# Exploring the effects of strategy-focused instruction in writing skills of 4TH grade students

**L. Rodríguez-Málaga** [1,2] · **M. Cueli** [2] · **C. Rodríguez** [2]

© Springer Science+Business Media, LLC, part of Springer Nature 2020

## Abstract

The improvement of writing skills is one of the aims of the educational system. The present study implemented strategy-focused instruction through the Cognitive Self-Regulation Instruction (CSRI) program which included three instructional components: Direct Teaching, Modelling, and Peer-Practice. The aim was to explore the short- and long-term effects of different sequences of the three CSRI instructional components on writing skills (reflected in greater Coherence, Structure and Quality of the text product). Six 4th grade primary education classes from three different schools ($N = 126$) were randomly allocated to the experimental conditions which differed in the order in which the instructional components were implemented: Direct Teaching, Peer-Practice, and Modelling (experimental condition 1, $n = 47$); or Modelling, Peer-Practice, and Direct Teaching (experimental condition 2, $n = 36$). A control condition ($n = 43$) was also included in which students received traditional instruction. Writing performance was measured through compare-contrast writing tasks. In addition, writing performance was assessed through the anchor text procedure (considering three variables: Coherence, Structure and Quality). Findings supported the effectiveness of strategy-focused instruction after four sessions of the CSRI program. Experimental conditions 1 and 2 showed a significant gain relative to the control condition and were equally effective for the improvement of writing skills over a short-term period. More specifically, the Peer-Practice component was associated with the largest gains in the 4th grade students' writing skills.

**Keywords** Writing skills · Strategy-focused instruction · Direct Teaching · Modelling

✉ C. Rodríguez
  rodriguezcelestino@uniovi.es

1   Department of Psychology, Sociology and Philosophy, Faculty of Education, University of León, León, Spain

2   Department of Psychology, Faculty of Psychology, University of Oviedo, Plaza Feijoo s/n, 33003 Oviedo, Asturias, Spain

## Introduction

Writing skills constitute a basic and fundamental condition of participation both in society and academic environments (Graham 2018). Becoming an effective writer demands not only mastery of transcription skills (e.g., handwriting and spelling) but also the deployment, executive control, and self-regulation of high level cognitive processes such as planning, drafting, and modifying ideas according to their correct linguistic form as well as revising and editing (e.g., Hayes 2012; van den Bergh et al. 2016).

However, given its multidimensional nature, mastering writing is a significant problem for new writers who face a "double challenge" (Rijlaarsdam and Couzijn 2000). According to these authors, students must perform two tasks during the writing activity: writing and learning to write. As writers, the aim is to produce text with the three main characteristics of Structure, Coherence, and Quality. As learners, the students' objective is to learn to write and consequently acquire writing skills (transcription skills, planning, revision and so forth).

In this sense, if students do not have sufficiently automatic transcription skills, they devote most of their cognitive resources to text production (e.g., word choice, spelling, sentence construction) and fewer resources to higher-level cognitive processes (e.g., planning, evaluation, review; Limpo and Alves 2018; McCutchen 2011; Rijlaarsdam et al. 2011). As Olive and Kellogg (2002) showed in their study with 3rd graders, when handwriting is not yet automatic, the process of transcription demands working memory resources, leaving few resources available for other high-level cognitive processes. Even in older students (upper-primary and high school), some studies have shown that students spend little time planning or reviewing and the use of these high level of cognitive processes is inefficient resulting in poor quality text (Albertson and Billingsley 2000; Beauvais et al. 2011; Limpo et al. 2014; López et al. 2019). Therefore, although it is important for students in all grades to receive effective instruction in writing skills, providing it in the first few years of primary education would prevent future difficulties in higher grades (Arrimada et al. 2019). In this context, different approaches and strategies have been developed for improving students' written skills. The present study focuses on examining the benefits of a specific strategy aimed at improving written skills, reflected in the production of text with Structure, Coherence and Quality.

