

Self-Regulated Strategy Instruction in College Developmental Writing

Charles A. MacArthur, Zoi A. Philippakos, and Melissa Ianetta
University of Delaware

The purpose of this study was to evaluate the effects of a curriculum for college developmental writing classes, developed in prior design research and based on self-regulated strategy instruction. Students learned strategies for planning, drafting, and revising compositions with an emphasis on using knowledge of genre organization to guide planning and self-evaluation. In addition to specific writing strategies, students learned strategies for self-regulation. This quasi-experimental study involved 13 instructors and 276 students in 19 developmental writing classes at 2 universities. The curriculum was taught for a full semester in 9 classes and compared with a business-as-usual control condition in 10 classes. Significant positive effects were found for overall quality of writing on a persuasive essay ($ES = 1.22$), and for length ($ES = .71$), but not for grammar. Significant positive effects were also found for self-efficacy and mastery motivation.

Keywords: writing instruction, strategy instruction, self-regulation, basic writing

Large numbers of students enter postsecondary programs with a high school diploma but underprepared for college in reading, writing, or math achievement. The most recent National Assessments of Educational Progress (NAEP) in writing in 2011 (NCES; National Center for Educational Statistics, 2012) found that only 27% of students in the final year of high school performed at or above a proficient level in writing; similar results were reported in 2007 and 2002 (Salahu-Din, Persky, & Miller, 2008). At the same time in 2011, 68% of high school graduates enrolled in college in the fall after graduation (National Center for Educational Statistics, 2013a). Clearly, a large proportion of new college students would not have met criteria for proficiency on the NAEP writing assessment. Statistics on participation in developmental or remedial courses reflect this pattern. According to the National Center for Educational Statistics (2013b), in 2011–2012, 33% of first-year college students took at least one remedial course in reading, writing, or math (40% in public 2-year and 30% in public 4-year institutions). Substantially higher proportions of minority students took developmental courses. For example, in 4-year public institutions, 20% of White students versus 33% of Black and 38% of Hispanic students took remedial courses; in 2-year colleges, the numbers were higher for all—32% of Whites, 54% of Blacks, and 45% of Hispanics. Although developmental courses offer an opportunity for underprepared students to attend college, only a small minority of students complete the required developmental courses

and go on to complete subsequent regular courses and finish a degree or certificate program (Attewell, Lavin, Domina, & Levey, 2006; Bailey, Jeong, & Cho, 2010; Bremer et al., 2013; Levin & Calcagno, 2008).

Despite the prevalence of developmental courses and the costs to institutions and students, little research has focused on evaluation of effective instructional approaches. Some research has investigated variations in the organization and sequencing of developmental courses, stretching them over two semesters or combining developmental and regular writing courses in the same semester, and some evidence indicates that these approaches may increase student persistence and retention (Adams, Gearhart, Miller, & Roberts, 2009; Glau, 2007). However, little research has focused on instructional methods or investigated effects on student writing achievement. A review of research on developmental literacy classes (Perin, 2013) found only 13 experimental or quasi-experimental studies, none of which focused on approaches to writing instruction. A review by a panel for the National Research Council (National Research Council, 2012) found seven instructional studies, but the only one focused on writing was a pilot study. Basic writing is an active field of scholarship, but it is part of the discipline of English with traditions of scholarship that have moved away from empirical research, particularly research on instructional methods, over the past 25 years (Haswell, 2005). Research on effective instruction for this large population is much needed.

The current study was the culmination of a 3-year design research project that addressed this problem by designing, implementing, and evaluating curriculum and instructional approaches based on self-regulated strategy instruction in developmental writing courses (MacArthur & Philippakos, 2012, 2013). In this curriculum, students learn strategies for planning, drafting, and revising compositions with an emphasis on using knowledge of genre to guide planning and self-evaluation. In addition, students learn strategies for self-regulation, including goal setting, task management, progress monitoring, and reflection.

This article was published Online First November 10, 2014.

Charles A. MacArthur and Zoi A. Philippakos, School of Education, University of Delaware; Melissa Ianetta, Department of English, University of Delaware.

Research reported in this article was supported by the Institute for Education Science, U. S. Department of Education, Grant R305A100614. The opinions expressed are those of the authors.

Correspondence concerning this article should be addressed to Charles A. MacArthur, School of Education, 015 Willard Hall, University of Delaware, Newark, DE 19716. E-mail: macarthur@udel.edu

Strategy instruction in writing, particularly self-regulated strategy development (SRSD), has been studied extensively with struggling and average writers at the elementary and secondary levels and has been found to have large effects on writing quality (Graham, Harris, & McKeown, 2013; Graham, McKeown, Kihara, & Harris, 2012; Graham & Perin, 2007). Graham and colleagues (2012) reported an effect size on writing quality of 1.02 for writing strategy instruction in elementary grades and an effect size of 0.50 for adding self-regulation strategies. A meta-analysis of SRSD writing studies in elementary and secondary schools (Graham et al., 2013) found an overall effect size of 1.75.

Little strategy instruction research has been conducted with college or adult learners. Two studies (Berry & Mason, 2012; MacArthur & Lembo, 2009) found positive effects of SRSD on writing by adult education learners, and another study found positive effects of tutoring in general self-regulation strategies with college students with learning disabilities (Butler, 2003). Although college composition courses and texts often include some instruction in planning strategies, we know of no research with college students on the effects of teaching writing strategies with or without self-regulation. The current study used a quasi-experimental design to test the effectiveness of the curriculum compared with a business-as-usual control group. To our knowledge, this study is the first experimental study of self-regulated strategy instruction with college students.

Theoretical and Empirical Background

Writing is a complex cognitive and social process that draws on writers' knowledge, skills, and strategies as well as their motivation and self-regulation (Hayes, 1996; Hayes & Flower, 1980; Hidi & Boscolo, 2006; Prior, 2006; Zimmerman, 1994). In addition to content knowledge, writers need rhetorical knowledge about the purposes of writing in various contexts and the genres that fit those contexts. They also need a repertoire of strategies for managing the complex processes involved in planning, drafting, evaluating, and revising. Strategies are needed to help writers use their discourse and content knowledge effectively. Complex writing also places substantial demands on self-regulation as writers set goals and manage their efforts. Because of its difficulty, writing requires substantial motivational resources.

Self-regulated strategy instruction is designed to address the demands of writing on discourse knowledge, strategy use, self-regulation, and motivation (Harris & Graham, 2009). Strategy instruction draws on multiple theories to address three key design issues: what strategies to teach, what is needed for self-regulated use, and how to teach strategies (MacArthur, 2011).

First, cognitive theories and models of composing help determine what strategies might be useful to teach (Hayes, 1996; Torrance, van Waes, & Galbraith, 2007). The core idea of cognitive strategy instruction is that it is possible to understand the strategies used by proficient writers and teach them in some form to developing writers. Proficient writers have considerable rhetorical and discourse knowledge about the purposes of writing in various contexts and the discourse forms, or genres, which fit those purposes. They also have a repertoire of strategies for managing the complex processes involved in planning, drafting, evaluating, and revising. For example, during planning, they engage in task analysis and goal setting, and use their knowledge of genres

strategically to help them generate and organize content. They also use knowledge about the characteristics of effective writing as criteria during evaluation and revision (MacArthur, 2012). Thus, most research on writing strategies has used task-specific strategies that integrate discourse knowledge (e.g., knowledge about text organization and evaluation criteria) with process knowledge about using that knowledge strategically (Englert, Raphael, Anderson, Anthony, & Stevens, 1991; Graham et al., 2013). In the current curriculum, discourse knowledge is integrated with strategies for both planning and revising. For each genre of writing, students learn the common text elements and organizational structures (e.g., thesis, reasons, evidence, and counterarguments for argumentative writing). These elements are then reflected in the design of graphic organizers that are used during planning and in the content of evaluation rubrics for revising.

Second, theories about self-regulation and motivation are critical to understanding what is required to help learners transfer strategies to independent use on other tasks in other settings. Proficient writing, at least for complex tasks, requires substantial self-regulation to manage multiple processes, maintain productivity, and learn from experience (Harris, Graham, MacArthur, Reid, & Mason, 2011; Schunk & Zimmerman, 2007). Writing strategies themselves support self-regulation by giving students a systematic way to approach complex tasks. In addition, research has found support for instruction in specific self-regulation strategies, including self-monitoring, self-evaluation, self-instructions, goal setting, self-reinforcement, and management of time and environment (Harris et al., 2011; Schunk & Zimmerman, 2007). Explicit attention to the development of self-regulation is a key component of the SRSD model (Graham et al., 2013; Harris & Graham, 2009).

