The published version can be found in: Limpo, T., & Alves, R. A. (2013). Teaching planning or sentence-combining strategies: Effective SRSD interventions at different levels of written composition. *Contemporary Educational Psychology, 38*, 328-341. doi: 10.1016/j.cedpsych.2013.07.004

Teaching Planning or Sentence-Combining Strategies: Effective SRSD Interventions at Different Levels of Written Composition

Teresa Limpo[a] and Rui A. Alves[a]

[a] University of Porto, Faculty of Psychology and Educational Sciences, Rua Alfredo Allen, 4200-392 Porto, Portugal

Author Note

This research was made possible by funds from the COST Action IS0703 – The European Research Network on Learning to Write Effectively (STSM-IS0703-3459) and from the Portuguese Foundation for Science and Technology (grants SFRH/BD/68548/2010 and PTDC/PSI-PCO/110708/2009 attributed to the first and second authors, respectively).

Correspondence concerning this article should be addressed to Teresa Limpo, University of Porto, Faculty of Psychology and Educational Sciences, Rua Alfredo Allen, 4200-392 Porto, Portugal. Tel.: +351 200 400 610. E-mail address: tlimpo@fpce.up.pt

Abstract

This study tested the effectiveness of two strategy-focused interventions aimed at promoting fifth and sixth graders' opinion essay writing. Over 12 weekly 90-min lessons, two groups of 48 and 39 students received, respectively, planning and sentence-combining instruction, which followed the Self-Regulated Strategy Development (SRSD) model. These intervention groups were compared with a practice control group of 39 students receiving standard writing instruction. The following main findings were noteworthy: (a) planning and sentence-combining instruction enhanced planning and sentence-construction skills, respectively; (b) both interventions increased opinion essay quality and text length; (c) planning instruction enhanced not only discourse-level writing but also some sentence- and word-level aspects of composition; (d) sentence-combining instruction enhanced not only sentence- and word-level writing but also some discourse-level aspects of composition; (e) after instruction, there was a correlation between self-efficacy and writing quality in both intervention groups; and (f) planning, but not sentence-combining, instructional effects generalized to summary writing.

Keywords: writing, strategy instruction, planning, sentence combining, self-regulated strategy development

Teaching Planning or Sentence-Combining Strategies: Effective SRSD Interventions at Different Levels of Written Composition

Given the importance of writing in present-day knowledge societies, it is of the utmost importance to develop evidence-based practices that promote students' writing performance (Graham, Gillespie, & McKeown, 2012). Based on multiple sources of information, Graham and Harris (2009) proposed that teaching planning and sentence-combining strategies are effective instructional practices in writing instruction. Nevertheless, little is known about the specific effects and comparative merits of planning and sentence-combining instruction. This was the main purpose of the present study, which examined the effectiveness of two strategy-focused interventions aimed at improving opinion essay writing in Grades 5-6 (age 10-12). One of the programs taught students a strategy to plan opinion essays, while the other program taught them a strategy to combine sentences. Self-regulation procedures to manage these writing strategies were embedded in both programs following the Self-Regulated Strategy Development (SRSD) model (Harris & Graham, 1996, 2009), which was adapted to the Portuguese culture and school system.

1.1. Writing Development and Strategy Instruction

The development of expertise in writing is a long and demanding process, heavily dependent upon changes in four areas: skills, strategies, knowledge, and motivation (Alexander, Graham, & Harris, 1998; Graham, 2006a). To produce a text, writers need sentence generation skills to transform their ideas into language representations, as well as transcription skills (i.e., spelling and handwriting) to transform those representations into written text (Berninger et al., 1992). Given the complexity of the writing process, text production also demands strategic processing (Pressley & Harris, 2006). Writers must rely on a repertoire of strategies to accomplish key cognitive processes (e.g., planning or revising) and to self-regulate their use. Moreover, writers need to access their long-term memory to retrieve content knowledge as well as discourse knowledge, which includes knowledge about intended audience, different genres, tasks schemas, and linguistic conventions (McCutchen, 1986, 2011). Another critical ingredient in writing

achievement is motivation, which comprises writers' will to engage in writing along with their own representations as writers (Boscolo & Hidi, 2007). In particular, self-efficacy seems to be one of the strongest predictors of writing performance (Pajares, 2003).

Despite the importance of these four components in writing development, they are not fully operational in developing writers. In Grades 4-6 (age 9-12), transcription skills represent a strong constraint to writing, indicating that they are not completely automatic (Graham, Berninger, Abbott, Abbott, & Whitaker, 1997; Limpo & Alves, 2013). Because of the heavy demands of these skills, students may have few attentional resources for self-regulatory processes (McCutchen, 1996). Bereiter and Scardamalia (1987) have proposed that youngsters cope with these challenges by engaging in knowledge telling: They compose text by retrieving content that is immediately written down. Their writing is mostly driven by what they know about the topic and how this can be fitted within genre constraints. This presumably effortless way of composing reflects poor strategic competence to deliberately activate and articulate key cognitive processes. Indeed, in Grades 4-6, Limpo and Alves (2013) found that planning and revising tend to play a minimal role not only in writing performance, but also in self-efficacy judgments. On the contrary, it was students' transcription ability that strongly influenced both their performance and self-efficacy in writing. Besides developing writers' skills, strategic behavior, knowledge, and motivation seem to increase with age and schooling, these components can be enhanced through appropriate instruction (Graham, 2006a).

Strategy instruction is a type of cognitive-oriented writing intervention, which aims to enhance conscious, goal-directed, and effortful processing in writing (Pressley & Harris, 2006). For that purpose, strategy instruction provides students with explicit and systematic teaching of strategies to accomplish writing-specific processes, such as planning, sentence generation, and revising. These strategies contain the procedural or the "how to" knowledge for carrying out a particular writing process. Usually, one way of promoting the memorization of strategy steps is the teaching of mnemonics (e.g., the PLANS strategy tells students that to plan a text they need to:

*P*ick goals, *L*ist ways to meet goals, *A*nd, make *N*otes, and *S*equence notes; Harris, Graham, Mason, & Friedlander, 2008). A common feature of different models of strategy instruction is the inclusion of self-regulation components, which are critical for facilitating strategic behaviors in writing (Alexander et al., 1998). These components promote not only the acquisition of the learned strategies but also their maintenance and generalization. Several meta-analyses indicated that strategy instruction is one the best teaching practices to promote writing quality in Grades 2–12 (Graham, McKeown, Kiuhara, & Harris, 2012; Graham & Perin, 2007; Rogers & Graham, 2008). In particular, the SRSD seems to be the most effective strategy instruction model, as its average effect size doubles that of the other ones.

The asset of SRSD over other approaches to strategy instruction is that it was designed to address the four critical ingredients of writing development: skills, strategies, knowledge, and motivation (Harris & Graham, 1996, 2009). Across six instructional stages (develop background knowledge; discuss it; model it; memorize it; support it; and independent performance) SRSD provides explicit teaching of writing strategies combined with self-regulation procedures to regulate strategies usage and writing behavior. In a meta-analytic review of true- and quasi-experimental studies examining SRSD effectiveness, Harris, Graham, Brindle, and Sandmel (2009) (see also Graham, 2006b; Graham & Harris, 2003) reported average effect sizes of 1.20 and 1.23 for writing quality at posttest (n = 15) and maintenance (n = 9), and 1.20 for generalization to untaught genres (n = 5). Besides writing quality, meaningful and lasting effects of SRSD had also been reported for other aspects of writing, such as, schematic structure (e.g., inclusion of genre-specific elements), approach to writing (e.g., time spent planning and writing), discourse knowledge, and self-efficacy beliefs. It is noteworthy that these improvements were observed across achievement level, grade level, cognitive process taught, target genre, and type of instructor (teacher or researcher).

1.2. Why Teach Strategies for Planning and Sentence Generation?

Planning, which is the generation and organization of ideas along with the formulation of goals for the task, is a critical ingredient in skilled writing (Hayes & Flower, 1980). Beauvais, Olive, and Passerault (2011) found that the longer the prewriting pause and planning time of undergraduates, the better the quality of their argumentative texts. However, beginning and developing writers barely engage in preplanning activities (Bereiter & Scardamalia, 1987; McCutchen, 2006). For instance, Torrance, Fidalgo, and García (2007) showed that, before writing, only 16% of sixth graders engaged in preplanning. Furthermore, during writing, they only spent 11% of their writing time thinking about content. In line with these findings, Limpo and Alves (2013) found that planning complexity did not contribute to writing quality in Grades 4-6 (age 9-12), but it did in Grades 7-9 (age 12-15). The authors suggested that, albeit being able to plan upon request, younger students lacked sufficient planning abilities to create a plan that could aid them to compose the text. Nonetheless, there is strong evidence that teaching strategies to plan ahead of writing is an effective way to promote young students' writing performance (Glaser & Brunstein, 2007; Graham, Harris, & Mason, 2005; Harris, Graham, & Mason, 2006; Torrance et al., 2007; Wong, Hoskyn, Jai, Ellis, & Watson, 2008). To enhance novice writers' planning abilities is critical because preplanning may help them to generate content and to create an organized structure for their compositions. In addition, the plan may function as an external memory where children store ideas to include in the text and outline action-plans to produce it (Graham & Harris, 2007; but see Kellogg, 1988). Consequently, the planning time during writing might be reduced, enabling students to focus on other writing processes, such as translation (Kellogg, 1988).

Translation refers to a foundational writing process through which writers transform their thoughts into written language (Hayes & Flower, 1980). The ability to construct syntactically correct and complex sentences is, therefore, a critical translating skill that characterizes expert writing (Beers & Nagy, 2009; Berninger, Nagy, & Beers, 2011). Improving young writers' sentence-construction skills is likely to boost writing expertise in several ways (Saddler, 2007;

Saddler & Asaro, 2008). The acquisition of fluency in sentence construction may free up attentional resources, so that students can concentrate on other aspects of composing, such as planning (Fayol, 1999; McCutchen, Covill, Hoyne, & Mildes, 1994). Moreover, students' with superior sentence-construction skills may have access to an enlarged syntactical repertoire for creating sentences, which not only facilitates translation (Hayes & Flower, 1986) but also set the stage for revision (Saddler & Graham, 2005). Additionally, the use of well-crafted, syntactically correct sentences may result in interesting and readable texts. An instructional method that provides direct practice with sentence-construction skills is sentence-combining (Saddler, 2007; Strong, 1986, 1996). Through sentence-combining exercises, students learn to transform basic sentences, such as "I think we should have homework" and "Homework helps us to study" into a more syntactically complex sentence, such as "I think we should have homework because it helps us to study". These exercises require that students manipulate sentences to create new syntactic structures preserving the original meaning. Evidence has been accumulated on the effectiveness of this approach to enhance syntactic complexity and overall compositional quality (Saddler, Behforooz, & Asaro, 2008; Saddler & Graham, 2005; for reviews and meta-analyses, see also Andrews et al., 2006; Graham & Perin, 2007; Hillocks, 1986).