One of the most effective approaches for achieving management of writing skills is strategy-focused instruction; alone or combined with self-regulation procedures (see the meta-analyses by Graham and Harris 2017; Graham et al. 2012). According to Zimmerman's Model (2000), self-regulation is a cyclical process involving three phases: a) a forethought phase allows students to set goals, analyze the task and plan strategies to achieve it; b) a performance phase includes self-control (meta-cognitive knowledge and specific strategies) and self-observation (self-recording and metacognitive monitoring); c) a self-reflection phase (after performance) refers to the evaluation of the performance and includes self-evaluation and self-reactions (Zimmerman 2013).

The most widely and successfully used approach combining strategy-focused instruction and self-regulation is the Self-Regulated Strategy Development model (SRSD; Graham and Harris 2017). SRSD is a strategy-focused instruction approach aimed at teaching students about writing and the strategies involved in the writing process (e.g., skills for planning, monitoring and reviewing their own writing process) (Graham et al. 2015). The benefits of programs developed using the SRSD model have been well demonstrated in different educational stages and in students both with and without learning difficulties (e.g., Brunstein and Glaser 2011; Gillespie and Kiuhara 2017; Limpo and Alves 2013a; McKeown et al. 2019;

Saddler et al. 2019). One of these programs is *Cognitive Self-Regulation Instruction* (CSRI) which was shown to be effective in improving the writing skills of 6th grade primary school students in regular classroom contexts (see review Fidalgo and Torrance 2017).

## Cognitive self-regulation instruction (CSRI)

CSRI was developed based on SRSD and Zimmerman's Model (see Harris and Graham 2017). The purpose of CSRI is to develop higher-levels of writing skills in upper-primary students (6th grade). With this aim, CSRI identified effective evidence-based practices including three instructional components: Direct Teaching, Modelling, and Peer-Practice (e.g., Graham and Perin 2007; Graham et al. 2012).

In Direct Teaching (before performance-task) the teacher focuses on teaching students metacognitive knowledge (different text types, characteristics of a quality text, and explicit knowledge about planning and drafting), self-regulation phases (planning, monitoring and review), and metacognitive strategies (related to planning and drafting through mnemonic rules). Meta-analyses (Graham and Harris 2017; Koster et al. 2015) provide evidence of the importance of teaching students metacognitive knowledge about the writing process and cognitive strategies. This knowledge will guide students through the key steps that the teacher will carry out in the subsequent component and will encourage correct consideration during the first phase of the self-regulation model.

The Modelling component consists of giving the students a model to copy (the teacher) who exemplifies the correct steps of the writing processes, the typical mistakes students make, and how to correct them. Through the Modelling component, students identify, analyze and evaluate the necessary steps that they will have to follow when writing their own texts. Research has shown that copying models (in contrast to mastery models) allow students to learn more and get better results (Braaksma et al. 2002; Zimmerman and Kitsantas 2002). The benefits of observational learning are well documented in the literature (e.g., Braaksma et al. 2010; van de Weijer et al. 2019). Through observation, students address all of their cognitive resources towards observing the model, who explains (how to write) and demonstrates the process (performs the writing task), stimulating the second phase of the self-regulation model, performance. To do this, the teacher uses thinking aloud protocols (a metacognitive technique that consists of the person doing the activity continuously verbalizing their thoughts while they do the activity; Montague et al. 2011). The inclusion of thinking aloud provides several advantages to both the teacher and the student. The teacher can evaluate what the students think and know and how they know it. Therefore, the teacher can give feedback about the writing process carried out by the student. For the student, the act of writing with thinking aloud can help them to achieve greater awareness and control over their own knowledge and skills. During composition, the students have to focus their attention on each step of the writing process displaying self-regulation skills such as self-control and self-monitoring (Bai 2018; Hartman 2001).