The curriculum in the current study places a strong emphasis on development of self-evaluation, which both teaches students discourse knowledge and enhances self-regulation. Reviews of writing research (Graham & Perin, 2007; Hillocks, 1986) have found support for teaching specific evaluation criteria and for peer review. However, peer review requires training to be effective (MacArthur, 2012). Peer review involves both giving and receiving feedback, but recent research (Cho & MacArthur, 2011; Philippakos, 2012) has found that giving feedback, in itself, has a positive impact on writing quality. In the current approach, students learn self-evaluation criteria and receive extensive guided practice in evaluating and revising papers written by unknown peers. This instruction prepares students for peer review and self-evaluation.

In addition to self-regulation, writing demands high levels of motivation. Motivation may be especially important for struggling writers. In a qualitative study of 34 community college students in freshman composition classes, Cox (2009) found that 80% mentioned fears of failure without being prompted, and that those anxieties often led to counterproductive behaviors, like withdrawing from classes or avoiding evaluation by not submitting assignments. Writing motivation has been studied using multiple theoretical constructs. One of the most frequently studied factors is self-efficacy (Pajares & Valiante, 2006); individuals' motivation to engage in activities, their persistence in the face of difficulties, and their affect are strongly influenced by expectations for success (Bandura, 1997). Research has consistently found correlations between writing self-efficacy and achievement (Bruning & Kauffman, in press; Pajares & Valiante, 2006). Another motivational construct that has been applied to writing is achievement goal

orientation; the trichotomous model of achievement goals (Elliot & Church, 1997) includes mastery, performance-approach, and performance-avoidance goals. For writing, mastery goals have been found to correlate positively, and performance-avoidance goals negatively, with self-efficacy and affect (Kauffman, Zhao, Dempsey, Zeleny, Wang, & Bruning, 2010; Pajares & Cheong, 2003). More recently, the effects of students' beliefs about writing have been studied (White & Bruning, 2005). Students who believe that writing is a meaningful way to explore ideas and communicate are more likely to persist and learn than writers who see writing as a school exercise (Hidi & Boscolo, 2006). Finally, affect may be important as students who like to write may engage in writing more often (Bruning & Kauffman, *in press*). In the current study, we assessed motivational outcomes of the instruction using several constructs: self-efficacy, beliefs about writing, goal orientation, and affect.

Third, a large body of research on strategy instruction in reading, writing, and math has contributed to our understanding of effective instructional methods for developing self-regulated strategic learners (MacArthur, 2011). The curriculum in the current study follows the core principles of SRSD (Graham et al., 2013; Harris & Graham, 2009). First, classes discuss the purposes of each genre and the reasons for the organizational elements using models of good and weak papers. Understanding of purposes is critical to generalization to other tasks and settings. Second, instructors explain the strategies and their value and model them while thinking aloud to make the cognitive processes visible. Self-regulation strategies are modeled as part of the process and discussed with the students; during the modeling process, students observe the instructors' use of the strategies to problem solve, cope with challenges, and self-regulate. Third, instructors scaffold student use of the strategies through collaborative and guided practice, gradually releasing control as students develop independence. Instructors provide guidance both on students' use of the strategies and on their writing. Fourth, self-evaluation is emphasized. Students learn evaluation criteria based on genre elements and practice evaluating and revising papers by unknown peers before engaging in peer review and self-evaluation. Finally, motivation and self-regulation are emphasized. The curriculum includes an overall self-regulation strategy that includes goal setting, task management, progress monitoring, and reflection. Students use their journals to reflect on their learning and set new goals, and instructors guide discussion of goals, strategies, and progress. Self-regulation strategies help students to be more successful on tasks and to feel more in control of their learning.

Research Questions

The project devoted 2 years to design research with three cycles of iterative curriculum development, implementation, and revision based on qualitative and quantitative analysis (MacArthur & Philippakos, 2012, 2013). In the final round of design research, eight instructors participated with a total of 13 classes at two levels of developmental writing. Complete pretest and posttest data were available for 115 students (51% male, 51% minority, and 20% nonnative English speakers). Measures of student outcomes included written compositions scored for quality on a 7-point scale; a motivation questionnaire that assessed learning goal orientation, beliefs about writing, self-efficacy, and affect; interviews with

students; and analysis of student writing in the class. Gains in writing quality from pretest to posttest were statistically significant ($p < .001$) and large, an average gain of 2.0 on the 7-point scale ($ES = 1.95$). All but 8 of the 115 students made some gain in quality over the semester, and 62% made gains of 2 or more points. In addition, students made large gains in ratings of conventions (i.e., grammar, usage, and mechanics, $ES < 1$) from pretest to posttest. Finally, students made statistically significant and large gains in self-efficacy and affect ($ESs < 1$) and small but statistically significant gains in mastery motivation and beliefs about the importance of substantive factors in writing. Gains in motivation were supported by qualitative analyses of interviews with students. The results from the design studies represent gains from participation in a developmental writing course, but without a control group that received writing instruction typical for the setting, it is not possible to conclude that the curriculum was responsible for the gains.

The quasi-experimental study described in this article addresses that limitation. Research questions included the following: First, compared with a business-as-usual control group, do treatment students make greater gains in overall writing quality and grammar? Based on the extensive research on SRSD in writing with younger students (Graham et al., 2013) and on our own design research, we anticipated substantial gains in overall writing quality. The curriculum does not emphasize grammar beyond including editing strategies, but it is an important outcome to consider. The design study found positive changes from pretest to posttest in grammar, but in the current study, we made no specific prediction because the emphasis on grammar in the control classrooms was an unknown.

Second, do treatment students make gains in motivation including self-efficacy, goal orientation, beliefs about writing, and affect? Motivational outcomes are especially important for a population at risk of dropping out of college. Based on the design research, we anticipated positive effects on self-efficacy and affect and increases in mastery goals and beliefs in the importance of substance in writing. In addition, mastery experiences and modeling that includes coping with difficulties contribute to self-efficacy (Bruning & Kauffman, *in press*), and SRSD includes such modeling and encourages students to assess their own progress toward mastery. Some prior research on SRSD in writing has reported no effect on self-efficacy (e.g., Graham, Harris, & Mason, 2005), possibly because students' pretest self-assessments were unrealistically high. In the case of developmental writing classes, students begin the semester knowing that they were placed in these classes based on low performance, so we anticipated improved self-efficacy.

Third, can college writing instructors implement the curriculum with fidelity based on the professional development provided? It is always important to document fidelity of implementation to ensure that observed outcomes are because of the planned treatment. In the current study, we also had a substantive interest in whether our professional development was effective.

Method

Participants and Sites

Recruitment. Institutions in the mid-Atlantic region were recruited in two ways: (a) direct contact via email and letters to administrators in colleges with substantial basic writing programs and (b) contacts with instructors via announcements on listservs

for college writing. Both forms of communication included an invitation to participate in a grant-supported research study and provided brief summaries of the instructional approach and the results of the design study. When instructors responded, we arranged to communicate with administrators and vice versa. At Site A, an instructor contacted us first, and at Site B, the first contact was with the head of the English department. Criteria for participation by an institution included participation of a minimum of four instructors, two each for the treatment group and business-as-usual control group. Both full-time and adjunct instructors were acceptable. The curriculum for the basic writing courses had to include persuasive writing. Instructors volunteered to participate in the treatment group, and others were recruited for the control with a promise that they would receive professional development and materials the following semester. Treatment instructors agreed to participate in 3 days of professional development before the semester and implement the curriculum as designed for research purposes; all instructors agreed to cooperate in data collection and permit observation in their classes. Both treatment and control instructors were paid a stipend for participation during the semester, and treatment instructors received an additional stipend for participation in professional development. Three 4-year colleges agreed to participate and met the requirements. One institution dropped out after professional development because one of the two required treatment instructors resigned.

Sites and instructors. The participants were 13 college instructors and 276 students who were recruited from two 4-year universities in the mid-Atlantic region of the United States. Three instructors from Site A and three instructors from Site B volunteered to participate in the study and received training on the experimental curriculum. Additional instructors (three from Site A and four from Site B) agreed to participate in a business-as-usual control condition. Control instructors confirmed that they included persuasive writing in their instruction and agreed to participate in data collection activities, but otherwise taught as they usually did. The instructors were White women except for one White man in the treatment group and one Black man in the control group. Of six treatment-group instructors, the highest degrees included one doctorate, four masters' in English, and one masters' in library science. Of the seven control-group instructors, the highest degrees included three doctorates, one masters' in English, and three bachelor's degrees. Experience teaching writing ranged from 0 to 16 years (median = 10) for the treatment group and from 5 to 30+ years (median = 8) for the control group.

Some instructors taught multiple sections. Site A included five treatment sections with 71 students and five control sections with 82 students. Site B included four treatment sections with 57 students and five control sections with 66 students. Both sites provided access to a writing center and encouraged students to seek outside support for their writing. Both sites offered basic writing classes at two levels. At Site A, the three lower level sections offered tutors assigned to the classes (one section with 11 students in control, two sections with 30 students in treatment). At Site B, two of the nine sections were a lower-level course (one section with 11 students in control, one section with 13 students in treatment). Students enrolled in sections individually online based on their own schedules.

