, Auburn, AL.

Harris, K., & Graham, S. (1996). *Making the writing process work: Strategies for composition and self-regulation* (2nd ed.). Cambridge: Brookline Books.

Harris, K. R., & Graham, S. (1999). Programmatic intervention research: Illustrations from the evolution of self-regulated strategy development. *Learning Disability Quarterly, 22,* 251–262.

*Hayes, B. L. (1984). *The effects of implementing process writing into a seventh grade English curriculum.* Unpublished doctoral dissertation, Delta State University, Cleveland, MS.

Hayes, J. R. (2000). A new framework for understanding cognition and affect in writing. In R. Indrisano & J. R. Squire (Eds.), *Perspectives on writing: Research, theory and practice* (pp 6–44). Newark, DE: International Reading Association. Originally published in C. M. Levy & S. Ransdell (Eds.) (1996). *The science of writing: Theories, methods, individual differences, and applications* (pp. 1–27). Mahwah, NJ: Erlbaum.

*Head, B. B. (2000). *Revision instruction and quality of writing by eighth grade students using paper and pencil or word processing.* Unpublished doctoral dissertation, Oakland University, Rochester, MI.

Hedges, L., & Olkin, I. (1985). *Statistical methods for meta-analysis.* Orlando, FL: Academic Press.

*Hill, B. G. (1990). *A comparison of the writing quality of paired and unpaired students composing at the computer.* Unpublished doctoral dissertation, University of Texas, Austin, TX.

Hillocks, G. (1986). *Research on written composition: New directions for teaching.* Urbana, IL: National Council of Teachers of English.

*Hillocks, G., Jr. (1979). The effects of observational activities on student writing. *Research in the Teaching of English, 13,* 23–35.

*Hillocks, G., Jr. (1982). The interaction of instruction, teacher comment,

and revision in teaching the composing process. *Research in the Teaching of English, 16,* 261–278.

*Hook, M. E. (1986). *An inquiry into the influence of expressive writing practice on the transactional writing abilities of selected fourth grade students.* Unpublished doctoral dissertation, Old Dominion University, Norfolk, VA.

*Howie, S. M. H. (1979). *A study: The effects of sentence combining practice on the writing ability and reading level of ninth grade students.* Unpublished doctoral dissertation, University of Colorado, Boulder, CO.

Huber, P. (1977). *Robust statistical procedures.* Philadelphia: Society for Industrial and Applied Mathematics.

*Jackiewicz, G. (1995). *The effect of computer based instruction on writing at the elementary level.* (ERIC Document Reproduction Service No. ED380802)

*Kanellas, R., Carifio, J., & Dagostino, L. (1998). *Improving the expository writing skills of adolescents.* Oxford, England: Oxford University Press.

Kerchner, L. B., & Kistinger, B. J. (1984). Language processing/ word processing: Written expression, computers and learning disabled students. *Learning Disability Quarterly, 7,* 329–335.

Keys, C. W. (2000). Investigating the thinking processes of eighth grade writers during the composition of a scientific laboratory report. *Journal of Research in Science Teaching, 37,* 676–690.

*Knight, J. E. (2003). *Helping able readers synthesize information from multiple sources in response to open-ended questions.* Unpublished doctoral dissertation, University of Maryland, College Park, MD.

*Knudson, R. E. (1989). Effects of instructional strategies on children's informational writing. *Journal of Educational Research, 83*(2), 91–96.

*Knudson, R. E. (1991). Effects of instructional strategies, grade, and sex on students' persuasive writing. *Journal of Experimental Education, 59,* 141–152.

*Lam, F. S., & Pennington, M. C. (1995). The computer vs. the pen: A comparative study of word processing in a Hong Kong secondary classroom. *Computer Assisted Language Learning, 8,* 75–92.

*Lerew, E. L. (1997). *The use of computers to improve writing skills among low-achieving Hispanic students.* Unpublished doctoral dissertation, University of La Verne, La Verne, CA.