The Peer-Practice component involves the students emulating the process performed by the teacher (model). Several studies have shown the effectiveness of Peer-Practice on writing performance (e.g., Boon 2016; Grenner et al. 2018; Grünke et al. 2019; Wigglesworth and Storch 2012). In this component, students can practice their own writing (using the thinking aloud technique) in a supportive context where the instructor guides students, clarifying, monitoring, and reinforcing (Graham et al. 2013; MacArthur 2017). Through thinking aloud, students externalize the process, allowing their peers and the teacher to comment and provide

scaffolding to guide the target behavior, while at the same time encouraging the third phase of the self-regulation model, self-reflection (Zimmerman 2013).

Despite the effectiveness of strategy-focused instruction in general (e.g., Bouwer et al. 2018; Koster et al. 2017) and SRSD in particular (e.g., Festas et al. 2015; Limpo and Alves 2013b; Palermo and Thomson 2018; Rosário et al. 2019; Saddler et al. 2019) there are two gaps in the literature about writing instruction. First, the different strategy-focused instruction programs are multicomponent in nature, combining different instructional content (i.e., knowledge linked to the text product, metacognitive knowledge about the process with mnemonic rules) and instructional components (Direct Teaching, Modelling, Peer-Practice; Fidalgo et al., 2017b). This multicomponent nature prevents us from understanding which of these instructional components or what combination of them is key for achieving better results in writing instruction (De la Paz 2007; MacArthur 2017). Identifying the main components and the most effective combinations of components could make it easier to include strategy-focused instruction in the normal curriculum. Teachers rarely adopt this practice for reasons of time, the number of students in the class, or the challenge associated with the implementation of such an unusual component as Modelling without the support of researchers (Kistner et al. 2010).

In this regard (the importance of identifying which components and what combinations are essential), some studies have focused on exploring and comparing the efficacy of the instructional components of strategy-focused instruction. For example, Fidalgo et al. (2015) carried out a study with three $6^{th}$ grade classes ($N = 62$) to assess the cumulative effects of a sequence of different instructional components of CSRI: (1) observation of a mastery model, (2) direct teaching, (3) peer feedback, and (4) only practice. Their results indicated that students who received the CSRI program improved the Structure, Coherence, and Quality of their written texts. This improvement was associated exclusively after the initial component, that is, the Modelling component (observation of a correct example of application of the strategies), without any type of Direct Teaching. The authors concluded that observation of a model followed by reflection promoted writing skills. However, Fidalgo et al. (2015) used a design in which the strategy-focused instruction always began with the Direct Teaching component.

A direct comparison of the benefits of Modelling and Direct Teaching was analyzed by López et al. (2017). Those authors randomly assigned 133 upper-primary students (10 to 12 years old) to one of two experimental conditions: Direct Teaching (students received explicit declarative knowledge of planning and drafting strategies), and Modelling (the teacher provided procedural knowledge of how to implement planning and drafting strategies). Their findings suggested similar improvements in both experimental conditions in all measures (Structure, Coherence and Quality of texts).

On similar lines, De Smedt and Van Keer (2018) investigated the combined effectiveness of explicit instruction and writing with peer-assistance in a study involving 206 fifth and sixth-grade students. These students were randomly assigned to one of four experimental conditions: (a) explicit instruction + individual writing; (b) explicit instruction + writing with peer assistance; (c) matched individual-practice comparison condition; and (d) matched peer-assisted practice comparison condition. The authors found that the first group of students outperformed the matched individual-practice students at posttest, highlighting the importance of explicit instruction. Furthermore, De Smedt and Van Keer (2018) did not find differences when comparing the individual writing and peer-assisted conditions.

A limitation of the studies noted above was that they used students in the final years of primary education, despite the aforementioned importance of providing effective instruction to

young children. Another significant gap in the literature is the previously noted absence of studies looking at the long-term effects of strategy-focused instruction. In this regard, Hacker et al. (2015) stated that few studies included maintenance measures, and most did not exceed eight weeks after the end of intervention. As de Boer et al. (2018) argued, metacognitive knowledge and strategies may need more time to carry out or demonstrate greater effectiveness in the long-term. Because the goal of teaching students strategies is to encourage effective autonomous learning over the long-term, (not just during an intervention), the study of maintenance strategies becomes a key focus for educators.