1.3. Why Teach Self-Regulation Procedures?

The complexity and cognitive demands of the composing process explains why skilled writing requires high levels of self-regulation (Graham & Harris, 2000). Self-regulation refers to the degree to which students are metacognitively, motivationally, and behaviorally strategic regulators of their own writing process (Zimmerman, 1995). Besides writers can use several self-regulation procedures (for a comprehensive list, see Zimmerman & Risemberg, 1997), particularly, goal setting, self-monitoring, self-reinforcement, and self-instructions seem to be important (Bandura, 1969; Harris et al., 2008).

As writing is a goal-directed activity (Hayes & Flower, 1986), goal setting is a critical component of effective text production by serving a self-regulatory function (Locke, Shaw, Saari,

& Latham, 1981; Zimmerman, 2009). The formulation of specific, proximal, and challenging goals provides clear information about task requirements and directs attention towards them. Also, it motivates the use of strategies to achieve goals, mobilizes effort, and increases persistence. Goal-setting procedures seem to be very effective in promoting writing quality (Graham & Perin, 2007), especially, when coupled with progress feedback (Schunk & Swartz, 1993). Students can track progress in achieving goals on their own through self-monitoring. First, they determine whether or not a target behavior has occurred and, then, they register the results (Nelson & Hayes, 1981). This procedure allows students to become aware of and responsible for their behavior (Mace, Belfiore, & Hutchinson, 2001; Rankin & Reid, 1995). Furthermore, it conveys to students that they are capable of meeting their goals, which may increase feelings of self-satisfaction and trigger self-administration of rewarding consequences (Bandura, 1969; Zimmerman, 2000). Self-consequences, such as self-reinforcement, enhance motivation and efforts to improve. Importantly, students can instruct themselves to set goals, self-monitor performance, and self-reinforce success by means of explicit self-instructions (Meichenbaum, 1977; Schunk, 2001). This form of self-speech enables writers to support and guide their behavior throughout the writing process.

The development of these four self-regulation procedures via explicit teaching and support seems to be one of the key features of SRSD. In a well-designed study, Glaser and Brunstein (2007) showed that teaching writing strategies combined with self-regulation practices was more effective than teaching writing strategies alone (see also Brunstein & Glaser, 2011).

1.4. The Present Study

The primary purpose of this study was to examine and compare the effectiveness of planning and sentence-combining interventions in Grades 5-6. Following the SRSD model, both programs taught a writing strategy in combination with self-regulation procedures (viz., goal-setting, self-monitoring, self-reinforcement, and self-instructions). While one of the programs taught a strategy for planning opinion essays, the other one taught a strategy for combining sentences in opinion essay writing. The two interventions only differed on the taught writing

strategy and everything else was held constant (e.g., program and lessons structure, self-regulation and instructional procedures, and number of writing tasks, including writing prompts). Meta-analyses findings showed that planning instruction lead to greater effect sizes than sentence-combining instruction (Graham & Perin, 2007; Rogers & Graham, 2008). Nevertheless, some caution is needed when interpreting these results. First, planning instruction frequently includes the teaching of self-regulation procedures, which is not the case of sentence-combining instruction. Second, interventions effects are usually compared regarding general, rather than specific, outcomes (e.g., writing quality). The present study aims to provide a more direct and controlled comparison of planning and sentence-combining interventions, which to the best of our knowledge has not been made.

We used the SRSD model in both planning and sentence-combining instruction, which allowed a fair comparison between the two interventions. Notably, this is the first study to couple SRSD with sentence-combining instruction. The development and evaluation of such an integrated instructional program is an important contribution to move forward both the SRSD and the sentence-combining literature. Given the value added impact of this model to writing instruction, the testing of how it can be effectively used to teach different skills in writing is a matter that still needs investigation. Although there is strong evidence that SRSD is effective to enhance planning skills, it has not been used to teach sentence-construction skills (Harris & Graham, 2009). Besides, there are only a few studies evaluating sentence-combining practices and none have taught self-regulation procedures to aid students in the production of sentences. As research has been demonstrating the effectiveness of both SRSD and sentence combining, the integration of these two forms of instruction seems a potentially sound instructional practice.

Furthermore, we have compared interventions effectiveness at three levels of written language (viz., discourse, sentence, and word). While planning instruction mainly taps discourse-level writing, sentence-combining instruction mainly taps sentence- and word-level writing. It is possible that holistic measures, such as overall quality, are not sufficiently sensitive to

discriminate instructional effects between these levels. The present study aims to provide a detailed comparison of the two interventions by testing specific effects on the target level and transfer effects across levels. Such a fine-grained analysis may provide helpful information to closely align writing instruction with students' writing needs.

1.5. Hypotheses

Before, in the middle (after Lesson 5), and after instruction (after Lesson 12), students planned and wrote an opinion essay, and completed a sentence-combining exercise. Intervention effects on writing were assessed on strategy-specific skills, writing performance, and levels of writing. Instructional effects on these three sets of measures are described next and were expected to emerge at posttest. As students did not write complete texts until midtest, only strategy-specific effects were expected there.

Planning and sentence-combining students were expected to improve the strategy-specific skills that were explicitly taught (viz., planning or sentence combining). Moreover, they were expected to write better and longer opinion essays than control students. These predictions stem from several studies indicating that strategy instruction combined with SRSD procedures is a highly effective practice in increasing the effectiveness of strategy usage, and the quality and length of texts produced (for reviews, see Graham, 2006b; Graham & Harris, 2003; Harris et al., 2009). In the present study, besides the systematic instruction in key writing processes, students received explicit teaching and scaffolded practice in a set of self-regulation procedures. By facilitating the strategic processing characteristic of skilled writers, both SRSD interventions should lead to meaningful gains in the quality and quantity of students' writing.

To provide a more fine-grained analysis of interventions effectiveness, students' opinion essays were analyzed at three levels of written language (Wagner et al., 2011; Whitaker, Berninger, Johnston, & Swanson, 1994). At the discourse level, we analyzed the inclusion of functional essay elements and coherence. At the sentence level, we analyzed the use of cohesion devices (Favart, 2005), and syntactic complexity via clause length (Beers & Nagy, 2009 showed

that clause rather than T-unit length was related to opinion essay quality). At the word level, we analyzed vocabulary diversity and use of modifiers. The two strategy interventions were expected to have different effects on these levels. On the one hand, we predicted that planning students would outperform their peers at the discourse level because they learnt a strategy to generate and organize ideas in a complete and coherent essay. Reviews of intervention studies have shown that this type of instruction increases the number of genre-specific elements included in texts (Graham, 2006b; Graham & Harris, 2003; Harris et al., 2009). On the other hand, we expected that sentence-combining students would outperform their peers at the sentence and word levels because they learnt a strategy to write syntactically complex sentences with cohesion devices, diverse vocabulary, and modifiers. After sentence-combining instruction, non-SRSD studies have reported increases not only in the number of connectives used in writing (Saddler et al., 2008) but also in the number of sentences combined through revision (Saddler & Graham, 2005). By incorporating sentence combining into the SRSD framework, we expected extensive improvements in students' writing at the sentence- and word-levels.

To examine instructional effects on motivation, at pretest and posttest, students filled out a self-efficacy scale. Even though some SRSD interventions failed to increase writing self-efficacy (Graham et al., 2005; Page-Voth & Graham, 1999; Sawyer, Graham, & Harris, 1992; but see Brunstein & Glaser, 2011), the use of self-regulation strategies is thought to increase students' beliefs about their capabilities (Pajares, 2003; Schunk, 2003; Zimmerman, 2000; Zimmerman & Risemberg, 1997). In the present study, the combination of goal setting with self-monitoring plus self-reinforcement was used to strength the link between strategy usage and enhanced performance, as well as to highlight students' progress over instruction. This should lead to an increase in their perceptions as competent writers. On this ground, we predicted that both SRSD interventions would increase students' self-efficacy for writing.

Finally, to examine generalization effects of the interventions, students summarized a text before and after instruction. The generalization measure of summary writing was chosen because it

enabled us to assess the transfer of strategies' core principles rather than the use of memorized routines (Shepard, 2000). Strategies transfer from opinion essay to summary writing should be facilitated by the SRSD. This model provides students critical self-regulatory tools to flexibly apply the learned knowledge and strategies to other writing tasks (Harris & Graham, 2009; Harris et al., 2009). We predicted that the teaching of a planning strategy to select and organize self-generated information would increase these students' ability to select relevant information to include in their summaries. We further anticipated that the teaching of a sentence-combining strategy would increase students' propensity to condense different ideas from the original text into single sentences in the summary.

2. Method

2.1. Participants and Design

Participants were 146 Portuguese native speakers in Grades 5-6 (3 classes per grade) from a public school located in an urban district in Northwest Portugal. Twenty students were excluded from the analyses based on one or more of the following criteria: absence in a testing session (12 students), absence in two or more lessons (4 students), and special education needs (5 students). Subsequent analyses were based on the data from 126 students. The study involved a pretest, midtest, posttest quasi-experimental design with three conditions: planning, sentence combining, and control. Within each grade level, each intact class was randomly assigned to one of the three conditions. Demographic data by condition is provided on Table 1.

Table 1 about here

2.2. Intervention Conditions

Students' Portuguese language teacher delivered writing instruction. Two teachers implemented each intervention program. The four female teachers (*M* age = 50.3 years, *SD* = 8.5) had a teaching experience of more than 16 years (*M* = 22.5 years, *SD* = 6.0). Writing instruction occurred during Portuguese language classes, in 12 weekly lessons of 90 min. In line with the

SRSD model (Harris & Graham, 1996, 2009), students were taught a writing strategy in tandem with self-regulation procedures. The following practices were used: development of knowledge for writing and self-regulation; explicit instruction, discussion, and modeling of the target strategy and self-regulation procedures; promotion of the mnemonic and self-instructions memorization; collaborative practice supported by teachers and guidance materials that were gradually faded; independent practice with teacher monitoring and support when needed.

2.2.1. Writing strategies. Both intervention groups learned a writing strategy, along with the necessary skills and knowledge to properly use it. In the planning condition, students were taught a strategy to plan opinion essays. They learned the mnemonic CRÊS, which stands for: tell what you believe, give 3 or more reasons, explain each reason, and wrap it up. This is the Portuguese adaptation of the mnemonic TREE (Harris et al., 2008). The strategy was practiced in isolation during the first five lessons. Then, it was embedded in text production.