Student participants. The 19 classes included 303 students. Of these, 9 could not give informed consent because they were

younger than 18, and 18 declined to participate. Of the 276 participants who participated in pretesting, 24 (8.6%) did not complete the posttest assessment. Ten were absent at posttest ($n = 5$ control, $n = 5$ treatment), and one did not complete the posttest because of a computer malfunction ($n = 1$ treatment). Furthermore, during the semester 11 participants dropped out of the course ($n = 4$ control, $n = 7$ treatment), and 2 dropped out of the study (control group). Of the 24 without a posttest assessment, 6 attended Site A ($n = 3$ control, $n = 3$ treatment) and 18 students attended Site B ($n = 8$ control, $n = 10$ treatment). Attrition was below 9% and nearly equal between conditions. Overall, 252 students completed the pretest and posttest assessments ($n = 137$ control, $n = 115$ treatment). Of this sample of 252, 54% were female, 52% were White (35% Black, 4% Asian, 6% Latino, and 6% other); 10% of the participants were not born in the United States or English was not the primary language spoken at home. Site B had a larger percentage of Black students than Site A, a larger proportion of nonnative English speakers, and a larger proportion of students older than 24. However, there were no significant demographic differences between treatment and control groups ($p > .3$) at either site (see Table 1).

The Writing Fluency and Writing Samples subtests from the *Woodcock-Johnson Tests of Achievement: Third Edition* (WJ-III, Woodcock, McGrew, & Mather, 2001, see Measures), were administered at pretest, and scores on the SAT Writing Test were available for Site A (see Table 1). A statistically significant difference favoring the control group was found for the WJ-III Writing Samples subtest ($F(1, 250) = 5.9, p < .05$, Cohen's $d = .30$). No differences were found for the Writing Fluency subtest or the SAT writing test ($p > .2$). The WJ-III scores were included as covariates in the analyses to control for preintervention differences on writing achievement.

Measures

Instructor measures. Treatment instructors were observed to provide ongoing professional development and to evaluate fidelity of treatment. Control instructors were observed to describe their instruction; they also provided copies of their assignments and samples of students' work. In addition, all instructors were interviewed before and after the semester.

Fidelity of implementation. Treatment instructors were observed at least four times during the semester. A trained observer took detailed field notes during the class and then completed the fidelity of implementation measure, which included a checklist of lesson components and ratings of quality of key elements of the instructional approach. The checklist included all the lesson activities; each activity was rated 0 to 2 (0 = *absent*, 1 = *done with modifications*, 2 = *done as in the lesson plan*). Quality of implementation was rated on 3-point rubrics for five key elements: introduction of a new genre, analysis of good and weak examples, explanation and modeling of the strategy, collaborative practice, and peer review. Only components relevant to the day's lesson were rated.

Observers were the first two authors and a graduate research assistant who was thoroughly familiar with the curriculum. The quality of implementation rubric was developed by the authors and applied to video-recorded lessons from prior work to refine the rubric and establish video anchors for training. Training of the

Table 1
Demographics and Pretest Standardized Writing Scores of Student Participants

	Site A		Site B		Total	
	Treatment <i>n</i> = 68	Control <i>n</i> = 79	Treatment <i>n</i> = 47	Control <i>n</i> = 58	Treatment <i>n</i> = 115	Control <i>n</i> = 137
Gender						
Female	58.8%	53.2%	55.3%	48.3%	57.4%	51.1%
Ethnicity						
Black	25.0%	27.8%	44.7%	48.3%	33.0%	36.5%
Asian	1.5%	1.3%	2.1%	1.7%	1.7%	1.5%
White	63.2%	62.0%	38.3%	34.5%	53.0%	50.4%
Latino	5.9%	2.5%	10.6%	6.9%	7.8%	4.4%
Other	4.4%	6.3%	4.3%	8.6%	4.3%	7.3%
Age < 24	100.0%	100.0%	83.0%	79.3%	93.0%	91.2%
Native English speakers	94.1%	97.5%	83.0%	82.8%	89.6%	91.2%
W-J III fluency ^a	25.1	24.8	20.7	20.8	23.3	23.1
<i>M</i> (<i>SD</i>)	(4.7)	(4.6)	(5.0)	(5.6)	(5.3)	(5.4)
W-J III samples ^a	15.4	17.3	14.3	14.0	15.0	15.9*
<i>M</i> (<i>SD</i>)	(2.8)	(2.2)	(3.0)	(3.3)	(2.9)	(3.2)
SAT writing	435	444	NA	NA	NA	NA
<i>M</i> (<i>SD</i>)	(44)	(35)				

^a Woodcock-Johnson Tests of Achievement: Third Edition.

* $p < .05$ by condition.

research assistant included discussion of the checklist and rubrics and practice using the video. For every observation, a second observer read the field notes and completed the checklist and quality ratings. Interrater reliability on the checklist of lesson components was 100%; reliability on the quality ratings was 94% exact agreement.

Interviews of treatment instructors. Instructors were interviewed before professional development to gather information on their education and experience and on their prior teaching practices. They were interviewed after the semester to gather information on their perspectives on the curriculum and their students' performance. Postinterviews included questions on instructors' overall evaluation of the curriculum, impact on student learning and motivation, feasibility for instructors, and evaluations of particular components.

Control observations, assignments, and writing samples. Control instructors were observed three times each, except for one who was observed only twice. An observer took detailed field notes and collected any materials used in the class. In addition, control instructors provided copies of all writing assignments together with the work completed by one typically performing student. The final assignment was collected for all students.

Interviews of control instructors. Instructors were interviewed before the semester to gather information on their education and experience and on their teaching practices. They were interviewed at the end of the semester for information on their instruction, including assignments and instructional methods.

Student measures. Students wrote persuasive essays at pretest and posttest that were scored for overall quality, length, and grammar. Also at pretest and posttest, they completed a motivation questionnaire tapping goals, beliefs, self-efficacy, and affect. At pretest only, standardized tests of writing were administered.

Essays. In the first week of class and in the final examination, students wrote persuasive essays. Although students learned to write in several genres, it was not feasible to administer multiple

posttest writing tasks and still give students time to write and revise. Persuasion was chosen because of its fundamental importance to academic writing (Nussbaum & Kardash, 2005; Wolfe, 2011). Two surveys of college instructors of composition, rhetoric, and English (ACT, Inc., 2012; Milewski, Johnson, Glazer, & Kubota, 2005) found that instructors gave the highest importance ratings to persuasive/argumentative writing. In addition, persuasive writing draws on the other genres; for example, causes and effects are common reasons, and personal experience can be used for support. Control instructors confirmed in advance that they taught persuasive writing, and data from control classrooms verified that they did. At each time, pretest and posttest, students had a choice of three topics, which had been pilot-tested in prior research (MacArthur & Philippakos, 2013). Although the writing prompts were comparable, the two testing occasions were not. Students wrote the pretest with the understanding that it was a baseline measure of their writing achievement; the posttest counted as part of their grade. In addition, all posttest essays were written on word processors, and all pretest essays were handwritten because classes did not have access to computers. Although the gains from pretest to posttest are not a valid measure of learning, the pretest measure is adequate for use as a covariate to control for individual differences.

Pretest essays were typed and checked for accuracy by research assistants. Spelling errors, including homonyms, apostrophes, abbreviations, and capitalization of proper nouns, as well as general spelling errors, were corrected before scoring for quality and grammar. Correcting such errors is good practice because spelling errors can bias raters' judgment of quality (Graham, Harris, & Hebert, 2011); in addition, in this case, students had access to a spell checker on the posttest. Spelling errors were also counted. One rater scored all papers, and 10% of the papers were scored independently by a second rater. Interrater reliability was adequate (Pearson $r = .93$).

Quality. Two independent raters, unaware of the purpose of the research study, rated papers for overall quality on a 7-point rubric. The holistic rubric directed raters to form an overall judgment of quality based on criteria for ideas or content, organization, word choice, sentence fluency, and errors in grammar and usage. Raters were trained using the rubric and anchor papers. Interrater reliability was adequate with a correlation of .82; exact agreement was 52% and agreement within one point was 92%.

Grammar, mechanics, and usage. First, two raters independently divided essays into T-units with the second rater marking 15% of the essays. A T-unit (Hunt, 1964) is defined as a main clause and all embedded subordinate clauses and phrases; it provides a measure of syntax independent of punctuation. Interrater reliability on the number of T-units was .99 (Pearson *r*). Subsequently, two different raters scored the essays for grammar errors; one scored all the papers, and the other scored a random 15%. Each T-unit was scored as correct or incorrect; incorrect T-units were then scored as containing errors in language and/or mechanics. Mechanics errors included all errors that would not be incorrect if the text were read aloud with appropriate prosody (primarily punctuation and capitalization since spelling had been corrected). Language errors included all errors that would remain errors on reading aloud (e.g., syntax, morphology, semantics, and usage). Interrater reliability for total errors was .92 (Pearson *r*); reliabilities for mechanics and language errors were .94 and .81, respectively.