The present study addressed these limitations by examining the short- and long-term effects of different sequences of the instructional components of CSRI on the writing skills of 4th grade students (aged between 9 and 10 years old). In order to do this, we designed two experimental conditions that differed in the sequence of instruction components. In experimental condition 1, the students first received the Direct Teaching component, followed by Peer-Practice, followed by Modelling. In contrast, in experimental condition 2, students first received the Modelling component, followed by Peer-Practice, followed by Direct Teaching. Similarly to López et al. (2017), we took both Modelling and Direct Teaching as starting points because writing skills can be improved by providing a framework of specific strategies (Direct Teaching) that guides attention to the key steps of the writing process performed by a model (Modelling) and the specific steps to follow (Peer-Practice). In this sense, prior research has already shown the positive effects of Direct Teaching about how to write and to plan and revise (e.g., Bouwer et al. 2018; Rietdijk et al. 2017). On the other hand, according to Fidalgo et al. (2015) it may be enough to provide students with the Modelling component to achieve writing performance. In this regard, successful interventions have included the Modelling component as an instructional mode to teaching writing strategies (Braaksma et al. 2010; Koster et al. 2017). Moreover, our design also explored whether the Peer-Practice component can have a key role in students' learning following the Direct Teaching component (De Smedt and Van Keer 2018), due to their complementary nature and the opportunity students have to learn from each other (Graham 2018), or whether the Peer-Practice component is unneccesary.

## Present study

The aim of this study was to explore: a) the short-term and long-term effectiveness of strategy-focused instruction (6 sessions) on the writing skills of 4th grade students; and b) the component, or order of implementation, that had the greatest effect on writing skills based on the cumulative effects of the instructional components.

The effectiveness of the strategy-focused instruction was assessed in the short-term at 4 time points: at the beginning of the intervention (Pretest), after the end (Posttest), after each component (Test 1 and Test 2). There was also an assessment eight months after finishing the intervention (Follow-Up) to check which group maintained the effects (or if there was any effect) of strategy-focused instruction in the long-term.

Our hypothesis was that strategy-focused instruction (CSRI) would be beneficial to writing skills (reflected in greater Structure, Coherence and Quality of the text product) over and above that afforded by traditional instruction, in both the short-term (Pre/Posttest), and the long-term (eight months after finishing the training). The writing from the students in the control

condition would be less coherent, less well structured, and of poorer quality than the writing from students in the experimental conditions. The study of the effect of these kinds of strategies on written skills will have an impact on classroom practice and on future teacher recommendations.

## Method

### Participants and design

The sample comprised a total of six classes in 4th grade primary education ($N = 126$) from three different mixed state- and privately-funded schools in the North of Spain. Students' ages ranged from 9 to 10 years old. Participants' details are shown in Table 1. The initial sample comprised 153 students. Sixteen students who had existing diagnoses of learning disabilities and special educational needs were not included in the statistical analysis although they did receive the same instruction as their peers. Furthermore, 11 students who did not attend all the evaluation time points were not considered in the statistical analysis [attrition rates of 0.72% were observed at Test 1 ($n = 1$); 2.20% at Test 2 ($n = 3$); 0.75% at Posttest ($n = 1$); and 4.54% at Follow-Up ($n = 6$)].

Prior to the intervention, all students followed the ordinary Spanish school education curriculum (focused mainly on teaching text production and rules for correct spelling and grammar).

The six classes were randomly assigned to one of the two experimental conditions or to a control condition, giving two classes in each condition. Both experimental conditions received the strategy-focused instruction (CSRI) but differed in the order in which instructional components were implemented: Direct Teaching, Practice and Modelling (experimental condition 1); Modelling, Practice and Direct Teaching (experimental condition 2). In the control condition students received examples of high-quality texts without any strategy-instruction.