Levy, M., & Ransdell, S. (1996). *The science of writing: Theories, methods, individual differences, and applications.* Mahwah, NJ: Erlbaum.

*Lichtenstein, N. (1996). *The effect of word processing on writing achievement.* (ERIC Document Reproduction Service No. ED394146)

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

*Loader, L. M. (1989). *The effects of the semantic organizer on writing ability and motivation with fourth grade students.* Unpublished doctoral dissertation, University of Connecticut, Storrs, CT.

*Lowther, D. L., Ross, S. M., & Morrison, G. M. (2003). When each one has one: The influences on teaching strategies and student achievement of using laptops in the classroom. *Educational Technology, Research and Development, 51*(3), 23–44.

*Lumbelli, L., Paoletti, G., & Frausin, T. (1999). Improving the ability to detect comprehension problems: From revising to writing. *Learning and Instruction, 9,* 143–166.

*Lytle, M. J. (1987). *Word processors and writing: The relation of seventh grade students' learner characteristics and revision behaviors.* Unpublished doctoral dissertation, University of Oregon, Eugene, OR.

MacArthur, C., Graham, S., & Fitzgerald, J. (2006). *Handbook of writing research.* New York: Guilford.

*MacArthur, C., Schwartz, S., & Graham, S. (1991). Effects of a reciprocal peer revision strategy in special education classrooms. *Learning Disability Research and Practice, 6,* 201–210.

*Miller, S. K. (1984). *Plugging your pencil into the wall: An investigation of word processing and writing skills at the middle school level.* Unpublished doctoral dissertation, University of Oregon, Eugene, OR.

*Moore, M. A. (1987). *The effect of word processing technology in a developmental writing program on writing quality, attitude towards composing, and revision strategies of fourth and fifth grade students.* Unpublished doctoral dissertation, University of South Florida, Tampa, FL.

*Moye, M. J. (1993). *The impact of a cognitive strategy on students' composing skills.* Unpublished doctoral dissertation, College of William and Mary, Williamsburg, VA.

Nagin, C. (2003). *Because writing matters: Improving student writing in our schools.* San Francisco, CA: Jossey-Bass.

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

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

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

*Olson, M. C., & DiStefano, P. (1980). Describing and testing the effectiveness of a contemporary model for in-service education in teaching composition. *Engineering Education, 12,* 69–76.

*Olson, V. B. (1990). The revising processes of sixth-grade writers with and without peer feedback. *Journal of Educational Research, 84,* 22–29.

*Page-Voth, V., & Graham, S. (1999). Effects of goal setting and strategy use on the writing performance and self-efficacy of students with writing and learning problems. *Journal of Educational Psychology, 91,* 230–240.

*Pantier, T. F. (1999). *A comparison of writing performance of fifth grade students using the process writing approach and the Shurley Method.* Unpublished doctoral dissertation, Oklahoma State University, Stillwater, OK.

*Pedersen, E. L. (1977). *Improving syntactic and semantic fluency in writing of language arts students through extended practice in sentence-combining.* Unpublished doctoral dissertation, University of Minnesota, Minneapolis, MN.

Perin, D. (2001). Academic-occupational integration as a reform strategy for the community college: Classroom perspectives. *Teachers College Record, 103,* 303–335.

Perin, D. (in press). Can community colleges protect both access and standards? The problem of remediation. *Teachers College Record.*

Persky, H. R., Daane, M. C., & Jin, Y. (2003). *The nation's report card: Writing 2002.* (NCES 2003–529). U.S. Department of Education. Institute of Education Sciences. National Center for Education Statistics. Washington, DC: Government Printing Office.

*Philhower, S. C. (1985). *The effects of the use of a word processing program on the writing skills of mildly handicapped, secondary students.* Unpublished doctoral dissertation, University of Iowa, Iowa City, IA.