Length. Length was calculated using the word processor's word-count feature.

Standardized writing test. Two subtests from the WJ-III (Woodcock et al., 2001) were used: Writing Fluency and Writing Samples. Writing Fluency is a timed test of how many correct sentences students can write using three given words and a picture. Writing Samples is an untimed test of sentence writing with increasing demands for vocabulary, concepts, and grammar. The WJ-III is a well-known test with generally adequate reliability and validity (Bradley-Johnson, Morgan, & Nutkins, 2004). A recent review of the adequacy of the test for older students, aged 16–25 (Krasa, 2007) found generally adequate reliability for the writing subtests. One rater scored all tests, and 10% were scored independently by a second rater; interrater reliabilities were .94 for Writing Fluency and .93 for Writing Samples (Pearson *r*).

Motivation questionnaire. The motivation questionnaire consisted of four scales: self-efficacy for writing, achievement goals for writing, beliefs about writing, and affect. The self-efficacy scale included three factors for writing tasks and processes (e.g., I can write a paragraph with a clear topic sentence), self-regulation (e.g., I can avoid distractions when I write), and grammar skills (e.g., I can use commas and semicolons in my sentences) and included 22 items. The goals scale included mastery (e.g., when I am writing in this class I am trying to better organize my ideas), performance (e.g., when I am writing in this class I am trying to get my teacher to think I am a good writer), and avoidance goals (e.g., when I am writing in this class I am trying to hide that I have a hard time writing) and consisted of 14 items. The beliefs scale assessed beliefs about the importance of substance (e.g., writing helps me think about my topic in a new way) and mechanics (e.g., good writers need little revision because they get it right the first time) with 12 items. The affect scale included five items about feelings and attitudes toward writing (e.g., I find writing satisfying). Prior research demonstrated adequate construct

validity and internal consistency (Cronbach α s ranging from .66 to .93) (Philippakos, MacArthur, & Uribe-Zarain, 2013). Internal consistencies for data in the current study ranged from .59 to .92.

Procedures

Treatment instruction. The curriculum included a consistent set of strategies for planning, drafting, evaluating, and revising that was used across all instructional units. In each genre-based unit of instruction, the basic writing strategy remained the same, whereas the purpose and text structure elements changed with the genre. Students gradually learned to apply the writing strategy flexibly as they applied it to varied types of writing. Planning included setting goals by analyzing the writing task for topic, audience, purpose, form, organizational elements (i.e., genre elements), and requirements (e.g., length), brainstorming, and organizing using a graphic organizer that differed by genre. Drafting included using the plan, writing main idea sentences, and providing supporting details. Revising included self-evaluation and peer review based on evaluation criteria reflecting the elements of the genre, as well as editing.

In addition to the writing strategy, the curriculum included a set of self-regulation strategies called *Strategies for Academic Success*. These strategies included goal setting, task management, progress monitoring, and reflection. Goal setting helps learners to direct their efforts on complex tasks like writing and motivates effort by helping them see progress (Zimmerman, 1994). Students were asked to discuss long-term goals, identify their short-term goals within each learning unit, and set specific personal goals for each assignment. Task management is a critical aspect of self-regulation; this strategy included a focus on management of time, place, and effort for completing work and, equally important, management of emotion and motivation. Checking progress asked students to consider two questions: Am I using the strategies? Are they helping me to improve my writing? Research shows that teacher feedback on both of these questions can improve student writing (Schunk & Swartz, 1993). The overall purpose was to improve motivation by convincing students that they can be successful when they identify and use appropriate strategies for achieving a specific goal. Finally, reflection occurred after completion of each assignment. Students were asked to consider what strategies they had used and whether the strategies were helpful, what they had learned about writing, and what goals to set for the next assignment. Overall, the self-regulation strategies were designed to encourage students to take responsibility for their own learning by setting goals, choosing strategies, managing their efforts, and reflecting on their progress. In the curriculum, journal entries, classroom discussions, and individual conferences were used to help students understand and apply these strategies. Teachers also reflected on their own use of these strategies on tasks that were challenging to them.

The curriculum was organized into units of instruction on various genres, or rhetorical modes. Instructors at each site were given flexibility in choosing which units to include in their courses. At Site A, instructors decided to focus on essay writing and to include personal narrative, cause-effect, and persuasive writing. At Site B, instruction began with paragraph writing with a switch to multiparagraph essays when instructors thought students were ready; instruction included narrative, procedural, cause-effect, and per-

suasive writing, although the lower-level classes omitted cause-effect writing.

Instruction within each unit followed a consistent pedagogical sequence—a *strategy for teaching strategies*: (a) Introduce the genre. The instructor led a discussion of the purpose of the genre and the text elements and features needed to achieve that purpose. (b) Analyze good and weak examples. The class analyzed a good example, discussed what made it good, and analyzed the genre elements. The elements were then presented as evaluation criteria that were applied to evaluate first the good paper and then the weak one. Students discussed how the papers could be improved. (c) Explain the strategy. The instructor briefly explained the writing strategies and their purposes. After the first unit, the explanation was mostly review with emphasis on how the strategy fit the new genre (e.g., the new graphic organizer). (d) Model the strategy. The instructor demonstrated the strategy using think-aloud modeling and including self-regulation statements. Instructors did some preparation for this modeling, but the emphasis was on live modeling to show the thought and struggle involved in planning, drafting, and revising. The student role was mostly to observe, but teachers often included them in generating ideas. (e) Collaborative practice. The class wrote a paper collaboratively with students providing all the content and the instructor guiding the use of the strategy, scaffolding student attempts, and scribing. (f) Guided practice. Students used the strategy to plan, draft, and revise their own papers with feedback from instructors on their use of the strategy and their writing. (g) Peer review and editing. Students learned to evaluate papers by applying the genre-specific evaluation criteria to papers written by unknown peers. They evaluated the papers on each criterion and made suggestions for improvement. Instructors guided this practice. Subsequently, students self-evaluated their papers and then met in pairs for peer review. (h) Support self-regulation. Instructors supported self-regulation throughout the course, using the Strategy for Academic Success as discussed earlier. They also included self-regulation comments (e.g., directing efforts, managing motivation) in their modeling. (i) Support learning the strategies. Instructors used review to ensure that students were able to explain the steps in the strategy and the elements of the genres. (j) Independence. Independent, self-regulated, use of the strategies was the goal. In most cases, students wrote a second paper in each genre with less instructor support.

Grammar, mechanics, and usage were taught in the context of editing. Research does not show a clear benefit of separate grammar instruction (Graham & Perin, 2007; Hillocks, 1986). The evaluation criteria used for revision emphasized the importance of writing clear sentences (e.g., clear thesis statement), an emphasis that might resolve many problems of sentence structure. In addition, each editing lesson included guided practice in fixing typical errors in the context of complete texts, followed by peer editing and self-editing with teacher support. For these editing lessons, the curriculum included a few sample texts and suggestions for instruction. However, individual instructors were encouraged to select the grammar and mechanics concepts their students needed and create texts for the editing practice. Instructors were also permitted to provide additional grammar instruction that they found appropriate, but to limit the time devoted to such instruction to 20 min a week. Based on posttest interviews and observations, only one treatment instructor provided extra grammar instruction;

she had her students use an online grammar practice site and went over the concepts in class.

Professional development. Instructors in the treatment group participated in 3 days of professional development (PD) during the summer and received ongoing support during the semester. Sessions were scheduled separately for the two sites to accommodate their schedules. PD was provided by the authors of this article and was designed to apply the principles of strategy instruction to PD. The key instructional principles and activities were explained as a *strategy for teaching strategies*. Thus, we explained and discussed the principles of self-regulated strategy instruction, modeled via demonstration teaching with the instructors as students, provided guided practice with feedback as the instructors taught the lessons, discussed application across the genres, and addressed questions and concerns. The first session was devoted to explanation of the curriculum and demonstration teaching (modeling) by the researchers. The day began with a presentation and discussion of the core principles of self-regulated strategy instruction, which was followed by an explanation of the strategies in the curriculum and an overview of the units of instruction. Instructors received copies of the teachers' guide, which included detailed lesson plans, and the student book. Then researchers modeled teaching one of the units, and the instructors participated as students, planning and drafting a paper. Instructors prepared to teach one of the units in the curriculum, which they did during the second day. The researchers provided guidance and feedback with further discussion of the pedagogy of self-regulated strategy instruction. In addition, time was devoted to discussion of how to support the development of the self-regulation strategies. The final day was devoted to questions about the curriculum and discussion of the experimental procedures and requirements.