The effect of the instruction was assessed by the change in the written skills across 5 time points: immediately prior to intervention (Pretest), following each component (Test 1 and Test 2), at the end of the intervention (Posttest), and eight months after finishing the intervention (Follow-Up). The design of the evaluation is outlined in Table 2.

### Training: CSRI program

The strategy-focused instruction used was the CSRI program (developed under the SRSD model; Harris and Graham 2017). This program is aimed at developing student strategies for setting appropriate production goals and process strategies focusing on expository compare-

**Table 1** Descriptive characteristics of the participants for the three conditions

|  | $N = 126$ | $n$ male ($n$ female) | Mean age ($SD$) |
|---|---|---|---|
| Experimental condition 1 | 47 | 26 (21) | 9.36 (0.48) |
| Experimental condition 2 | 36 | 16 (20) | 9.56 (0.58) |
| Control condition | 43 | 21 (22) | 9.27 (0.46) |

$SD$ = Standard Deviation

Table 2 Research design showing instructional sequences and writing assessment probe for each condition

| Condition | Week 1–2 | | | Week 3–4 | | Week 5–6 | | 8 months later |
|---|---|---|---|---|---|---|---|---|
| EC 1 | PRE | Direct Teaching | Test1 | Peer-Practice | Test2 | Modelling | POST | Follow-Up |
| EC 2 | PRE | Modelling | Test1 | Peer-Practice | Test2 | Direct Teaching | POST | Follow-Up |
| CC | PRE | Control | Test1 | Control | Test2 | Control | POST | Follow-Up |

EC1 = Experimental condition 1; EC2 = Experimental condition 2; CC = Control condition; PRE = Pretest; POST = Posttest

contrast texts (see Fidalgo et al., 2017a). Specifically, the CSRI attempts to develop students' strategies for planning and drafting appropriate compare-contrast texts. The program had three components (Direct Teaching, Modelling and Peer-Practice) with two sessions for each component (there were a total of six sessions, following a schedule of one session of 1 h per week). The three components varied in terms of how the content was delivered. The content of the CSRI program and the content of the control condition were both controlled so that the students received the same amount of instruction in Structure, Coherence and textual Quality in 1 h (the normal duration of the intervention session). It was emphasized that the teachers had to adhere to the step-by-step content scripted by the researchers, and that they could not repeat the content or divide the session into several parts (nor change the focus of the instruction according to their own criteria). In the same way, the students were not allowed to take the work material to study or memorize at home. Moreover, we also checked that at the beginning of the sessions all the teachers reminded students of the content taught in the previous sessions. In this way, no student started the session at a different level and with different knowledge from the rest of the class.

The features of the three training conditions are described in Table 3 and are summarized below.

Table 3 Summary of features of the three training conditions

| | Condition 1 | Condition 2 | Control |
|---|---|---|---|
| **Instructional content** | | | |
| Product goals (the oaiue mnemonic rule) | + | + | − |
| Planning and drafting writing strategies | + | + | − |
| Structural and linguistic features of compare–contrast essays | + | + | + |
| **Instructional approach** | | | |
| Direct Teaching of cognitive writing strategies through mnemonics | + | + | − |
| Modelling of the planning and drafting processes through thinking aloud (without explicit reference to the cognitive writing strategies) | − | + | − |
| Modelling of the planning and drafting processes through thinking aloud (with explicit reference to the cognitive writing strategies) | + | − | − |
| Analysis of high-quality compare-contrast texts | + | + | + |
| Practice in pairs | + | + | + |
| Motivation support | + | + | + |
| **Kind of knowledge provided** | | | |
| Self-Regulated approach | + | + | − |
| Declarative knowledge | + | + | + |
| Procedural knowledge | + | + | − |