In the sentence-combining condition, students were taught a strategy to combine sentences. We developed the mnemonic DICA that is the Portuguese acronym to: what do you want to say?, what is the idea (addition, contrast, or cause)?, choose the best connective, and enrich with adjectives and adverbs. During the first five lessons students performed sentence-combining exercises, which evolved from highly cued to uncued. Then, they started writing opinion essays. In accordance with Strong (1986), three procedures were used to promote transfer of sentence-combining skills from exercises to composition. First, we used whole-discourse exercises, that is, the sentences within an exercise formed an opinion essay. Second, students learned opinion markers to help them to organize the sentences within the essay. Third, students were provided explicit instruction and guided practice to integrate DICA into composition.

2.2.2. Self-regulation procedures. These procedures were introduced in a stepwise fashion and similarly across intervention conditions. *Goal setting* was introduced in Lesson 1. Students had a general goal (viz., to write good opinion essays) plus strategy-specific goals. Planning students had to write complete opinion essays, while sentence-combining students had to

write well-crafted sentences with connectives, opinion markers, and adjectives/adverbs. *Self-monitoring* was introduced in Lesson 2. Students were given a "self-monitoring sheet" where they: (a) set the goal for the task, (b) registered and counted the number of essay parts (planning condition), or the number of connectives, opinion markers, and adjectives/adverbs (sentence-combining condition), and (c) wrote a self-reinforcement statement. They were also given a "progress sheet" to register and track their performance during instruction. *Self-instructions* were introduced in Lesson 3. Using a "writing flowchart" students developed self-instructions to set goals, use the strategy, and check if goals were met.

2.2.3. Lessons summary. In Lesson 1, students discussed the importance of planning or constructing well-crafted sentences, respectively, in planning and sentence-combining condition. The target strategy was introduced in Lesson 2. Using the self-monitoring and progress sheets, planning and sentence-combining students had to find and register the essay parts or connectives included in an exemplar opinion essay, and, then, in their own pretest essays. From this session on, they filled out the progress sheet anytime they worked individually.

Lessons 3-5 involved modeling of strategy implementation followed by collaborative and independent practice. In Lesson 3, teachers modeled how to use the strategy and, after discussing it, students came up with their self-instructions. In Lesson 4, the class emulated the teacher modeling and repeated it at home. In Lesson 5, they applied the writing strategy, individually.

Lessons 6-8 involved modeling of writing strategy integration within text production followed by collaborative and independent practice. In Lesson 6, the teacher modeled how to use the strategy in writing. Then, students came up with updated self-instructions. In Lesson 7, the class emulated the teacher modeling. In addition, teachers prompted strategies transfer to different situations. In Lesson 8, students wrote their first opinion essay, individually.

In Lesson 9, each teacher grouped students that faced similar difficulties and gave them individualized feedback. Lessons 10 and 11 involved independent practice in opinion essay writing with minimal support. In Lesson 12, students examined and discussed their progress sheet.

2.2.4. Treatment fidelity. Five procedures guaranteed that both interventions were delivered as intended. First, teachers participated in an 8-hr pre-intervention workshop, in which they were introduced to the writing strategies and self-regulation procedures to be taught. They also received the instructional manuals and discussed lessons' procedures. Second, teachers had weekly meetings with the first author to practice the next lesson and to discuss the previous one. Deviations from instructional plans were rare and usually involved missed steps. Third, teachers were provided with a checklist with all lessons' steps and they were asked to check them off when completed. Teachers of both strategy conditions completed 98% of the proposed steps. Whenever possible, missed steps were addressed in the next lesson. Fourth, the first author observed one third of the lessons and filled out the same checklist as the teachers. Planning and sentence-combining teachers completed 97% and 96% of the proposed steps, respectively. Fifth, the quality of these observed lessons was evaluated on five items: (a) level of students' engagement, (b) students' responses to questions and participation in discussion, (c) teachers' responses to students' questions, (d) efficiency of instruction, and (e) pacing of instruction (based on Saddler & Graham, 2005). The average quality for planning and sentence-combining instruction was 3.7 and 3.6, respectively (0 = *very low*; 4 = *very high*).

2.3. Control Condition

Writing instruction of control students followed the standard writing curriculum and was delivered by their regular Portuguese language teachers, not implementing the interventions. After instruction, the two teachers were interviewed to determine their approach to the teaching of writing. They reported to allot between 45 and 90 min to writing instruction weekly. This writing time was predominately devoted to grammar instruction and to independent text production. The teaching of grammar was based on traditional whole-class teaching methods. Teachers made no references to the use of sentence combining. Regarding text production, they reported to use the process approach that was recently included in the Portuguese language curriculum (Reis et al., 2009). Nevertheless, no references were made to the explicit and systematic teaching of either

writing strategies or self-regulation procedures to accomplish specific writing processes. In sum, writing instruction delivered to control students greatly differed from the one delivered to intervention students. Additionally, control students were asked to write the same number of opinion essays and on the same topic as the other students.

2.4. Testing Sessions

One week before and after instruction all students completed a pretest and a posttest, respectively. Testing sessions occurred in regular classroom groups and lasted 90 min. Students started by filling out a self-efficacy scale (Pajares & Valiante, 1999). Then, the experimenter presented the opinion essay topic (pretest: "Do you think teachers should give homework every days?"; and posttest: "Do you think children should work out every days?"), and gave students a blank sheet, in which they could write everything that would help them to write the essay (for a similar procedure see Berninger, Whitaker, Feng, Swanson, & Abbott, 1996). Students had 8 min to plan the text and 16 min to write it. Afterwards, students did a sentence-combining exercise, in which they combine four pairs of kernel sentences into a syntactically correct sentence. Lastly, students summarized a text. Only at posttest, intervention students filled out a scale to assess social validity.

After Lesson 5, all students completed a 60-min midtest. With a similar procedure to pretest and posttest, they were asked to plan and write an opinion essay ("Do you think children should go to bed early every days?") and to perform a sentence-combining exercise.

2.5. Measures

Except writing quality and variables calculated with the Computerized Language Analysis software (CLAN; MacWhinney, 2000), all other measures were scored by the first author. A second judge rescored one third of the measures at each testing time.

2.5.1. Strategy-specific measures. The developmental maturity of students' *planning* behavior was measured with a scale ranging from 1 (*low*) to 6 (*high*), which was based on the scales of Whitaker et al. (1994) and Olinghouse and Graham (2009). The scores 1 and 2 were

attributed to plans representing no and minimal preplanning, respectively. Plans summarizing the text received a score of 3, and plans with topics slightly elaborated in the text received a score of 4. The scores 5 and 6 were attributed to plans with emergent subordination (i.e., rudimentary macrostructure) and structural relationships (e.g., graphic organizers), respectively. At pretest, midtest, and posttest, inter-rater reliability using Cohen's weighted Kappa, was .87, .81, and .82, respectively. *Sentence-combining* skills were measured with a four-item exercise. For each item, one point was awarded if the student produced a syntactically correct sentence. An extra point was awarded if the sentence also included relevant changes that improved the quality of the original sentences (e.g., use of pronouns). The final score was the sum of the points awarded per item (max = 8). At pretest, midtest, and posttest, inter-rater reliability using Cohen's weighted Kappa was .95, .91, and .94, respectively.

2.5.2. Writing performance measures. Two measures were obtained from students' opinion essays: writing quality and text length. Two graduate students, blind to study purposes, assessed *quality*. To avoid biased judgments all texts were previously typed and corrected for spelling, punctuation, and capitalization errors (Berninger & Swanson, 1994). Using a scale ranging from 1 (*low*) to 7 (*high*), judges rated ideas quality, coherence, syntax, and vocabulary. The average across these factors was calculated for each rater. At pretest, midtest, and posttest, inter-rater reliability using the Intraclass Correlation Coefficient (ICC) was .95, .96, and .96, respectively. The final quality score was the average across raters. CLAN was used to obtain the number of words in the essays (*text length*).

2.5.3. Discourse-level measures. Opinion essays were scored for the presence and elaboration of four functional essay elements: premise, reasons, elaborations, and conclusion (based on Harris & Graham, 1996). Information off-topic or with no rhetorical purpose was rated as non-functional. For *premise* and *conclusion*, it was awarded one point if they were present, and two points if they were present and elaborated. For *reasons*, one point was awarded for each unique reason justifying the premise. For *elaborations*, one point was awarded for each reason

explained in depth (e.g., use of examples). At pretest, midtest, and posttest, inter-rater reliability of these measures using Cohen's weighted Kappa was greater than .75, .78, and .75, respectively. We also calculated *coherence* by dividing the number of functional elements by the number of functional plus non-functional elements.

2.5.4. Sentence-level measures. We evaluated the ratio of connective and opinion clauses, variety of connectives and opinion markers, and clause length. A clause was defined as a unit with a unified predicate and expressing a single situation (Berman & Slobin, 1994). *Ratio of connective clauses* and *ratio of opinion clauses* were calculated by dividing the number of connectives or opinion markers by the number of clauses. *Variety of connectives* and *variety of opinion markers* was obtained from the number of different connectives or opinion markers. *Clause length* (i.e., number of words per clause) was calculated with CLAN. At pretest, midtest, and posttest, inter-rater reliability of these measures using ICC was greater than .85, .89, and .85, respectively.

2.5.5. Word-level measures. We measured vocabulary diversity, and ratio and variety of modifiers (viz., adjectives and adverbs with the suffix –ly). *Vocabulary diversity* was assessed with a corrected type-token ratio (Carroll, 1964), which was calculated by dividing different words by the square root of two times the total words. *Ratio of modifiers* was the proportion of modifiers to text length. *Variety of modifiers* was obtained from the number of different modifiers. At pretest, midtest, and posttest, inter-rater reliability of these measures using ICC was greater than .92, .95, and .89, respectively.

2.5.6. Motivational measure: Self-efficacy. Students' self-efficacy beliefs were assessed with the Writing Skills Self-Efficacy scale (Pajares & Valiante, 1999; adapted to Portuguese by Limpo & Alves, 2013). The scale has 10 items, which measure students' confidence about being able to accomplish specific writing skills. The answers are given in a scale ranging from 0 (*no chance*) to 100 (*completely certain*). As confirmatory factor analysis on students' pretest responses showed a good fit of the data to a single-factor solution (CFI = .95, RMSEA = .09), the final score was the average across all items ($\alpha_{pretest}$ = .93; $\alpha_{posttest}$ = .92).

2.5.7. Generalization measure: Summary writing. The texts to be summarized reported an experiment showing animals' intelligence. Summaries were scored for sensitivity to importance and sentences transformation (based on Friend, 2001; Garner & McCaleb, 1985). For *sensitivity to importance*, we considered six elements: thesis, experiment description, conclusion, irrelevant ideas excluded, and misinterpretations. For the first three elements, it was awarded one point if they were present but incomplete, and two points if they were present and complete. As the stimulus text contained two irrelevant ideas, one point as awarded for each one that was excluded. We also counted the number of ideas that were misinterpretations of the original text. The final score was the total points awarded for thesis, experiment, conclusion, and irrelevant ideas excluded, minus the number of misinterpretations. At pretest and posttest, inter-rater reliability using Cohen's weighted Kappa was .83 and .83, respectively. For *sentences transformation*, one point was awarded for each syntactically correct sentence that subsumed two or more sentences from the original text. The final score was the proportion of correct transformations by the number of sentences. At pretest and posttest, inter-rater reliability using ICC was .91 and .87, respectively.