During the semester, project staff observed treatment classrooms to provide continued professional development as well as to rate fidelity of implementation; on request, staff met with instructors outside of class to help with planning. When possible, observers met briefly with instructors right after class to answer questions and provide feedback. In addition, observers sent written feedback. Feedback was guided by the rubric for key elements of strategy instruction on the fidelity of treatment rating. Each treatment instructor was observed at least four times.

Control instruction. Control instructors continued with their prior approaches to instruction. As part of selection for participation, they all confirmed that they included persuasive writing in their courses. Furthermore, they looked at the baseline writing prompts and agreed that students should learn this type of writing in their classes. The following description of their instruction is based on observations, collection of course assignments and writing samples, and interviews at the end of the semester (see Measures). All instructors included at least four writing assignments in at least three different genres. All but one instructor included a writing assignment that required persuasive writing. The exception included a classification and comparison essay that required students to state a thesis and support it with evidence, which required many of the skills needed for persuasive writing. All instructors required regular writing in a journal. Consistent with the choices made by treatment instructors, the control instructors at Site A all asked students to write essays, whereas those at Site B asked students to write paragraphs initially and essays later. Control

classes met on the same weekly schedules and for the same total time as treatment classes at the same site.

Most of the control instructors (6 of 7) taught some strategies for planning. Five provided some instruction on brainstorming and clustering or using a graphic organizer; four taught outlining. For these six instructors, the final writing assignments submitted by students included some written evidence of planning. The same six instructors required students to submit multiple drafts and engaged their students in peer review. However, only three of them reported teaching students how to evaluate each other's work before peer review. As for self-regulation strategies, when asked if they focused on time management, only two control instructors reported that they spent substantial time discussing work habits, time management, or goals.

All of the control instructors assigned homework exercises on grammar from a textbook and devoted some class time to instruction in grammar and conventions. Their particular practices varied. Three reported including grammar and conventions on quizzes. Two others reported that they taught grammar before peer editing sessions and asked students to help each other edit for errors. One had students use a sentence-combining workbook. Grammar and mechanics were not the major focus in any of the control classes.

Analysis

Considering that the students were nested within sections taught by different instructors, hierarchical linear modeling (HLM) was used as it allowed the examination of the effects of this nesting (Bryk & Raudenbush, 1992). The effects of the curriculum on the quality, length, and grammatical errors of written responses were estimated in three-level HLM models, where students (Level 1) were nested within sections (Level 2) that were nested within instructors (Level 3). Site was included in the model as a fixed factor because it was crossed with treatment condition; additionally, we wanted to determine whether the curriculum was differentially effective across sites with different characteristics. Covariates were included in the model to control for preintervention differences between treatment and control on writing achievement. Specifically, each analysis controlled for the relevant pretest scores and for performance on the two W-J III writing measures. Finally, treatment condition was included as a fixed factor at Level 3 to estimate the impact of the intervention.

Analysis of covariance (ANCOVA) was used to analyze data on motivation. Analyses were conducted for each construct—goals, beliefs, self-efficacy, and affect. Within each construct, separate analyses for each factor were interpreted with Bonferroni adjustments. Because of the multiple analyses, there was insufficient power for analysis using HLM.

Results

Fidelity of Implementation

Fidelity of implementation was measured using a checklist of lesson components and ratings of quality of key elements of strategy instruction (see Measures). Each component in the lesson observed was coded as completed as planned (2), completed with modifications (1), or omitted (0). Overall fidelity was high. Across all instructors and components, the average rating was 1.63, and

only 10% of lesson components were omitted. Average ratings for instructors ranged from 1.2 to 2.0, with only one instructor below 1.6; components omitted ranged from 0 to 15%.

Quality of key elements was rated on a 3-point scale with 3 representing very good instruction. Quality of instruction was generally high. Across all instructors and elements, the average rating was 2.7 ($SD = .6$). Ratings were averaged for each element and for each instructor. Average ratings for the elements across teachers ranged from 2.4 for collaborative practice to 2.8 for explanation and modeling of planning strategies. Average ratings for instructors ranged from 2.0 to 3.0.

Writing Outcomes

Quality. Data on quality and length are presented in Table 2. The HLM analysis for quality found statistically significant effects for the covariates of pretest essay quality and the W-J III Writing Fluency score ($p < .001$ and $p = .052$, respectively), but not for the Writing Samples score ($p = .195$). A main effect of condition was found, $F(1, 7.3) = 40.0$, $p < .001$ with the treatment group producing higher quality essays compared with the control. The effect size for condition was large (Glass's $\Delta = 1.22$). Glass's Δ , which uses the posttest control standard deviation, was appropriate because the variance at posttest increased for the treatment group. A main effect of site was also found, $F(1, 8.3) = 15.5$, $p = .004$, but no interaction was found between site and condition ($p = .41$). A further analysis was conducted with class level (lower and higher classes) as a fixed factor to examine the effects on the model. The results indicated no effect of class level ($p = .78$) and no interaction with condition ($p = .20$).

Descriptively, the mean gains from pretest to posttest on the 7-point quality scale were 1.0 for the control group and 2.5 for the treatment group. The mean gain scores for all classes in the treatment group were higher than for all classes in the control group, except for one tie (i.e., one treatment and one control class had mean gains of 1.7). At the individual level, in the control group, 19% of students made no gain; 63% gained at least 1.0; 20% gained at least 2.0; and only 3% gained at least 3.0. In the treatment group, only one student made no gain; 98% gained at least 1.0; 76% gained at least 2.0; and 37% gained at least 3.0.

Length. A similar HLM analysis was conducted for length. The analysis found statistically significant effects for the covariates of pretest essay length and the W-J III Writing Fluency score ($p < .001$ and $p = .05$, respectively), but not for the Writing Samples score ($p > .4$). A main effect of condition was found, $F(1, 7.6) = 7.52$, $p = .027$ with the treatment group producing longer essays than the control. The effect size for condition was moderate (Glass's $\Delta = 0.71$). A main effect of site was found ($p = .010$), but no interaction between site and condition ($p > .2$).

Grammar. A similar analysis was conducted for grammar. The analysis found statistically significant effects for the covariates of pretest essay grammar and the W-J III Writing Fluency score ($p < .001$ and $p = .002$, respectively), but not for the Writing Samples score ($p = .057$). There was a main effect of site ($p = .011$), but no effect of condition ($p = .086$), and no interaction between site and condition ($p = .30$). The effect-size for condition was small (Glass's $\Delta = 0.18$). Both groups made errors in about one third of their sentences.

Table 2
Means and SDs for Essay Quality, Length, and Grammar

	Site A		Site B		Total	
	Treatment <i>n</i> = 68	Control <i>n</i> = 79	Treatment <i>n</i> = 47	Control <i>n</i> = 58	Treatment <i>n</i> = 115	Control <i>n</i> = 137
Quality ^a						
Pretest	2.53	3.31	2.30	2.48	2.44	2.96
<i>M</i> (<i>SD</i>)	(0.80)	(0.74)	(0.67)	(0.70)	(0.75)	(0.83)
Posttest	5.40	4.42	4.34	3.39	4.97	3.98
<i>M</i> (<i>SD</i>)	(0.89)	(0.78)	(0.99)	(0.68)	(1.07)	(0.90)
Adjusted posttest	5.37	4.13	4.52	3.58	4.94***	3.86***
Length ^b						
Pretest	211	338	204	217	207	286
<i>M</i> (<i>SD</i>)	(93)	(105)	(83)	(72)	(89)	(110)
Posttest	826	721	544	459	710	609
<i>M</i> (<i>SD</i>)	(182)	(207)	(188)	(142)	(230)	(223)
Adjusted posttest	862	627	590	509	726*	568*
Grammar ^c						
Pretest	.683	.680	.613	.524	.655	.614
<i>M</i> (<i>SD</i>)	(.16)	(.14)	(.22)	(.20)	(.19)	(.19)
Posttest	.725	.722	.642	.566	.691	.656
<i>M</i> (<i>SD</i>)	(.14)	(.13)	(.16)	(.18)	(.15)	(.17)
Adjusted posttest	.699	.688	.668	.620	.684	.654

^a Overall quality was rated on a 7-point scale. ^b Length = number of words. ^c Grammar is the proportion of T-units without errors of grammar, mechanics, or usage.
* *p* < .05. *** *p* < .001.

Motivation

Scores on the motivation measure are presented in Table 3. ANCOVA was used to analyze data for each construct—goals, beliefs, self-efficacy, and affect—with the relevant pretest score as a covariate. Within each construct, separate analyses for each factor were interpreted with Bonferroni adjustments.

Goals. A statistically significant effect was found for mastery goals ($F(1, 249) = 7.01, p = .009$, partial $\eta^2 = .027$, Cohen’s $d = 0.29$). The students in the treatment condition scored higher than the students in the control condition. No effects were found for performance or avoidance goals (both *ps* > .4).