*Pisano, R. C. (1980). *The effectiveness of an intervention study in critical thinking skills designed to improve written composition in eleventh and twelfth graders.* Unpublished doctoral dissertation, Rutgers: The State University of New Jersey, New Brunswick, NJ.

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

*Prater, D. L., & Bermúdez, A. B. (1993). Using peer response groups with limited English proficient writers. *Bilingual Research Journal, 17,* 99–116.

Pressley, M., Graham, S., & Harris, K. R. (2006). The state of educational intervention research. *British Journal of Educational Psychology, 76,* 1–19.

Pressley, M., Yokoi, L., Rankin, J., Wharton-McDonald, R., & Mistretta, J. (1997). A survey of the instructional practices of grade 5 teachers nominated as effective in promoting literacy. *Scientific Studies of Reading, 1,* 1–16.

Prior, P. (2006). A sociocultural theory of writing. In C. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (pp. 54–66). New York: Guilford.

*Pritchard, R. J. (1987). Effects on student writing of teacher training in the National Writing Project Model. *Written Communication, 4,* 51–67.

Pritchard, R., J., & Honeycutt, J. (2006). Process writing. In C. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of Writing Research* (pp. 275–290). New York: Guilford.

*Pritchard, R. J., & Marshall, J. C. (1994). Evaluation of a tiered model for staff development in writing. *Research in the Teaching of English, 28,* 259–285.

*Reece, J. E., & Cumming, G. (1996). Evaluating speech-based composition methods: Planning, dictation, and the listening word processor. In C. M. Levy & S. Ransdell (Eds.), *The science of writing: Theories, methods, individual differences, and applications (pp. 361*–380). Mahwah, New Jersey: Lawrence Erlbaum Associates.

*Reedy, J. E., Jr. (1964). *A comparative study of two methods of teaching the organization of expository writing to ninth-grade pupils.* Unpublished doctoral dissertation, Boston University School of Education, Boston, MA.

*Reimer, M. (2001). *The effect of a traditional, a process writing, and a combined talking and writing instructional approach on the quality of secondary English students' written response.* Unpublished master's thesis, University of Manitoba, Canada.

*Reynolds, C., Hill, D., Swassing, R., & Ward, M. (1988). The effects of revision strategy instruction on the writing performance of students with learning disabilities. *Journal of Learning Disabilities, 21,* 540–545.

*Roberts, C. (2002). *The influence of teachers' professional development at the Tampa Bay Area Writing Project on student writing performance.* Unpublished doctoral dissertation: University of South Florida: Tampa, FL.

*Roberts, R. J. (1999). *Use of computer dictation by students with learning disabilities.* Unpublished doctoral dissertation, Syracuse University, Syracuse, NY.

*Robinson, M. E. (1986). *The writing performance and revision behavior of fifth grade process and non-process writing students during one-day and two-day writing sessions.* Unpublished doctoral dissertation, University of Maine, Orono, ME.

Rosenbaum, N. J. (1987). *A study to determine the value of computers in the revision process of written communication.* Unpublished doctoral dissertation, Temple University, Philadelphia, PA.

Rosenberg, M., Adams, D., & Gurevitch, J. (2000). *MetaWin: Statistical soft-ware for meta-analysis* (Version 2). Sunderland, MA: Sinauer Ass.

*Saddler, B., & Graham, S. (2005). The effects of peer-assisted sentence-combining instruction on the writing performance of more and less skilled young writers. *Journal of Educational Psychology, 97,* 43–54.

*Sawyer, R., Graham, S., & Harris, K. R. (1992). Direct teaching, strategy instruction, and strategy instruction with explicit self-regulation: Effects on the composition skills and self-efficacy of students with learning disabilities. *Journal of Educational Psychology, 84,* 340–352.

*Scannella, A. M. (1982). *A writing-as-process model as a means for improving compositions and attitudes toward composition in the high school.* Unpublished doctoral dissertation, Rutgers: The State University of New Jersey, New Brunswick, NJ.