2.5.8. Social validity. At the end of the study, intervention students were asked to fill out a scale to assess the perceived value of the strategies. They rated their level of agreement (1 = *strongly disagree*; 6 = *strongly agree*) with these statements: (a) The strategy helped me to write better opinion essays; (b) The strategy helped me to write better in general; (c) The strategy was difficult; (d) I will continue to use the strategy; (e) The strategy should be taught to other students; and (f) I would like to learn strategies for other genres (based on MacArthur & Philippakos, 2010).

3. Results

In a set of preliminary analyses, firstly, we tested if our data met the normality assumption of parametric procedures. The inspection of the skewness and kurtosis of all pretest, midtest, and posttest scores revealed no distributional problems, as the absolute values of these indexes did not exceed 3.0 and 10.0, respectively (Kline, 2005). Secondly, as students were nested within classrooms, we tested if there were differences between classrooms for all dependent measures at

pretest. One-way analyses of variance (ANOVAs) showed no differences across classrooms (*p*s > .06), except for planning, text length, and self-efficacy. Thirdly, because our sample included fifth and sixth graders, we tested if there were differences between grades for all dependent measures at pretest. One-way ANOVAs using grade level as a between-subjects factor revealed no grade effects (*p*s > .13).

For planning, text length, and self-efficacy we conducted two-way nested Analyses of Covariance (ANCOVAs) with condition as a fixed factor, classroom as a random factor, and the respective pretest score as a covariate. Both at midtest and posttest, these analyses revealed no effects of classroom nested within condition (*F*s < 1.27, *p*s > .29). Thus, for all variables, we conducted 3 x 3 (Condition [planning, sentence combining, control] x Testing Time [pretest, midtest, posttest]) ANOVAs with repeated measures on the last factor. Table 2 provides means and standard deviations for all dependent variables and Table 3 presents ANOVAs' results. Significant Condition x Testing Time interactions were examined by means of tests of simple main effects, which are described in the text as follows. First, we report results regarding differences between conditions within each testing time. Then, we present results regarding differences between testing times within each condition. Significant simple effects were followed-up through pairwise comparisons. Table 4 reports Cohen's *d* for significant pairwise comparisons between conditions at midtest and posttest (Cohen, 1988). Because either in simple effects analysis or in pairwise comparisons we were making three comparisons at a time, to control for Type I error inflation, we used a Bonferroni adjustment with an alpha level of .017 ($\alpha = .05/3$).

Table 2, 3 and 4 about here

3.1. Strategy-Specific Measures

3.1.1. Planning. There were statistically significant effects for condition, testing time, and the interaction between the two. Tests of simple main effects for the interaction revealed differences between conditions at midtest, $F(2, 123) = 35.99$, $p = .001$, $\eta^2_p = 0.37$, and posttest,

$F(2, 123) = 210.05$, $p < .001$, $\eta^2_p = 0.77$. At midtest and posttest, follow-up analyses showed that planning students wrote more complex plans than sentence-combining and control students (ps < .001). Tests of simple main effects also revealed differences between testing sessions for the planning group, $\Lambda = .39$, $F(2, 122) = 97.52$, $p < .001$, $\eta^2_p = 0.62$. Follow-up analyses showed that planning skills increased from pretest to midtest, and from midtest to posttest (ps < .001).

3.1.2. Sentence combining. There were statistically significant effects for condition and the interaction between condition and testing time. Tests of simple main effects for the interaction revealed differences between conditions at midtest, $F(2, 123) = 5.53$, $p = .005$, $\eta^2_p = 0.08$, and posttest, $F(2, 123) = 13.02$, $p < .001$, $\eta^2_p = 0.18$. At midtest and posttest, follow-up analyses showed that sentence-combining students correctly combined more sentences than planning and control students (ps < .006). Tests of simple main effects also revealed differences between testing sessions for the sentence-combining group, $\Lambda = .88$, $F(2, 122) = 8.07$, $p < .001$, $\eta^2_p = 0.12$. Follow-up analyses showed an increase in sentence-combining skills from midtest to posttest ($p = .01$).

3.2. Writing Performance Measures

3.2.1. Quality. There were statistically significant effects for condition, testing time, and the interaction between the two. Tests of simple main effects for the interaction revealed differences between conditions at midtest, $F(2, 123) = 9.58$, $p < .001$, $\eta^2_p = 0.14$, and posttest, $F(2, 123) = 10.03$, $p < .001$, $\eta^2_p = 0.14$. At midtest and posttest, follow-up analyses showed that planning and sentence-combining students wrote qualitatively better opinion essays than control students (ps < .004). No differences were found between planning and sentence-combining conditions. Tests of simple main effects also revealed differences between testing sessions for the planning group, $\Lambda = .91$, $F(2, 122) = 6.39$, $p = .002$, $\eta^2_p = 0.10$, and for the sentence-combining group, $\Lambda = .92$, $F(2, 122) = 5.52$, $p = .005$, $\eta^2_p = 0.08$. For both groups, follow-up analyses showed that writing quality increased from pretest to posttest ($p = .001$).

3.2.2. Text length. Significant differences between conditions were found for text length at pretest. Thus, treatment effects on text length at midtest and posttest were analyzed with two one-

way ANCOVAs, with pretest scores as the covariate. For both analyses, we found no interactions between pretest scores and condition (*p*s > .19) meaning that the assumption of homogeneous regression slopes was met. The effect of condition was significant at midtest $F(2, 120) = 9.98$, $p < .001$, $\eta^2_p = 0.14$, and posttest $F(2, 120) = 4.08$, $p = .02$, $\eta^2_p = 0.06$. At both testing times, follow-up analyses revealed that planning and sentence-combining students wrote longer texts than control students (*p*s < .001).

3.3. Discourse-Level Measures

3.3.1. Premise. We found statistically significant effects of condition and testing time, but no interaction between the two. Planning students outperformed sentence-combining and control students ($p < .001$). Also, there was an increase from pretest to midtest ($p < .001$).

3.3.2. Reasons. There were statistically significant effects for condition, testing time, and the interaction between the two. Tests of simple main effects for the interaction revealed differences between conditions at midtest, $F(2, 123) = 14.52$, $p < .001$, $\eta^2_p = 0.19$, and posttest, $F(2, 123) = 22.03$, $p < .001$, $\eta^2_p = 0.26$. At midtest and posttest, follow-up analyses showed that planning students wrote more reasons than sentence-combining and control students (*p*s < .005). Only at posttest, sentence-combining students wrote more reasons than control students ($p = .003$). Tests of simple main effects also revealed differences between testing sessions for the planning group, $\Lambda = .65$, $F(2, 122) = 33.15$, $p < .001$, $\eta^2_p = 0.35$. Follow-up analyses showed an increase in the number of reasons from pretest to midtest ($p < .001$).

3.3.3. Elaborations. There were statistically significant effects for condition, testing time, and the interaction between the two. Tests of simple main effects for the interaction revealed differences between conditions at midtest, $F(2, 123) = 4.98$, $p = .008$, $\eta^2_p = 0.08$, and posttest, $F(2, 123) = 17.94$, $p < .001$, $\eta^2_p = 0.23$. At midtest and posttest, follow-up analyses showed that planning students wrote more elaborated reasons than control students (*p*s < .003). Moreover, only at posttest, planning students wrote more elaborated reasons than sentence-combining students ($p < .001$). Tests of simple main effects also revealed differences between testing sessions for the

planning group, $\Lambda = .70$, $F(2, 122) = 25.81$, $p < .001$, $\eta^2_p = 0.30$. Follow-up analyses showed that the number of elaborations progressively increased from pretest to midtest, and from midtest to posttest (ps < .001).

3.3.4. Conclusion. There were statistically significant effects for condition, testing time, and the interaction between the two. Tests of simple main effects for the interaction revealed differences between conditions at midtest, $F(2, 123) = 7.60$, $p = .001$, $\eta^2_p = 0.11$, and posttest, $F(2, 123) = 34.16$, $p < .001$, $\eta^2_p = 0.36$. At midtest and posttest, follow-up analyses showed that both planning and sentence-combining students surpassed control students (ps < .003). Tests of simple main effects also revealed differences between testing sessions for the planning group, $\Lambda = .71$, $F(2, 122) = 24.75$, $p < .001$, $\eta^2_p = 0.29$, and for the sentence-combining group, $\Lambda = .65$, $F(2, 122) = 33.54$, $p < .001$, $\eta^2_p = 0.36$. Follow-up analyses showed that both groups' conclusion scores increased from pretest to midtest (ps < .001).

3.3.5. Coherence. There were statistically significant effects of testing time and the interaction between condition and testing time. Tests of simple main effects for the interaction revealed differences between conditions at posttest, $F(2, 123) = 9.38$, $p < .001$, $\eta^2_p = 0.13$. Follow-up analyses showed that planning students wrote more coherently than sentence-combining and control students (ps < .001). Tests of simple main effects also revealed differences between testing sessions for the planning group, $\Lambda = .87$, $F(2, 122) = 9.29$, $p < .001$, $\eta^2_p = 0.13$. Follow-up analyses showed an increase in coherence from pretest to midtest ($p = .002$).

3.4. Sentence-Level Measures

3.4.1. Ratio of connective clauses. There were no statistically significant effects.

3.4.2. Variety of connectives. There were statistically significant effects for condition, testing time, and the interaction between the two. Tests of simple main effects for the interaction revealed differences between conditions at midtest, $F(2, 123) = 11.43$, $p < .001$, $\eta^2_p = 0.16$, and posttest, $F(2, 123) = 54.02$, $p < .001$, $\eta^2_p = 0.47$. At midtest and posttest, follow-up analyses showed that both sentence-combining and planning students wrote a wider variety of connectives

than control students (*p*s < .001). Moreover, only at posttest, sentence-combining students outperformed planning students ($p < .001$). Tests of simple main effects also revealed differences between testing sessions for the sentence-combining group, $\Lambda = .48$, $F(2, 122) = 65.38$, $p < .001$, $\eta^2_p = 0.52$. Follow-up analyses showed a progressive increase in the variety of connectives from pretest to midtest, and from midtest to posttest (*p*s < .001).