Self-efficacy. A statistically significant effect was found for self-efficacy for tasks/processes ($F(1, 249) = 7.58, p = .006$, partial $\eta^2 = .030$, Cohen’s $d = 0.27$). The treatment students

scored higher than control students. No effects were found for grammar (*p* = .18) or self-regulation (*p* > .3).

Beliefs. No statistically significant effects were found for substance or mechanics (both *ps* > .2).

Affect. No statistically significant effect was found for affect (*p* = .067).

Treatment Instructors’ Perspectives

Instructors’ perspectives are an important social validity outcome; it is unlikely that instructors will use a curriculum unless they perceive it as feasible and effective. Postinterviews were analyzed for overall evaluation of the curriculum, impact on student learning and motivation, feasibility for instructors, and comments and suggestions on particular components. Two researchers

Table 3
Means and SDs for Motivation Factors

Scale	Factors	Treatment (<i>n</i> = 115)			Control (<i>n</i> = 137)		
		Pre <i>M</i> (<i>SD</i>)	Post <i>M</i> (<i>SD</i>)	Adjusted post	Pre <i>M</i> (<i>SD</i>)	Post <i>M</i> (<i>SD</i>)	Adjusted post
Goals	Performance	2.9 (.9)	3.2 (1.0)	3.29	3.2 (.9)	3.3 (.9)	3.29
	Avoidance	2.5 (.9)	2.7 (1.0)	2.70	2.7 (.9)	2.7 (.9)	2.63
	Mastery	4.0 (.6)	4.3 (.6)	4.26**	4.1 (.6)	4.1 (.6)	4.08**
Self-efficacy	Tasks/processes	67.2 (14.5)	77.7** (12.8)	77.1**	65.34 (15.0)	72.7 (15.5)	73.2**
	Grammar	59.5 (20.4)	68.9 (16.7)	69.3	60.8 (18.6)	67.5 (18.0)	67.2
	Self-regulation	66.4 (17.9)	73.1 (16.1)	72.8	65.3 (18.3)	70.9 (18.9)	71.2
Beliefs	Substance	3.8 (.6)	4.0 (.6)	3.99	3.7 (.6)	3.9 (.6)	3.93
	Mechanics	2.5 (.6)	2.5 (.7)	2.50	2.6 (.6)	2.5 (.7)	2.49
Affect	Affect	2.9 (.9)	3.3 (.9)	3.37	3.0 (.8)	3.3 (.8)	3.22

Note. All measures are on a scale of 1–5, except for self-efficacy with a scale of 0–100.
** *p* < .01.

independently analyzed the interviews and then discussed their conclusions.

All six instructors had an overall positive response to the curriculum. They thought that most of their students had learned to use the writing strategies and had made gains in writing achievement and motivation. All instructors commented positively on the organized set of writing strategies that integrated genre elements but remained consistent across the genres. They agreed that the strategies provided a systematic way for students to approach writing assignments that was helpful for these basic writers. Except for the one new instructor, all had taught strategies in the past but not in such an organized way. One of the new instructional aspects for the instructors was think-aloud modeling; only one instructor reported having used some modeling previously. All found it initially quite challenging, but agreed that it was a critical aspect of the instruction. All six instructors also found the *Strategies for Academic Success* to be effective in helping students gain more control of their writing assignments. Some commented on the value of integrating these self-regulation strategies with the writing course instead of teaching them in a separate workshop on study skills. The instructors reported that peer review was initially difficult for students, but that instruction in how to evaluate papers made the process effective. Instructors did identify some areas for improvement. Two instructors suggested combining think-aloud modeling with collaborative practice in later units. Two reported that students found the task analysis procedure too repetitive in later units. Instructors in the higher performing classes requested longer and higher quality essays to use as good examples.

Discussion

This study was the first experimental or quasi-experimental evaluation of self-regulated strategy instruction with college developmental writers. Treatment classes used a curriculum based on SRSD that included systematic instruction in genre-specific strategies for planning, drafting, evaluating, and revising together with self-regulation strategies for goal setting, task management, progress monitoring, reflection, and self-evaluation of writing. The results indicate that the self-regulated strategy curriculum had a large positive effect on the overall quality of students' writing ($ES = 1.22$). On average, treatment students gained 2.5 points on a 7-point quality scale, while control students gained 1.0 point (see Table 2). Furthermore, although performance differed between the two universities, the treatment was equally effective at both sites. The large effect size is consistent with prior research on SRSD (Graham et al., 2013, 2012).

In addition, the students in the treatment condition had higher self-efficacy for tasks and processes ($d = 0.27$) and higher mastery goals ($d = 0.29$). Self-efficacy has consistently been found to correlate with writing performance (Pajares & Valiante, 2006). SRSD emphasizes mastery of writing strategies and increasing motivation by showing students their improved writing performance, and mastery experiences are the clearest way to increase self-efficacy (Bandura, 1997; Bruning & Kauffman, in press). In addition, modeling that includes coping with difficulties is another feature of SRSD that has been shown to enhance self-efficacy (Zimmerman & Kitsantas, 2002). Furthermore, the development of self-regulation strategies should increase students' sense of control and self-efficacy (Harris & Graham, 2009). However, gains in

self-efficacy have not always been found as a result of instruction even when writing improved (Graham et al., 2005). The positive findings in this study are important if we consider that lower-performing students tend to be less confident and struggle with motivation (Cox, 2009; Hidi & Boscolo, 2006; Schunk & Zimmerman, 2007). The improvement in motivation may significantly affect the students' performance in future writing classes and their academic success.

The results revealed no statistically significant effects on grammar, mechanics, and usage. The curriculum did not include instruction in grammar except for brief teacher-designed lessons as part of editing. However, we had anticipated that instruction focused on helping students to express their ideas more clearly would have some impact on grammar. Control classrooms included somewhat more grammar instruction, but did not place a major emphasis on grammar either. Differences between the treatment and control groups and between pretest and posttest were quite small in educational terms. The current findings contrast with the results in the design research (MacArthur & Philippakos, 2013), which found large gains in grammar and mechanics over the semester. However, in the design research, we used a rating scale rather than counts of errors, and raters may have been influenced by impressions of quality (Graham et al., 2011).

In general, it appears that students were in need of some instruction on sentence construction; on average, regardless of time or condition, about one third of T-units had errors in grammar, usage, or mechanics. Grammar instruction is a controversial topic (Gribbin, 2005) since previous research reviews have not found positive effects of grammar instruction on students' writing (Graham & Perin, 2007). However, there is clearly a need for research on effective instruction in sentence construction. Some promising research on sentence combining (Saddler & Graham, 2005) and embedded grammar instruction (Myhill & Jones, 2007) may provide successful approaches. Further research is needed.

The results show that the treatment instructors implemented the curriculum with high fidelity. This finding suggests that the curriculum is feasible to implement in college classes and that the model of professional development used to introduce the program and provide feedback to instructors was effective. The instructors were able to implement the lessons as intended and to use pedagogical methods such as think-aloud modeling, scaffolding student use of the strategies, and supporting self-regulation strategies. Despite diversity in educational background and experience, all instructors implemented the curriculum with acceptable fidelity. It should be noted, however, that different amounts and levels of support were required across different instructors. Additional research is needed to further examine professional development on the application of strategy instruction at the college level. In fact, further research on professional development for SRSD is needed at all levels of schooling (Harris et al., 2012).

Furthermore, interviews with treatment instructors indicate that they found the curriculum to be effective in improving students' writing and motivation. Although they had taught strategies in the past, they had not used the pedagogical methods associated with self-regulated strategy instruction. They all found think-aloud modeling challenging as instructors but valuable to their students. The instructors believed that the self-regulation strategies helped their students gain more control of their writing and increase

motivation. The comments made by these instructors echo the comments of the instructors during the design research.

In addition, the value that the instructors saw in this curriculum and its principles of instruction is evident from their decision to continue its use after the completion of the study. One of the sites officially adopted the curriculum across all developmental sections. In the second site, the three treatment instructors have continued to use the curriculum for three semesters with some thoughtful adaptations.

Observations, collection of assignments, and interviews in the control condition revealed some overlap in instruction between treatment and control. First, all control instructors assigned writing in multiple genres, and all but one specifically assigned persuasive writing. All assigned regular writing in journals. Nearly all of the control instructors (6 of 7) provided some instruction in strategies for planning, including brainstorming, clustering or using a graphic organizer, and outlining, and they expected students to engage in some form of planning. In fact, the final assignments submitted by students in these classes did show evidence of planning. However, the control instructors did not model how to use the planning strategies to write a paper. The same six instructors required students to submit multiple drafts of each paper and engaged students in peer review during class. However, only three of the instructors provided any training in how to evaluate their peers' papers. Only two instructors reported class discussions about time management and other self-regulation issues. All of the control instructors taught some grammar, but it was not a major emphasis in any of the classes. It is difficult to say whether these practices are typical of developmental courses in other institutions. Although some survey data (ACT, Inc., 2012; Milewski et al., 2005) are available on the types of writing assigned in first-year college composition classes and instructors' views on what outcomes are important, we have not found any reliable information from surveys about instructional practices in developmental writing. In this study, at least, both treatment and control students were exposed to writing in multiple genres, planning strategies, peer review, and modest amounts of grammar. The differences are not about content or instructional goals but rather about systematic instruction that shows students how to use specific planning and revising strategies, supports them as they learn, and develops their self-regulation strategies.