*Scardamalia, M., & Bereiter, C. (1985). Development of dialectical processes in composition. In D. R. Olson, N. Torrance, & A. Hildyard (Eds.), *Literacy, language, and learning: The nature and consequences of reading and writing* (pp. 307–329). Cambridge: Cambridge University Press.

*Scardamalia, M., Bereiter, C., & Steinbach, R. (1984). Teachability of reflective processes in written composition. *Cognitive Science, 8,* 173–190.

*Scardamalia, M., & Paris, P. (1985). The function of explicit discourse knowledge in the development of text representations and composing strategies. *Cognition and Instruction, 2,* 1–39.

*Schunk, D. H., & Swartz, C. W. (1993a). Goals and progress feedback: Effects on self-efficacy and writing achievement. *Contemporary Educational Psychology, 18,* 337–354.

*Schunk, D. H., & Swartz, C. W. (1993b). Writing strategy instruction with gifted students: Effects of goals and feedback on self-efficacy and skills. *Roeper Review, 15*(4), 225–230.

*Sengupta, S. (2000). An investigation into the effects of revision strategy instruction on L2 secondary school learners. *System, 28,* 97–113.

Shanahan, T. (2004). Overcoming the dominance of communication: Writing to think and to learn. In T. L. Jetton & J. A. Dole (Eds.). *Adolescent literacy research and practice* (pp. 59–73). New York: Guilford.

*Shinn, J. A. (1986). *The effectiveness of word processing and problem solving computer use on the skills of learning disabled students.* Unpublished doctoral dissertation, United States International University, San Diego, CA.

*Silver, N. W., & Repa, J. T. (1993). The effect of word processing on the quality of writing and self-esteem of secondary school English-as-Second-Language students: Writing without censure. *Journal of Educational Computing Research, 9,* 265–283.

*Simmons, D. C., Kameenui, E. J., Dickson, S., Chard, D., Gunn, B., & Baker, S. (1994). Integrating narrative reading comprehension and writing instruction for all learners. In D. J. Leu, & C. K. Kinzer (Eds.), *Multidimensional aspects of literacy research, theory, and practice: 43rd yearbook of The National Reading Conference* (pp. 572–582). Chicago: National Reading Conference.

Smyth, J. (1998). Written emotional expression: Effect sizes, outcome types, and moderating variables. *Journal of Consulting and Clinical Psychology, 66,* 174–184.

Sperling, M., & Freedman, S. W. (2002). Research on writing. In V. Richardson (Ed.), *Handbook of research on teaching* (4th ed., pp. 370–389). Washington, DC: American Educational Research Association.

*Stoddard, E. P. (1982). *The combined effect of creative thinking and sentence-combining activities on the writing ability of above average ability fifth and sixth grade students.* Unpublished doctoral dissertation, University of Connecticut, Storrs, CT.

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

Swedlow, J. (1999). The power of writing. *National Geographic, 196,* 110–132.

*Thibodeau, A. E. (1964). *Improving composition writing with grammar and organization exercises utilizing differentiated group patterns.* Unpublished doctoral dissertation, Boston University School of Education, Boston, MA.

*Thibodeau, A. L. (1964). *A study of the effects of elaborative thinking and vocabulary enrichment exercises on written composition.* Unpublished doctoral dissertation, Boston University School of Education, Boston, MA.

*Tienken, C. H., & Achilles, C. M. (2003). Changing teacher behavior and improving student writing achievement. *Planning and Changing, 34*(3/4), 153–168.

*Troia, G., & Graham, S. (2002). The effectiveness of a highly explicit, teacher-directed strategy instruction routine: Changing the writing performance of students with learning disabilities. *Journal of Learning Disabilities, 35,* 290–305.

*Umbach, B. T. (1990). *A comparison of two methods of teaching written language to low-performing fourth graders in two rural schools.* Unpublished doctoral dissertation, Auburn University, Auburn, AL.