3.4.3. Ratio of opinion clauses. There were statistically significant effects for condition, testing time, and the interaction between the two. Tests of simple main effects for the interaction revealed differences between conditions at posttest, $F(2, 123) = 35.97$, $p < .001$, $\eta^2_p = 0.37$. Follow-up analyses showed that sentence-combining and planning students wrote more clauses with opinion markers than control students (*p*s < .003). Moreover, sentence-combining students outperformed planning students ($p < .001$). Tests of simple main effects also revealed differences between testing sessions for the sentence-combining group, $\Lambda = .60$, $F(2, 122) = 40.42$, $p < .001$, $\eta^2_p = 0.40$. Follow-up analyses showed that the ratio of opinion clauses progressively increase from pretest to midtest, and from midtest to posttest (*p*s < .001).

3.4.4. Variety of opinion markers. There were statistically significant effects for condition, testing time, and the interaction between the two. Tests of simple main effects for the interaction revealed differences between conditions at midtest, $F(2, 123) = 13.63$, $p < .001$, $\eta^2_p = 0.18$, and posttest, $F(2, 123) = 119.05$, $p < .001$, $\eta^2_p = 0.66$. At midtest and posttest, follow-up analyses showed that sentence-combining and planning students used more diversified opinion markers than control students (*p*s < .011). Moreover, sentence-combining students outperformed planning students (*p*s < .004). Tests of simple main effects also revealed differences between testing sessions for the sentence-combining group, $\Lambda = .26$, $F(2, 122) = 172.47$, $p < .001$, $\eta^2_p = 0.74$, and for the planning group, $\Lambda = .77$, $F(2, 122) = 18.72$, $p < .001$, $\eta^2_p = 0.24$. Follow-up analyses showed a progressive increase in the variety of connectives from pretest to midtest, and from midtest to posttest (*p*s < .007).

3.4.5. Clause length. There was a statistically significant interaction between condition and testing time. Tests of simple main effects for the interaction revealed differences between conditions at posttest, $F(2, 123) = 12.38$, $p < .001$, $\eta^2_p = 0.17$. Follow-up analyses showed that sentence-combining students wrote longer clauses than planning and control students (ps < .001). Tests of simple main effects also revealed differences between testing sessions for the sentence-combining group, $\Lambda = .89$, $F(2, 122) = 7.52$, $p < .001$, $\eta^2_p = 0.11$. Follow-up analyses showed that the number of words per clause increased from pretest to midtest ($p = .005$).

3.5. Word-Level Measures

3.5.1. Vocabulary diversity. There were statistically significant effects for condition, testing time, and the interaction between the two. Tests of simple main effects for the interaction revealed differences between conditions at midtest, $F(2, 123) = 7.48$, $p < .001$, $\eta^2_p = 0.11$, and posttest, $F(2, 123) = 11.17$, $p < .001$, $\eta^2_p = 0.15$. At midtest and posttest, follow-up analyses showed that sentence-combining and planning students used more varied vocabulary than control students (ps < .016). Moreover, only at posttest, sentence-combining students outperformed planning students ($p = .013$). Tests of simple main effects revealed differences between testing sessions for the sentence-combining group, $\Lambda = .74$, $F(2, 122) = 21.29$, $p < .001$, $\eta^2_p = 0.26$. Follow-up analyses showed a progressive increase in vocabulary diversity from pretest to midtest, and from midtest to posttest (ps < .001).

3.5.2. Ratio of modifiers. There was a statistically significant effect of testing time, with a progressive increase from pretest to midtest, and from midtest to posttest ($p < .012$). The Condition x Testing Time interaction was not reliable.

3.5.3. Variety of modifiers. There were statistically significant effects for condition, testing time, and the interaction between the two. Tests of simple main effects for the interaction revealed differences between conditions at posttest, $F(2, 123) = 12.70$, $p < .001$, $\eta^2_p = 0.17$. Follow-up analyses showed that sentence-combining and planning students used more diversified modifiers than control students (ps < .003). Tests of simple main effects also revealed differences

between testing sessions for the sentence-combining group, $\Lambda = .58$, $F(2, 122) = 44.86$, $p < .001$, $\eta^2_p = 0.42$, and for the planning group, $\Lambda = .74$, $F(2, 122) = 21.30$, $p < .001$, $\eta^2_p = 0.26$. Follow-up analyses showed that the variety of modifiers increased from pretest to midtest, and from midtest to posttest in the sentence-combining group (ps < .001), but only increased from midtest to posttest in the planning group ($p < .001$).

3.6. Motivational Measure: Self-Efficacy

There was a statistically significant main effect of testing time, with self-efficacy beliefs increasing from pretest to posttest. No interaction between condition and testing time was found. To explore the relationship between self-efficacy and writing quality we calculated the correlation between these two variables for each condition before and after instruction. At pretest, we did not find statistically significant correlations in any group (rs < .29, ps > .08). Notably, at posttest, self-efficacy and writing quality were correlated in both planning ($r = .46$, $p < .001$) and sentence-combining ($r = .51$, $p < .001$) groups, but not in the control group ($r = .20$, $p = .22$).

3.7. Generalization Measure: Summary Writing

3.7.1. Sensitivity to importance. There was a statistically significant interaction between condition and testing time. Tests of simple main effects for the interaction revealed differences between conditions at posttest, $F(2, 123) = 5.26$, $p = .006$, $\eta^2_p = 0.08$. Follow-up analyses showed that planning students were better at discriminating information relevance than sentence-combining and control students (ps < .007). Tests of simple main effects also revealed a statistically significant increase in sensitivity to importance from pretest to posttest for the planning group, $\Lambda = .90$, $F(2, 122) = 13.18$, $p < .001$, $\eta^2_p = 0.10$.

3.7.2. Sentences transformation. Although there was a statistically significant main effect of testing time, the Condition x Testing Time interaction was not reliable.

3.8. Social Validity

Intervention students were very positive about the perceived value of the taught strategies. Almost all students agreed (5) or strongly agreed (6) that: the strategy helped them to write better

in general (92%; *M* = 5.44, *SD* = 0.33) and opinion essays in particular (99%; *M* = 5.91, *SD* = 0.33), they would continue to use the strategy (92%; *M* = 5.64, *SD* = 0.67), the strategy should be taught to other students (99%; *M* = 5.82, *SD* = 0.41), and they would like to learn strategies for other genres (89%; *M* = 5.48, *SD* = 0.78). Also, 66% of the students disagreed (2) or strongly disagreed (1) that the strategy was difficulty (*M* = 2.27, *SD* = 1.38).

3.9. Additional Analyses: Classroom Effects

Because interventions were delivered to intact classrooms, we further explored if there were differences across classrooms participating in the same instructional program. For that, we conducted separate 2 (classroom) x 3 (testing time) ANOVAs for planning and sentence-combining conditions. Except for coherence in planning instruction, $\Lambda = .66$, $F(2, 45) = 11.52$, $p < .001$, $\eta^2_p = 0.34$, and variety of modifiers in sentence-combining instruction, $\Lambda = .76$, $F(2, 36) = 5.82$, $p = .006$, $\eta^2_p = 0.24$, instructional effects were not moderated by classroom (Fs < 2.38, ps < .12). Regarding coherence, the two classrooms receiving planning instruction significantly increased across instruction, $\Lambda = .65$, $F(2, 45) = 12.21$, $p < .001$, $\eta^2_p = .35$ versus $\Lambda = .67$, $F(2, 45) = 10.90$, $p < .001$, $\eta^2_p = .33$, and showed no differences at posttest ($F < 1$). Nevertheless, at midtest, one of the classes wrote more coherent texts than the other one, $F(1, 46) = 13.27$, p = .001, $\eta^2_p = .22$. A similar patter was found for variety of modifiers in sentence-combining instruction. Both classrooms significantly improved across instruction, $\Lambda = .45$, $F(2, 36) = 22.30$, $p < .001$, $\eta^2_p = .55$ versus $\Lambda = .46$, $F(2, 36) = 20.92$, $p < .001$, $\eta^2_p = .54$, and showed no posttest differences ($F < 1$). Still, at midtest, one of the classes used more diverse modifiers than the other one, $F(1, 37) = 8.70$, $p = .005$, $\eta^2_p = .19$.

4. Discussion

The present study evaluated the effectiveness of two strategy-focused interventions aimed to boost fifth and sixth graders' opinion essay writing by teaching them a planning or a sentence-combining strategy. Instructional effects were assessed on strategy-specific skills, writing

performance, discourse-, sentence-, and word-levels of written language, self-efficacy beliefs, and summary writing.

4.1. Strategy-Specific Effects

In line with previous findings (e.g., Glaser & Brunstein, 2007; Saddler & Graham, 2005), we showed that planning and sentence-combining instruction increased students' planning and sentence-construction skills, respectively. In comparison to sentence-combining and control students, planning students created more complex plans, both at midtest and posttest. These students' planning skills largely and progressively increased throughout instruction. Conversely, sentence-combining students correctly combine more sentences than planning and control students, both at midtest and posttest. Although students' ability to combine sentences showed an upward trend during instruction, it only significant increased from midtest to posttest. This result suggests that the first five lessons did not provide students enough practice in sentence combining. Furthermore, it reinforces Strong's (1986) claim that sentence-combining skills may benefit from providing students extended opportunities to apply them in composition.

4.2. Writing Performance Effects

The hypotheses regarding the impact of strategy instruction on writing performance were partially confirmed. We found that the interventions focused either on planning or on sentence combining had a positive impact on opinion essay quality. These effects were found at posttest, and also at midtest. The fostering of self-regulation might have promoted the integration of the target strategy into composition, even before students were explicitly instructed in how to do it. Yet, we think that the instructional component focused on this integration was decisive given that only posttest quality was superior to pretest quality. Reproducing meta-analyses findings (Graham & Perin, 2007), we found stronger effect sizes for planning than sentence-combining instruction, even though the interventions did not significantly differ between them. Notably, this is the first study showing the effectiveness of teaching a sentence-combining strategy coupled with self-

regulation procedures to promote writing quality. This result supports the use of the SRSD instructional model to foster composing processes besides planning and revising.

As predicted, strategy instruction also resulted in longer opinion essays than control instruction at posttest. Although increases in text length had already been reported after planning instruction (De La Paz & Graham, 2002; Graham et al., 2005; Harris et al., 2006), Saddler and Graham (2005) found no changes in text length after sentence-combining instruction. Despite some differences between their study and ours (e.g., target genre, duration of instruction, participants' grade), the promotion of self-regulation in our program might explain the very large posttest effect size favoring sentence-combining instruction.