Limitations

There are a number of limitations that should be noted. The most significant limitations arise from the quasi-experimental design. Without random assignment of instructors, it is always possible that factors other than the treatment affected the outcomes. First, treatment instructors volunteered to participate in the study whereas control instructors agreed only to participate as control classes. Although the two groups were equivalent in demographics, education, and experience, volunteers may be more motivated and interested in improving their instruction. Second, as noted under Participants, there was an imbalance at Site A in the number of lower-level classes with a tutor. We were unaware of the difference when we selected control classes because all sections had the same course number. To address the problem, we reran the quality analysis with level of class as a factor and found no effect of level or interaction with treatment. Third, three control instruc-

tors with a total of 44 students did not give academic credit for the final exam. These students did take the exam and received their financial incentive, but their motivation to do well may have been affected. We reran the analysis without these students and found the same significant effects. We are aware of that these limitations pose potentially significant problems and that our post hoc analytical strategies may not be entirely adequate. Overall, these problems with the research design were overwhelmed by the size of the treatment effect. The study was underpowered for the HLM analysis, but we found a significant difference in writing quality because of the large effect size. Further research using randomized control trials is needed to verify the current findings.

Another limitation, unrelated to the quasi-experimental design, is that the posttest of writing quality was based on a single persuasive essay. Time did not permit writing multiple essays with opportunity for planning and revising, so we chose to use persuasive writing. Researchers (Nussbaum & Kardash, 2005) and college faculty (ACT, Inc., 2012; Milewski et al., 2005) agree that persuasive writing is critical to college success, but it would have been preferable to assess writing in multiple genres, including a genre that was not taught explicitly in the treatment.

Further Research

The most immediate need for further research is to evaluate the instructional approach using randomized control trials. In addition, the need for additional research on instructional methods for teaching sentence construction and grammar was mentioned above. The need for research on professional development has also been mentioned.

In addition, further research could also investigate components of the approach, in particular, the approach to teaching self-evaluation and peer review. Relatively few SRSD studies have included instruction in both planning and revising strategies (Graham et al., 2013). The current study integrated instruction in planning and evaluating/revising by connecting both to genre elements. The focus on evaluation was based on prior research on teaching evaluation criteria (for reviews, see Graham & Perin, 2007; MacArthur, 2012) and research on the effects of giving feedback (Cho & MacArthur, 2011; Philippakos, 2012). Research might investigate the effects of this evaluation and revision component separately as a way to support more effective peer review and develop self-evaluation. It might be easier for instructors to integrate this specific practice into their courses than to use the entire approach.

Another direction for future research is to investigate how the instructional approach might be used in alternative approaches to delivering developmental education such as accelerated learning programs (Adams, Gearhart, Miller, & Roberts, 2009) and stretch programs (Glau, 2007). These approaches are spreading, and some evidence indicates that they improve retention. However, the approaches do not address pedagogical methods directly.

Closing Thoughts

The positive findings from this research add to the body of research on SRSD, expanding the range of its application to underprepared college writers. Developmental education represents a large and growing challenge for educators, policymakers,

and researchers. The current study offers promise that effective instruction can help developmental writers to increase their knowledge about writing strategies; to develop self-regulation strategies for setting goals, managing their efforts, and reflecting on their progress; and, thereby, to increase their writing achievement. Overall, a great deal of research is needed to design and evaluate evidence-based instructional practices to meet the needs of this diverse population.

References

- ACT, Inc. (2012). *ACT national curriculum survey 2012: English language arts*. Iowa City, IA: Author.
- Adams, P., Gearhart, S., Miller, R., & Roberts, A. (2009). The Accelerated Learning Program: Throwing open the gates. *Journal of Basic Writing*, 28, 50–69.
- Attewell, P., Lavin, D., Domina, T., & Levey, T. (2006). New evidence on college remediation. *The Journal of Higher Education*, 77, 886–924. <http://dx.doi.org/10.1353/jhe.2006.0037>
- Bailey, T., Jeong, D. W., & Cho, S.-W. (2010). Referral, enrollment, and completion in developmental education sequences in community colleges. *Economics of Education Review*, 29, 255–270. <http://dx.doi.org/10.1016/j.econedurev.2009.09.002>
- Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York, NY: Freeman.
- Berry, A. B., & Mason, L. H. (2012). The effects of self-regulated strategy development on the writing of expository essays for adults with written expression difficulties: Preparing for the GED. *Remedial and Special Education*, 33, 124–136. <http://dx.doi.org/10.1177/0741932510375469>
- Bradley-Johnson, S., Morgan, S. K., & Nutkins, C. (2004). Review of the Woodcock-Johnson tests of achievement: Third edition. *Journal of Psychoeducational Assessment*, 22, 261–274.
- Bremer, C. D., Center, B. A., Opsal, C. L., Medhanie, A., Jang, Y. J., & Geise, A. C. (2013). Outcome trajectories of developmental students in community colleges. *Community College Review*, 41, 154–175. <http://dx.doi.org/10.1177/0091552113484963>
- Bruning, R., & Kauffman, D. F. (in press). Self-efficacy beliefs and motivation in writing development. In C. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (2nd ed., pp. 96–114). New York, NY: Guilford Press.
- Bryk, A. S., & Raudenbush, S. W. (1992). *Hierarchical linear models: Applications and data analysis methods*. Newbury Park, CA: Sage.
- Butler, D. L. (2003). Structuring instruction to promote self-regulated learning by adolescents and adults with learning disabilities. *Exceptionality*, 11, 39–60. http://dx.doi.org/10.1207/S15327035EX1101_4
- Cho, K., & MacArthur, C. (2011). Learning by reviewing. *Journal of Educational Psychology*, 103, 73–84. <http://dx.doi.org/10.1037/a0021950>
- Cox, R. D. (2009). “It was just that I was afraid”: Promoting success by addressing students’ fear of failure. *Community College Review*, 37, 52–80.
- Elliot, A. J., & Church, M. A. (1997). A hierarchical model of approach and avoidance achievement motivation. *Journal of Personality and Social Psychology*, 72, 218–232. <http://dx.doi.org/10.1037/0022-3514.72.1.218>
- Englert, C. S., Raphael, T. E., Anderson, L. M., Anthony, H. M., & Stevens, D. D. (1991). Making writing strategies and self-talk visible: Cognitive strategy instruction in writing in regular and special education classrooms. *American Educational Research Journal*, 28, 337–372. <http://dx.doi.org/10.3102/00028312028002337>
- Glau, G. R. (2007). Stretch at 10: A progress report on Arizona State University’s Stretch Program. *Journal of Basic Writing*, 26, 30–48.
- Graham, S., Harris, K. R., & Hebert, M. (2011). It is more than just the message: Presentation effects in scoring writing. *Focus on Exceptional Children*, 44, 1–13.
- Graham, S., Harris, K. J., & Mason, L. (2005). Improving the writing performance, knowledge, and self-efficacy of struggling young writers: The effects of self-regulated strategy development. *Contemporary Educational Psychology*, 30, 207–241. <http://dx.doi.org/10.1016/j.cedpsych.2004.08.001>
- Graham, S., Harris, K. R., & McKeown, D. (2013). The writing of students with learning disabilities: Meta-analysis of self-regulated strategy development writing intervention studies and future directions: Redux. In H. L. Swanson, K. Harris, & S. Graham (Eds.), *Handbook of learning disabilities* (2nd ed., pp. 565–590). New York, NY: Guilford Press.
- Graham, S., McKeown, D., Kihara, S., & Harris, K. R. (2012). A meta-analysis of writing instruction for students in the elementary grades. *Journal of Educational Psychology*, 104, 879–896. <http://dx.doi.org/10.1037/a0029185>
- Graham, S., & Perin, D. (2007). A meta-analysis of writing instruction for adolescent students. *Journal of Educational Psychology*, 99, 445–476. <http://dx.doi.org/10.1037/0022-0663.99.3.445>
- Gribbin, B. (2005). Our ambivalence toward teaching grammar. *English Journal*, 94, 17–23. <http://dx.doi.org/10.2307/30046408>
- Harris, K. J., & Graham, S. (2009). Self-regulated strategy development in writing: Premises, evolution, and the future. In V. Connelly, A. L. Barnett, J. E. Dockrell, & A. Tolmie (Eds.), *Teaching and learning writing (British Journal of Educational Psychology Monograph Series II: Pt. 6, pp. 113–135)*. Leicester, United Kingdom: British Psychological Society.
- Harris, K. R., Graham, S., MacArthur, C. A., Reid, R., & Mason, L. (2011). Self-regulated learning processes and children’s writing. In B. Zimmerman & D. H. Schunk (Eds.), *Handbook of self-regulation of learning and performance* (pp. 187–202). New York, NY: Routledge.
- Harris, K. R., Lane, K. L., Graham, S., Driscoll, S. A., Sandmel, K., Brindle, M., & Schatschneider, C. (2012). Practice-based professional development for self-regulated strategies development in writing: A randomized controlled study. *Journal of Teacher Education*, 63, 103–119. <http://dx.doi.org/10.1177/0022487111429005>
- Haswell, R. H. (2005). NCTE/CCCC’s Recent War on Scholarship. *Written Communication*, 22, 198–223. <http://dx.doi.org/10.1177/0741088305275367>
- Hayes, J. R. (1996). A new framework for understanding cognition and affect in writing. In C. M. Levy & S. Ransdell (Eds.), *The science of writing* (pp. 1–27). Mahwah, NJ: Erlbaum.
- Hayes, J., & Flower, L. (1980). Identifying the organization of writing processes. In L. W. Gregg & E. R. Steinberg (Eds.), *Cognitive processes in writing* (pp. 3–30). Hillsdale, NJ: Erlbaum.
- Hidi, S., & Boscolo, P. (2006). Motivation and writing. In C. A. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (pp. 144–157). New York, NY: Guilford Press.
- Hillocks, G. (1986). Research on written composition: New directions for teaching. Urbana, IL: National Conference on Research in English/ERIC Clearinghouse on Reading and Communication Skills.
- Hunt, K. W. (1964). *Differences in grammatical structures written at three grade levels, the structures to be analyzed by transformational methods, Report# CRP-1998*. Tallahassee, FL: Florida State University
- Kauffman, D. F., Zhao, R., Dempsey, M. S., Zeleny, M. G., Wang, S., & Bruning, R. H. (2010). *Achievement goals’ impact on 11th grade students’ affect, self-efficacy, and writing achievement*. Paper presented at the American Educational Research Association annual meeting.
- Krasa, N. (2007). Is the Woodcock-Johnson III a test for all seasons? Ceiling and item gradient considerations in its use with older students. *Journal of Psychoeducational Assessment*, 25, 3–16. <http://dx.doi.org/10.1177/0734282906291768>