*Utay, C., & Utay, J. (1997). Peer-assisted learning: The effects of cooperative learning and cross-age peer tutoring with word processing on

writing skills of students with learning disabilities. *Journal of Computing in Childhood Education, 8,* 165–185.

*Varble, M. E. (1990). Analysis of writing samples of students taught by teachers using whole language and traditional approaches. *Journal of Educational Research, 83*(5), 245–251.

*Vinson, L. L. N. (1980). *The effects of two prewriting activities upon the overall quality of ninth graders' descriptive paragraphs.* Unpublished doctoral dissertation, University of South Carolina, Columbia, SC.

*Walser, T. M. (2000). *Teaching self-management of writing strategies to promote transfer.* Unpublished doctoral dissertation, Utah State University, Logan, UT.

*Welch, M. (1992). The PLEASE strategy: A metacognitive learning strategy for improving the paragraph writing of students with mild disabilities. *Learning Disability Quarterly, 15,* 119–128.

*Welch, M., & Jensen, J. (1990). Write, PLEASE: A video-assisted strategic intervention to improve written expression of inefficient learners. *Remedial and Special Education, 12,* 37–47.

*Widvey, L. I. H. (1971). *A study of the use of a problem-solving approach to composition in high school English.* Unpublished doctoral dissertation, University of Nebraska, Lincoln, NE.

*Wienke, J. W. (1981). *Strategies for improving elementary school students' writing skills.* (ERIC Document Reproduction Service No. ED209679)

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

*Wong, B. Y. L., Butler, D. L., Ficzere, S. A., & Kuperis, S. (1996). Teaching low achievers and students with learning disabilities to plan, write, and revise opinion essays. *Journal of Learning Disabilities, 29,* 197–212.

*Yarrow, F., & Topping, K. J. (2001). Collaborative writing: The effects of metacognitive prompting and structured peer interaction. *British Journal of Educational Psychology, 71,* 261–282.

*Yeh, S. (1998). Empowering education: Teaching argumentative writing to cultural minority middle-school students. *Research in the Teaching of English, 33,* 49–83.

*Zellermayer, M., Salomon, G., Globerson, T., & Givon, H. (1991). Enhancing writing-related metacognitions through a computerized writing partner. *American Educational Research Journal, 28,* 373–391.

Zimmerman, B., & Risemberg, R. (1997). Become a self-regulated writer: A social cognitive perspective. *Contemporary Educational Psychology, 22,* 73–101.