4.3. Discourse-, Sentence-, and Word-Level Effects

As expected, we found that planning students wrote more complete and coherent essays than sentence-combining and control students. Thus, teaching a planning strategy plus self-regulation procedures seems to be an effective way to promote students' writing at the discourse level. Importantly, planning students register a progressive growth in all discourse-level measures across testing sessions, even though only pretest-midtest increases were larger enough to be statistically significant. It seems that practicing the planning strategy alone was sufficient for students to increase the completeness and coherence of their opinion essays. This is not to say, however, that composing opportunities are worthless. Actually, three findings suggest that planning students beneficiated from lessons integrating the planning strategy with composition: (a) the strongest effect sizes favoring planning instruction occurred at posttest; (b) the superiority of planning over sentence-combining instruction was mainly found at posttest; and (c) only one of the classes receiving planning instruction wrote more coherent texts than their peers at midtest. It is also noteworthy that planning students surpassed control students at the sentence and word levels. It is likely that by generating and organizing their ideas before writing, they were able to focus on sentence and word-level concerns during writing (Graham & Harris, 2007; Kellogg, 1988). The finding that clause length remained uninfluenced by planning instruction was not surprising. Even

traditional grammar instruction was shown to have no effect on syntactic complexity (Andrews et al., 2006). The instruction provided by the sentence-combining program seems to be needed to boost such a specific skill as the ability to produce complex syntactic structures.

Proving the effectiveness of sentence-combining instruction to influence students' writing at the sentence and word levels, we found that sentence-combining students displayed better sentence-construction and vocabulary skills than planning and control students. This result might be explained by the teaching of a sentence-combining strategy in tandem with self-regulation procedures, along with the use of specific practices to promote the transfer of the taught skills to composition (see section 2.2.1. for a description of these practices). Importantly, we also found that the majority of sentence-combining students' sentence- and word-level scores progressively increased from pretest to midtest, and from midtest to posttest. Still, only at posttest have sentence-combining students clearly outperformed their peers in almost all sentence- and word-level measures. This pattern of findings indicates that the use of isolated exercises to enhance sentence-combining skills might not be enough for students to transfer those skills to composition. Explicit teaching and systematic training in employing sentence-combining skills in writing seems to be needed not only to increase these skills (see the above section 4.1.), but also to apply them in text production. It is also noteworthy that sentence-combining instruction favored some discourse-related aspects of composition. The increase in sentence construction fluency might have enabled sentence-combining students to attend to other aspects of composition, such as text content and structure (Fayol, 1999; McCutchen et al., 1994; Strong, 1986). Probably, sentence-combining instruction did not influence elaborations and coherence because their development may require explicit teaching and guided practice.

4.6. Motivational Effects

Contrary to our predictions, the interventions failed to increase students' self-efficacy (for similar results, see Graham et al., 2005; Page-Voth & Graham, 1999; Sawyer et al., 1992), probably, because it was being overestimated at pretest. Before instruction, self-efficacy and

writing quality were not correlated. These results are consistent with Limpo and Alves (2013), who found that self-efficacy did not contribute to writing quality in Grades 4-6. Of great import, this pattern changed at posttest, where we found moderate correlations between self-efficacy and writing quality for both planning and sentence-combining groups. Strategy instruction seemed to reduce the discrepancy between students' self-efficacy beliefs and their actual performance, which might be explained by the teaching of self-regulation procedures (Klassen, 2002; Schunk, 2003). In particular, self-monitoring might have helped students to gain conscious access to their successes and failures, turning their perception of ability more realistic and adjusted to their current performance.

4.7. Generalization Effects

As anticipated, planning instructional effects transferred to summary writing. Planning instruction seemed to be beneficial for students to discriminate between information that should and should not be included in a summary. This result suggests that besides mastering the planning routine, they also mastered the core principles of selecting and organizing information (Shepard, 2000). This enhanced sensitivity to importance via planning instruction was an important finding because this strategic skill is related to reading comprehension (Winograd, 1983). Refuting our hypothesis, sentence-combining effects did not transfer to sentences transformation in summary writing. Possibly, sentence-combining skills were not sufficiently acquired to be generalized to a different task (Chi & VanLehn, 2012). Indeed, students only obtained 4.52 points out of 8 in the posttest sentence-combining task. Alternatively, they might have failed in detecting any link between the sentence-combining and summary tasks (Perkins & Salomon, 2012). Perhaps due to poor procedural knowledge about summarizing, they did not know that they should have transformed rather than reproduced the original sentences.

5. Limitations and Future Directions

Six limitations of the current study should guide future research. First, as standardized writing tests in Portuguese are lacking, we only used researcher-constructed tests. Future studies

should also include standardized measures to assess instructional effects more comprehensively. Second, although participants were nested within classrooms, given the few classes involved (N = 6), we used participants as the unit of analysis. Overall, supplementary analyses suggested that interventions effectiveness did not differ across classrooms for the majority of outcomes assessed. Still, large-scale intervention studies using multilevel analyses are warranted to explore instructional effects both at the student and classroom levels. Third, writing prompts were not counterbalanced across testing sessions. Nevertheless, it was found that young students produce texts of similar length and quality in response to different opinion essay prompts (Harris et al., 2006). Indeed, we found no differences over time regarding control students opinion essays. Fourth, as youngsters are unlikely to preplan spontaneously (McCutchen, 2006), to avoid pretest floor effects, we asked them to preplan in all testing sessions. Nonetheless, more research is needed to examine the influence of writing instruction on the management of writing processes, similarly to Torrance et al. (2007; see also Fidalgo, Torrance, & García, 2008). Fifth, time constraints in the testing sessions impeded us to analyze instructional effects on revision, which should be address in the future because sentence combining can also be a revising tool (Zimmerman & Kitsantas, 2002). Finally, the present study did not explicitly test the incremental effect of self-regulatory training above and beyond sentence-combining instruction. Further research should compare teaching sentence combining plus self-regulation procedures with teaching sentence combining alone.

6. Conclusions

The present study adds to a growing body of research (Graham & Harris, 2007; Graham & Perin, 2007; Saddler, 2007) that using the SRSD model to teach key writing processes, such as planning and translation, is an effective way to foster students' writing. Of critical relevance was our finding that whereas planning instruction primarily boosted the discourse-related aspects of composition, sentence-combining instruction primarily enhanced writing at the sentence and word levels. The educational implication of this finding is twofold. On the one hand, by knowing

students' writing needs, writing instruction can be specifically tailored to the levels of writing that are a struggle for them. On the other hand, if the instructional goal is to promote text production in a comprehensive way, it will be beneficial to target several levels of written composition, simultaneously. As the present study compared two strategies targeting different writing processes occurring at different moments of text production, we believe that planning and sentence-combining instruction can be integrated within a single writing program. This program could taught students not only to generate and organize their ideas ahead of writing but also to transform them in interesting and mature sentences in writing. Indeed, although instructional effects showed some transfer across levels, by coupling planning with sentence-combining instruction, one can expect far-reaching gains in students' ability to write proficiently at several levels of text production.

References

Alexander, P. A., Graham, S., & Harris, K. R. (1998). A perspective on strategy research: Progress and prospects. *Educational Psychology Review, 10*, 129-154. doi: 10.1023/A:1022185502996

Andrews, R., Torgerson, C., Beverton, S., Freeman, A., Locke, T., Low, G., . . . Zhu, D. (2006). The effect of grammar teaching on writing development. *British Educational Research Journal, 32*, 39-55. doi: 10.1080/01411920500401997

Bandura, A. (1969). *Principles of behavior modification*. New York, NY: Holt, Rinehart and Winston.

Beauvais, C., Olive, T., & Passerault, J.-M. (2011). Why are some texts good and others not? Relationship between text quality and management of the writing processes. *Journal of Educational Psychology, 103*, 415-428. doi: 10.1037/a0022545

Beers, S. F., & Nagy, W. E. (2009). Syntactic complexity as a predictor of adolescent writing quality: Which measures? Which genre? *Reading and Writing: An Interdisciplinary Journal, 22*, 185-200. doi: 10.1007/s11145-007-9107-5

Bereiter, C., & Scardamalia, M. (1987). *The psychology of written composition*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Berman, R. A., & Slobin, D. I. (Eds.). (1994). *Relating events in narrative: A crosslinguistic developmental study*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Berninger, V. W., Nagy, W. E., & Beers, S. F. (2011). Child writers' construction and reconstruction of single sentences and construction of multi-sentence texts: Contributions of syntax and transcription to translation. *Reading and Writing: An Interdisciplinary Journal, 24*, 151-182. doi: 10.1007/s11145-010-9262-y

Berninger, V. W., & Swanson, H. L. (1994). Modifying Hayes and Flower's model of skilled writing to explain beginning and developing writing. In E. C. Butterfield (Ed.), *Children's writing: Toward a process theory of the development of skilled writing* (Vol. 2, pp. 57-81). Greenwich, Connecticut: JAI Press.

Berninger, V. W., Whitaker, D., Feng, Y., Swanson, H. L., & Abbott, R. D. (1996). Assessment of planning, translating, and revising in junior high writers. *Journal of School Psychology, 34*, 23-52. doi: 10.1016/0022-4405(95)00024-0

Berninger, V. W., Yates, C. M., Cartwright, A. C., Rutberg, J., Remy, E., & Abbott, R. D. (1992). Lower-level developmental skills in beginning writing. *Reading and Writing: An Interdisciplinary Journal, 4*, 257-280. doi: 10.1007/BF01027151

Boscolo, P., & Hidi, S. (2007). The multiple meanings of motivation to write. In S. Hidi & P. Boscolo (Eds.), *Writing and motivation* (pp. 1-14). Amsterdam: Elsevier.

Brunstein, J. C., & Glaser, C. (2011). Testing path-analytic mediation model of how self-regulated writing strategies improve fourth graders' composition skills: A randomized controlled trial. *Journal of Educational Psychology*, Advanced online publication. doi: 10.1037/a0024622

Carroll, J. B. (1964). *Language and thought*. Englewood Cliffs, NJ: Prentice-Hall, Inc.

Chi, M. T. H., & VanLehn, K. A. (2012). Seeing deep structure from the interactions of surface features. *Educational Psychologist, 47*, 177-188. doi: 10.1080/00461520.2012.695709

De La Paz, S., & Graham, S. (2002). Explicitly teaching strategies, skills, and knowledge: Writing instruction in middle school classrooms. *Journal of Educational Psychology, 94*, 687-698. doi: 10.1037//0022-0663.94.4.687

Favart, M. (2005). Les marques de cohésion: Leur rôle fonctionnel dans l'acquisition de la production écrite de texte. *Psychologie Française, 50*, 305-322. doi: 10.1016/j.psfr.2005.006

Fayol, M. (1999). From on-line management problems to strategies in written composition. In M. Torrance & G. Jeffery (Eds.), *The cognitive demands of writing: Processing capacity and working memory effects in text production* (pp. 13-23). Amsterdam: Amsterdam University Press.