- Levin, H. M., & Calcagno, J. C. (2008). Remediation in the community college: An evaluator's perspective. *Community College Review*, 35, 181–207. <http://dx.doi.org/10.1177/0091552107310118>
- MacArthur, C. A. (2011). Strategies instruction. In K. R. Harris, S. Graham, & T. Urdu (Eds.), *Educational psychology handbook: Vol. 3. Applications of educational psychology to learning and teaching* (pp. 379–401). Washington, DC: American Psychological Association.
- MacArthur, C. A. (2012). Evaluation and revision processes in writing. In V. W. Berninger (Ed.), *Past, Present, and Future Contributions of Cognitive Writing Research to Cognitive Psychology* (pp. 461–483). London: Psychology Press.
- MacArthur, C. A., & Lembo, L. (2009). Strategy instruction in writing for adult literacy learners. *Reading and Writing*, 22, 1021–1039. <http://dx.doi.org/10.1007/s11145-008-9142-x>
- MacArthur, C. A., & Philippakos, Z. A. (2012). Strategy instruction with college basic writers: A design study. In C. Gelati, B. Arfé, & L. Mason (Eds.), *Issues in writing research* (pp. 87–106). Padova: CLEUP.
- MacArthur, C. A., & Philippakos, Z. A. (2013). Self-regulated strategy instruction in developmental writing: A design research project. *Community College Review*, 41, 176–195. <http://dx.doi.org/10.1177/0091552113484580>
- Milewski, G. B., Johnson, D., Glazer, N., & Kubota, M. (2005). *A survey to evaluate the alignment of the new SAT Writing and Critical Reading sections to curricula and instructional practices*. New York, NY: College Entrance Examination Board.
- Myhill, D., & Jones, S. (2007). More than just error correction. Students' perspectives on their revision processes during writing. *Written Communication*, 24, 323–343. <http://dx.doi.org/10.1177/0741088307305976>
- National Center for Education Statistics. (2012). *The nation's report card: Writing 2011 (NCES 2012–470)*. Washington, DC: NCES, Institute of Education Sciences, U. S. Department of Education.
- National Center for Education Statistics. (2013a). *Digest of education statistics: 2012*. Washington, DC: NCES, Institute for Education Sciences, U. S. Department of Education. Retrieved from http://nces.ed.gov/programs/digest/d12/tables/dt12_236.asp
- National Center for Education Statistics. (2013b). *2011–12 National Post-secondary Student Aid Study (NPSAS:12)*. Washington, DC: NCES, Institute for Education Sciences, U. S. Department of Education. Retrieved from <http://nces.ed.gov/datalab/tableslibrary/viewtable.aspx?tableid=9420>
- National Research Council. (2012). *Improving adult literacy instruction: Options for practice and research*. Washington, DC: The National Academies Press.
- Nussbaum, E. M., & Kardash, C. M. (2005). The effects of goal instructions and text on the generation of counterarguments during writing. *Journal of Educational Psychology*, 97, 157–169. <http://dx.doi.org/10.1037/0022-0663.97.2.157>
- Pajares, F., & Cheong, Y. F. (2003). Achievement goal orientations in writing: A developmental perspective. *International Journal of Educational Research*, 39, 437–455. <http://dx.doi.org/10.1016/j.ijer.2004.06.008>
- Pajares, F., & Valiante, G. (2006). Self-efficacy beliefs and motivation in writing development. In C. A. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (pp. 158–170). New York, NY: Guilford Press.
- Perin, D. (2013). Literacy skills among academically underprepared students. *Community College Review*, 41, 118–136. <http://dx.doi.org/10.1177/0091552113484057>
- Philippakos, Z. A. (2012). *Effects of reviewing on fourth and fifth-grade students' persuasive writing and revising* (Doctoral dissertation, University of Delaware, 2012). ProQuest Dissertations and Theses, 260.
- Philippakos, Z. A., MacArthur, C. A., & Uribe-Zarain, X. (2013). *Writing motivation: Validation of a measure for college writers*. Paper presented at the Annual Conference of the Literacy Research Association, Houston, TX.
- Prior, P. (2006). A sociocultural theory of writing. In C. A. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (pp. 54–66). New York, NY: Guilford Press.
- 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. <http://dx.doi.org/10.1037/0022-0663.97.1.43>
- Salahu-Din, D., Persky, H., & Miller, J. (2008). *The nation's report card: Writing 2007 (NCES 2008–468)*. Washington, DC: National Center for Education Statistics, Institute of Education Sciences, U. S. Department of Education.
- Schunk, D. H., & Swartz, C. W. (1993). Goals and progress feedback: Effects on self-efficacy and writing achievement. *Contemporary Educational Psychology*, 18, 337–354. <http://dx.doi.org/10.1006/ceps.1993.1024>
- Schunk, D. H., & Zimmerman, B. J. (2007). Influencing children's self-efficacy and self-regulation of reading and writing through modeling. *Reading & Writing Quarterly: Overcoming Learning Difficulties*, 23, 7–25. <http://dx.doi.org/10.1080/10573560600837578>
- Torrance, M., van Waes, L., & Galbraith, D. (Eds.). (2007). *Writing and cognition: Research and applications, studies in writing* (Vol. 20). Amsterdam: Elsevier. [http://dx.doi.org/10.1108/S1572-6304\(2007\)0000020002](http://dx.doi.org/10.1108/S1572-6304(2007)0000020002)
- White, M. J., & Bruning, R. (2005). Implicit writing beliefs and their relation to writing quality. *Contemporary Educational Psychology*, 30, 166–189. <http://dx.doi.org/10.1016/j.cedpsych.2004.07.002>
- Wolfe, C. R. (2011). Argumentation across the curriculum. *Written Communication*, 28, 193–219. <http://dx.doi.org/10.1177/0741088311399236>
- Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). *Woodcock-Johnson III: Tests of achievement*. Itasca, IL: Riverside.
- Zimmerman, B. J. (1994). Dimensions of academic self-regulation: A conceptual framework for education. In B. J. Zimmerman & D. H. Schunk (Eds.), *Self-regulation of learning and performance: Issues and educational applications* (pp. 3–21). Hillsdale, NJ: Erlbaum.
- Zimmerman, B. J., & Kitsantas, A. (2002). Acquiring writing revision and self-regulatory skill through observation and emulation. *Journal of Educational Psychology*, 94, 660–668. <http://dx.doi.org/10.1037/0022-0663.94.4.660>

Received April 9, 2014

Revision received August 21, 2014

Accepted August 22, 2014 ■