Appendix

Writing Instruction Treatments That Included Three or Fewer Effect Sizes

Study	Grade	Participant type	*n*	Genre	Treatment	Publication type	Effect size
Individualized							
Utay and Utay (1997)	4–6	LD	47	N	older students with learning disabilities tutored younger students with learning disabilities versus regular writing instruction in computer lab	J	−0.40
Eagleton (1973)	6	full range	87	N	tutors used self-initiating activities to enhance written communication versus no-treatment control	D	−0.02
Farrell (1977)	11	full range	84	E	tutor provided instruction to a small group of students to help them with their composition versus traditional composition instruction	D	0.08
Environmental							
Tienken and Achilles (2003)	4	full range	98	N	training in implementing an environmental mode of writing instruction versus unspecified control	J	0.41
Bui (2002)	5	full range	113	?	teachers implemented a model with specific learning objectives for acquiring genre and process knowledge and students worked together around specific writing tasks to acquire these skills versus unspecified control	D	0.26
Presentational							
Caplan and Keech (1980)	12	full range	129	N	teacher-led discussion, assignments that involved using models that clarified differences between showing and telling, and feedback from the teacher versus unspecified control condition	R	0.12
Dictation							
De La Paz and Graham (1997)	5–7	LD	22	E	students dictated plans and text versus writing plans and text	J	0.35
Process goals							
Schunk and Swartz (1993b), Exp. 2	4	high	22	N, E	process goals to learn strategy plus feedback versus goals to write a specific type of paragraph	J	1.69
Schunk and Swartz (1993a), Exp. 2	5	full range	30	N, E	process goals to learn strategy plus feedback versus goals to write specific type of paragraph	J	2.14
Schunk and Swartz (1993a), Exp. 2	5	full range	20	N, E	process goals to learn strategy plus feedback versus goals to write a specific type of paragraph	J	1.12
Process approach and strategy instruction							
Curry (1997)	4	LD	44	N	strategy instruction combined with the process writing approach versus the process writing approach	D	0.69
Grammar and spell checkers							
Espinoza (1992)	6	full range	61	E	grammar and spell checkers added to a word-processing program versus word processing	D	0.10
Parental training							
Guastello (2001)	4	full range	167	N	parents taught about process writing approach and scoring rubric used in their child's class versus no-treatment control	J	2.19
Observations							
Braaksma (2002)	8	full range	214	E	students observed and analyzed how a stronger and a weaker writer composed text versus writing text	D	−0.06
Couzijn (1999)	9	full range	60	E	students observed the processes a peer used to write text correctly versus writing text	B	0.65
Duke (2003)	10–12	full range	164	E	planning and revising strategies were modeled versus feedback on written products	D	0.61
Planning and dictation							
Reece and Cumming (1996), Study 4	5–6	full range	20	?	students prompted to plan when dictating following a brief demonstration on how to do so versus no-treatment condition	B	0.58
Reece and Cumming (1996), Study 5	5–6	full range	20	?	students prompted to plan when composing by means of speech synthesis following brief instruction in how to do so versus no-treatment condition	B	−0.19
De La Paz and Graham (1997)	5–7	LD	21	E	students taught planning strategy to use while dictating plans and text versus strategy instruction while handwriting text	J	1.12
Revising instruction							
V. B. Olson (1990)	6	full range	86	N	students provided with instruction on adding, deleting, substituting, paraphrasing, and rearranging ideas versus grammar instruction and process writing	J	0.24
Head (2000)	8	full range	46	E	students taught how to make substantive revisions versus traditional writing instruction	D	0.37
Sengupta (2000)	HS	SLL	100	E	students taught how to revise papers so that they were more reader friendly versus minimal feedback	J	0.08

(*Appendix continues*)

Appendix (*continued*)

Study	Grade	Participant type	*n*	Genre	Treatment	Publication type	Effect size
Rubric instruction							
Collopy and Bowman (2005)	4	full range	100	N	teachers implemented a six-traits analytic model and students learned to identify each trait and evaluate their writing versus unspecified control condition	P	0.26
Guastello (2001)	4	full range	167	N	students taught to score rubric for evaluating their writing versus no-treatment control	J	1.27
Andrade and Boulay (2003)	7–8	full range	119	N,E	students taught to self-assess writing by means of scoring versus familiarization with the rubric	J	0.00
Scales: Minimal instruction							
Knudson (1991)	4,6, 8	full range	82	E	students used scales and questions to guide their writing versus free writing	J	−0.23
Knudson (1991)	4,6, 8	high	82	E	students used scales and questions to guide their writing versus free writing	J	−0.29
Vocabulary							
Duin and Graves (1987)	7	full range	53	E	students provided with intensive vocabulary instruction versus traditional vocabulary instruction	J	1.21

Note. HS = high school. For participant type, LD = students with learning disabilities; full range = normal variation in regular classroom, high = above average writers, SLL = second language learners, SNL = special needs learners (students with learning disabilities and speech and language difficulties). For genre, N = narrative, E = expository. For publication type, D = dissertation or thesis, J = journal, B = book, R = report, P = conference presentation. Studies with two asterisks in front of them were computed with gain scores. A question mark indicates that information for a particular category was unavailable.

Received February 9, 2006
Revision received March 27, 2007
Accepted April 16, 2007 ■