Fidalgo, R., Torrance, M., & García, J.-N. (2008). The long-term effects of strategy-focussed writing instruction for grade six students. *Contemporary Educational Psychology, 33*, 672-693. doi: 10.1016/j.cedpsych.2007.09.001

Friend, R. (2001). Effects of strategy instruction on summary writing of college students. *Contemporary Educational Psychology, 26*, 3-24. doi: 10.1006/ceps.1999.1022

Garner, R., & McCaleb, J. L. (1985). Effects of text manipulations on quality of written summaries. *Contemporary Educational Psychology, 10*, 139-149. doi: 10.1016/0361-476X(85)90014-1

Glaser, C., & Brunstein, J. C. (2007). Improving fourth-grade students' composition skills: Effects of strategy instruction and self-regulation procedures. *Journal of Educational Psychology, 99*, 297-310. doi: 10.1037/0022-0663.99.2.297

Graham, S. (2006a). Writing. In P. A. Alexander & P. H. Winne (Eds.), *Handbook of educational psychology*. Mahwah, NJ: Erlbaum.

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

Graham, S., Berninger, V. W., Abbott, R. D., Abbott, S. P., & Whitaker, D. (1997). Role of mechanics in composing of elementary school students: A new methodological approach. *Journal of Educational Psychology, 89*, 170-182. doi: 10.1037/0022-0663.89.1.170

Graham, S., Gillespie, A., & McKeown, D. (2012). Writing: Importance, development, and instruction. *Reading and Writing: An Interdisciplinary Journal*(Advanced online publication). doi: 10.1007/s11145-012-9395-2

Graham, S., & Harris, K. R. (2000). The role of self-regulation and transcription skills in writing and writing development. *Educational Psychologist, 35*, 3-12. doi: 10.1207/S15326985EP3501_2

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

Graham, S., & Harris, K. R. (2007). Best practices in teaching planning. In S. Graham, C. A. MacArthur & J. Fitzgerald (Eds.), *Best practices in writing instruction*. New York, NY: The Guilford Press.

Graham, S., Harris, K. R., & Mason, L. H. (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. doi: 10.1016/j.cedpsych.2004.08.001

Graham, S., McKeown, D., Kiuhara, S., & Harris, K. R. (2012). A meta-analysis of writing instruction for students in the elementary grades. *Journal of Educational Psychology, 104*, 879-896. doi: 10.1037/a0029185

Graham, S., & Perin, D. (2007). A meta-analysis of writing instruction for adolescent students. *Journal of Educational Psychology, 99*, 445-476. doi: 10.1037/0022-0663.99.3.445

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

Harris, K. R., & Graham, S. (2009). Self-regulated strategy development in writing: Premises, evolution, and the future. *British Journal of Educational Psychology Monograph Series, II*(6), 113-135. doi: 10.1348/978185409X422542

Harris, K. R., Graham, S., Brindle, M., & Sandmel, K. N. (2009). Metacognition and children's writing. In D. J. Hacker, J. Dunlosky & A. C. Graesser (Eds.), *Handbook of metacognition in education* (pp. 131-153). Mahwah, NJ: Erlbaum.

Harris, K. R., Graham, S., & Mason, L. H. (2006). Improving the writing, knowledge, and motivation of struggling young writers: Effects of self-regulated strategy development with and without peer support. *American Educational Research Journal, 43*, 295-340. doi: 10.3102/00028312043002295

Harris, K. R., Graham, S., Mason, L. H., & Friedlander, B. (2008). *Powerful writing strategies for all students* Baltimore: Brookes.

Hayes, J. R., & Flower, L. (1980). Identifying the organization of writing processes. In L. W. Gregg & E. R. Steinberg (Eds.), *Cognitive processes in writing* (pp. 3-29). Hillsdale, NJ: Lawrence Erlbaum Associates.

Hayes, J. R., & Flower, L. (1986). Writing research and the writer. *American Psychologist, 41*, 1106-1113. doi: 10.1037/0003-066X.41.10.1106

Hillocks, G. (1986). *Research on written composition*. Urbana, IL: ERIC Clearing-house on Reading and Communications Skills.

Kellogg, R. T. (1988). Attentional overload and writing performance: Effects of rough draft and outline strategies. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14*, 355-365. doi: 10.1037/0278-7393.14.2.355

Klassen, R. (2002). A question of calibration: A review of the self-efficacy beliefs of students with learning disabilities. *Learning Disability Quarterly, 25,* 88-102. doi: 10.2307/1511276

Kline, R. B. (2005). *Principles and practice of structural equation modeling* (2nd ed.). New York, NY: The Guilford Press.

Limpo, T., & Alves, R. A. (2013). Modeling writing development: Contribution of transcription and self-regulation to Portuguese students' text generation quality. *Journal of Educational Psychology, 105,* 401-413. doi: 10.1037/a0031391

Locke, E., Shaw, K., Saari, L., & Latham, G. (1981). Goal setting and task performance: 1969-1980. *Psychological Bulletin, 90,* 125-152.

MacArthur, C. A., & Philippakos, Z. (2010). Instruction in a strategy for compare-contrast writing. *Council for Exceptional Children, 76,* 438-456.

Mace, F. C., Belfiore, P. J., & Hutchinson, J. M. (2001). Operant theory and research on self-regulation. In D. H. Schunk & B. J. Zimmerman (Eds.), *Self-regulated learning and academic achievement* (pp. 39-65). Mahwah, NJ: Lawrence Erlbaum Associates.

MacWhinney, B. (2000). *The Childes project: Tools for analyzing talk.* Mahwah, NJ: Lawrence Erlbaum Associates.

McCutchen, D. (1986). Domain Knowledge and linguistic knowledge in the development of writing ability. *Journal of Memory and Language, 25,* 431-444.

McCutchen, D. (1996). A capacity theory of writing: Working memory in composition. *Educational Psychology Review, 8,* 299-325. doi: 10.1007/BF01464076

McCutchen, D. (2006). Cognitive factors in the development of children's writing. In C. A. MacArthur, S. Graham & J. Fitzgerald (Eds.), *Handbook of writing research* (pp. 115-130). New York, NY: The Guilford Press.

McCutchen, D. (2011). From novice to expert: Implications of language skills and writing-relevant knowledge for memory during the development of writing skill. *Journal of Writing Research, 3,* 51-68.

McCutchen, D., Covill, A., Hoyne, S. H., & Mildes, K. (1994). Individual differences in writing: Implications of translating fluency. *Journal of Educational Psychology, 86*, 256-266.

Meichenbaum, D. (1977). *Cognitive-behavior modification: An integrative approach*. New York, NY: Plenum Press.

Nelson, R. O., & Hayes, S. C. (1981). Theoretical explanations for reactivity in self-monitoring. *Behavior Modification, 5*, 3-14. doi: 10.1177/014544558151001

Olinghouse, N. G., & Graham, S. (2009). The relationship between the discourse knowledge and the writing performance of elementary-grade students. *Journal of Educational Psychology, 101*, 37-50. doi: 10.1037/a0013248

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.

Pajares, F. (2003). Self-efficacy beliefs, motivation, and achievement in writing: A review of the literature. *Reading and Writing Quarterly, 19*, 139-158. doi: 10.1080/10573560390143085

Pajares, F., & Valiante, G. (1999). Grade level and gender differences in the writing self-beliefs of middle school students. *Contemporary Educational Psychology, 24*, 390-405. doi: 10.1006/ceps.1998.0995

Perkins, D. N., & Salomon, G. (2012). Knowledge to go: A motivational and dispositional view of transfer. *Educational Psychologist, 47*, 248-258. doi: 10.1080/00461520.2012.693354

Pressley, M., & Harris, K. R. (2006). Cognitive strategies instruction: From basic research to classroom instruction. In P. A. Alexander & P. H. Winne (Eds.), *Handbook of educational psychology* (pp. 265-286). New York: Macmillan.

Rankin, J. L., & Reid, R. (1995). The SM rap - Or, here's the rap on self-monitoring. *Intervention in school and clinic, 30*, 181-188. doi: 10.1177/105345129503000309

Reis, C., Dias, A. P., Cabral, A. T., Silva, E., Viegas, F., Bastos, G., . . . Pinto, M. O. (2009). *Programas de português do ensino básico*. Lisboa: Ministério da Educação.

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

Saddler, B. (2007). Improving sentence construction skills through sentence-combining practice. In S. Graham, C. A. MacArthur & J. Fitzgerald (Eds.), *Best practices in writing instruction*. New York, NY: The Guilford Press.

Saddler, B., & Asaro, K. (2008). Beyond noun-verb: The use of sentence combining to improve sentence writing ability. *Insights on Learning Disabilities, 5*, 41-50.

Saddler, B., Behforooz, B., & Asaro, K. (2008). The effects of sentence-combining instruction on the writing of fourth-grade students with writing difficulties. *The Journal of Special Education, 42*, 79-90. doi: 10.1177/0022466907310371

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. J., 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.

Schunk, D. H. (2001). Social cognitive theory and self-regulated learning. In D. H. Schunk & B. J. Zimmerman (Eds.), *Self-regulated learning and academic achievement*. Mahwah, NJ: Lawrence Erlbaum Associates.

Schunk, D. H. (2003). Self-efficacy for reading and writing: Influence of modeling, goal setting, and self-evaluation. *Reading and Writing Quarterly, 19*, 159-172.

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

Shepard, L. A. (2000). The role of assessment in a learning culture. *Educational Researcher, 29*, 4-14. doi: 10.3102/0013189X029007004

Strong, W. (1986). *Creative approaches to sentence combining*. Urbana, IL: ERIC Clearinghouse on Reading and Communication Skills and the National Council of Teachers of English.

Strong, W. (1996). *Writer's toolbox: A sentence-combining workshop*. New York, NY: McGraw-Hill.

Torrance, M., Fidalgo, R., & García, J.-N. (2007). The teachability and effectiveness of cognitive self-regulation in sixth-grade writers. *Learning and Instruction, 17*, 265-285. doi: 10.1016/j.learninstruc.2007.02.003

Wagner, R. K., Puranik, C. S., Foorman, B., Foster, E., Wilson, L. G., Tschinkel, E., & Kantor, P. T. (2011). Modeling the development of written language. *Reading and Writing: An Interdisciplinary Journal, 24*, 203-220. doi: 10.1007/s11145-010-9266-7

Whitaker, D., Berninger, V. W., Johnston, J., & Swanson, H. L. (1994). Intraindividual differences in levels of language in intermediate grade writers: Implications for the translating process. *Learning and Individual Differences, 6*, 107-130. doi: 10.1016/1041-6080(94)90016-7

Winograd, P. N. (1983). Strategic difficulties in summarizing texts. *Reading Research Quarterly, 19*, 404-425. doi: 10.2307/747913

Wong, B. Y. L., Hoskyn, M., Jai, D., Ellis, P., & Watson, K. (2008). The comparative efficacy of two approaches to teaching sixth graders opinion essay writing. *Contemporary Educational Psychology, 33*, 757.784. doi: 10.1016/j.cedpsych.2007.12.004

Zimmerman, B. J. (1995). Self-regulation involves more than metacognition: A social cognitive perspective. *Educational Psychologist, 30*, 217-221. doi: 10.1207/s15326985ep3004_8

Zimmerman, B. J. (2000). Attaining self-regulation: A social-cognitive perspective. In M. Boekaerts, P. R. Pintrich & M. Zeidner (Eds.), *Self-regulation: Theory, research, and applications* (pp. 13-39). Orlando, FL: Academic Press.

Zimmerman, B. J. (2009). Goal setting: A key proactive source of academic self-regulation. In D. H. Schunk & B. J. Zimmerman (Eds.), *Motivation and self-regulated learning: Theory, research, and applications* (pp. 267-295). New York, NY: Routledge.

Zimmerman, B. J., & Kitsantas, A. (2002). Acquiring writing revision and self-regulatory skill through observation and emulation. *Journal of Educational Psychology, 94*, 660-668.

Zimmerman, B. J., & Risemberg, R. (1997). Becoming a self-regulated writer: A social cognitive perspective. *Contemporary Educational Psychology, 22*, 73-101. doi: 10.1006/ceps.1997.0919

Table 1

Demographic data for the participating students by condition

	Condition		
Measure	Planning	Sentence Combining	Control
Grade (*N*s)			
Fifth	24	20	25
Sixth	24	19	14
Gender (*N*s)			
Girls	27	15	20
Boys	21	24	29
Age (in years)			
M (*SD*)	11.2 (0.7)	11.1 (0.6)	11.2 (0.6)
Min-Max	10.2–13.7	10.2–12.6	10.1–12.6
Mother's educational level (%)			
Grade 4 or below	4	10	5
Grade 9 or below	15	18	31
High school	44	46	28
College or above	31	21	26
Unknown	6	5	10
School marks (1-5)			
$M_{\text{Portuguese}}$ (*SD*)	3.6 (0.9)	3.5 (1.0)	3.1 (0.8)
$M_{\text{Mathematics}}$ (*SD*)	3.7 (0.8)	3.7 (1.1)	3.3 (0.9)

Note. For school marks, 1 = *lowest score* and 5 = *highest score*.

Table 2

Means (and standard deviations) for all measures in each condition by testing time

	Pretest			Midtest			Posttest		
Measure	Planning	Sentence Combining	Control	Planning	Sentence Combining	Control	Planning	Sentence Combining	Control
Strategy specific									
Planning	2.58 (1.38)	2.38 (1.29)	2.00 (0.92)	3.73 (1.63)	2.08 (0.81)	1.74 (0.75)	5.42 (0.74)	1.97 (1.11)	1.77 (0.99)
Sentence combining	2.94 (1.87)	3.28 (1.64)	2.92 (1.98)	2.54 (1.75)	3.67 (1.68)	2.56 (1.79)	2.75 (1.97)	4.54 (1.86)	2.51 (1.97)
Writing performance									
Quality	3.97 (1.16)	3.75 (1.34)	3.70 (0.91)	4.26 (1.05)	3.97 (1.23)	3.23 (1.04)	4.59 (1.10)	4.39 (1.49)	3.47 (1.03)
Text length	99.90 (45.96)	67.56 (34.36)	80.28 (34.78)	118.29 (37.77)	93.23 (35.65)	63.95 (38.85)	131.88 (42.50)	116.69 (33.97)	67.67 (32.09)
Text length (adjusted *M*)				110.72	102.40	66.71	126.74	124.86	68.56
Discourse level									
Premise	1.46 (0.50)	1.33 (0.53)	1.23 (0.43)	1.69 (0.51)	1.72 (0.46)	1.54 (0.56)	1.71 (0.46)	1.44 (0.50)	1.33 (0.48)
Reasons	1.71 (0.80)	1.92 (0.93)	1.85 (0.84)	2.77 (0.91)	2.15 (1.18)	1.62 (0.91)	3.02 (0.60)	2.44 (1.14)	1.74 (0.91)
Elaborations	0.65 (0.64)	0.51 (0.64)	0.74 (0.55)	1.19 (0.98)	0.79 (0.83)	0.64 (0.63)	1.73 (1.13)	0.85 (0.59)	0.72 (0.72)
Conclusion	0.65 (0.73)	0.28 (0.61)	0.46 (0.68)	1.19 (0.87)	1.10 (0.82)	0.54 (0.76)	1.56 (0.74)	1.44 (0.75)	0.36 (0.67)
Coherence	0.85 (0.15)	0.87 (0.13)	0.89 (0.14)	0.93 (0.11)	0.91 (0.12)	0.93 (0.13)	0.96 (0.07)	0.89 (0.11)	0.87 (0.15)
Sentence level									
Ratio of connective clauses	0.42	0.48	0.44	0.48	0.49	0.44	0.47	0.48	0.47

	(0.18)	(0.22)	(0.16)	(0.17)	(0.16)	(0.23)	(0.14)	(0.13)	(0.19)
Variety of connectives	3.19 (1.18)	2.79 (1.20)	2.87 (1.03)	3.31 (1.15)	3.72 (1.12)	2.49 (1.21)	3.83 (0.86)	5.82 (2.26)	2.49 (0.76)
Ratio of opinion clauses	0.16 (0.14)	0.14 (0.11)	0.18 (0.12)	0.17 (0.09)	0.23 (0.16)	0.20 (0.15)	0.23 (0.13)	0.40 (0.14)	0.14 (0.14)
Variety of opinion markers	1.33 (0.83)	1.15 (0.84)	1.51 (0.88)	1.88 (1.10)	2.69 (1.89)	1.15 (0.63)	2.58 (1.35)	5.36 (1.40)	1.13 (0.83)
Clause length	6.70 (1.19)	6.26 (1.30)	6.77 (1.12)	6.69 (1.11)	7.00 (1.37)	6.60 (1.27)	6.21 (0.82)	7.22 (1.16)	6.21 (1.19)
Word level									
CTTR	4.43 (0.71)	4.10 (0.73)	4.11 (0.64)	4.46 (0.60)	4.54 (0.67)	4.03 (0.65)	4.53 (0.75)	4.90 (0.67)	4.16 (0.63)
Ratio of modifiers	0.02 (0.02)	0.02 (0.02)	0.02 (0.02)	0.02 (0.02)	0.03 (0.02)	0.03 (0.03)	0.04 (0.03)	0.05 (0.02)	0.04 (0.03)
Variety of modifiers	1.69 (1.68)	1.38 (1.29)	1.26 (1.19)	2.06 (1.67)	2.64 (2.24)	1.77 (2.01)	4.08 (2.43)	5.26 (2.98)	2.46 (1.85)
Self-efficacy	69.65 (17.78)	72.46 (16.79)	63.92 (26.92)	–	–	–	79.05 (14.75)	78.41 (14.72)	72.97 (21.02)
Summary									
Sensitivity to importance	3.31 (1.64)	3.46 (1.94)	3.41 (2.25)	–	–	–	4.50 (1.56)	3.36 (1.98)	3.38 (2.17)
Sentences transformation	0.12 (0.23)	0.08 (0.20)	0.07 (0.15)	–	–	–	0.14 (0.28)	0.17 (0.26)	0.14 (0.23)

Note. See Method's section for description of measures.

Table 3

Results of the 3 (condition) x 3 (testing time) repeated measures ANOVAs

	ME of condition	ME of testing time		Interaction		
Measure	$F(2, 123)$	Λ	$F(2, 122)$	Λ	$F(4, 244)$	η^2_p
Strategy specific						
Planning	85.38***	.78	17.39***	.45	30.16***	0.33
Sentence combining	8.37***	.97	1.74	.88	4.07**	0.06
Writing performance						
Quality	8.27***	.91	6.00**	.87	4.31**	0.07
Text length[a]						
Discourse level						
Premise	7.11***	.83	12.83***	.96	1.27	0.02
Reasons	17.43***	.80	15.41***	.76	8.88***	0.13
Elaborations	13.23***	.83	12.39***	.82	6.54***	0.10
Conclusion	23.45***	.64	35.07***	.74	10.06***	0.14
Coherence	1.56	.92	5.59**	.89	3.82**	0.06
Sentence level						
Ratio of connective clauses	< 1	.99	< 1	.99	< 1	0.01
Variety of connectives	31.57***	.67	30.04***	.59	18.36***	0.23
Ratio of opinion clauses	11.44***	.77	18.05***	.68	12.89***	0.17
Variety of opinion markers	51.26***	.41	89.52***	.36	40.22***	0.40
Clause length	2.06	.98	1.52	.83	5.78***	0.09
Word level						
CTTR	7.32***	.85	10.73***	.83	6.07***	0.09
Ratio of modifiers	3.05	.58	44.05***	.98	< 1	0.01
Variety of modifiers	8.00***	.51	59.00***	.84	5.58***	0.08
Self-efficacy[b]	1.83	.76	39.68***	.99	< 1	0.01
Summary[b]						
Sensitivity to importance	1.57	.98	3.03	.93	4.54**	0.07
Sentences transformation	< 1	.96	5.21*	.99	< 1	0.01

[a]Instructional effects were calculated with ANCOVAs (see section 3.2.2.). [b]As data was collected only at pretest and posttest, degrees of freedom for condition, testing time, and the interaction were $F(2, 123)$, $F(1, 123)$ and $F(2, 123)$, respectively. ME = main effect.

* $p < .05$. ** $p < .01$. *** $p < .001$.

Table 4

Effect sizes (Cohen's d) computed for statistical significant pairwise comparisons (α = .017) between conditions at midtest and posttest

	Planning vs. Sentence Combining		Planning vs. Control		Sentence Combining vs. Control	
Measure	Midtest	Posttest	Midtest	Posttest	Midtest	Posttest
Strategy-specific						
Planning	1.28	3.66	1.57	4.18		
Sentence combining	-0.66	-0.93			0.64	1.06
Writing performance						
Quality			0.99	1.05	0.65	0.72
Text length[a]			1.15	1.55	0.96	1.70
Discourse level						
Reasons	0.59	0.64	1.26	1.66		0.68
Elaborations		0.98	0.67	1.07		
Conclusion			0.80	1.70	0.71	1.52
Coherence		0.76		0.77		
Sentence level						
Variety of connectives		-1.16	0.69	1.65	1.06	1.98
Ratio of opinion clauses		-1.26		0.67		1.86
Variety of opinion markers	-0.52	-2.02	0.81	1.29	1.09	3.68
Clause length		-1.01				0.86
Word level						
CTTR		-0.52	0.69	0.53	0.77	1.14
Variety of modifiers				0.75		1.13
Summary						
Sensitivity to importance	–	0.64	–	0.59	–	

Note. Effect sizes for premise, number of modifiers, self-efficacy, and sentences transformation were not presented because the Condition x Testing Time interaction was not significant.

[a]Effect sizes calculated from adjusted means.