**The first experimental condition** included the three components in the sequence: Direct Teaching, Peer-Practice and Modelling. The Direct Teaching component comprised two sessions in which students were introduced to the strategy planning process (Session 1) and then to the drafting process (Session 2). At the beginning of Session 1, the instructor taught a metacognitive matrix identifying the nature, purpose and central features of effective planning processes. Then students were introduced to the mnemonic POD + the vowels OAIUE, as a scaffold for planning their compare-contrast texts. POD stands for each of the steps: (1) *Think of ideas -Pensar-* before writing; (2) *Organize* your thoughts with the OAIUE mnemonic rule [O *(Objective)* prompted students to identify the purpose of different text types; A *(Audience)* prompted students to capture future readers' interest and attention, motivate them to read, and make it easier to understand; I *(Ideas)* prompted students to think of ideas, brainstorm or search for other documentary sources, and differentiate between main and secondary ideas and examples; U *(Union)* reminded students to connect the ideas in the text, joining thematic ideas (e.g., similarities vs. differences in comparison - contrast text); E *(Esquema-Plan)* reminded students to make a *plan* including ideas about the introduction, development and conclusion]; and (3) *Develop* the text (Fig. 1).

In the second session of the Direct Teaching component, students were given a metacognitive matrix identifying the nature, purpose and central features of effective drafting processes. Then, the instructor taught the mnemonic IDC + the vowels OAIUE (see Fig. 1). IDC encourages the organization and structure of a comparison-contrast text: (1) *Introduction* in which students should present the topic, the purpose of the text and capture the readers' interest; (2) *Development* in which students were instructed to develop the ideas and examples to explain these ideas and; (3) *Conclusion* which reminded students to make a personal contribution to the text, an overall point of view, or reflection of everything discussed in the text. Again, the vowels provided criteria about the content during all three of the IDC
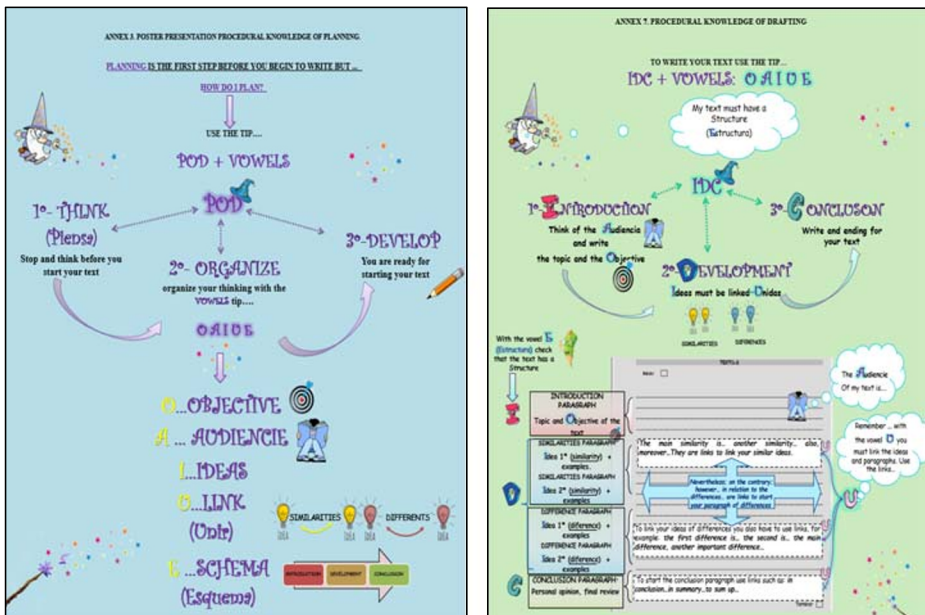


Fig. 1 Example of graphic organizer of POD + the vowels OAIUE and IDC + the vowels OAIUE

production phases. Both strategies were supported by illustrated summaries showing the POD and IDC+ vowels mnemonics to facilitate students' learning.

The Peer-Practice component required students to work in pairs planning (Session 3) and drafting (Session 4) a compare-contrast text. The first session (Session 3) started with the instructor reminding students of the content of previous sessions and giving students some practice in thinking aloud. The instructor selected students with similar abilities and paired them. Students were assigned to writer or helper roles. The more extrovert student or the one more likely to do the thinking aloud was the writer and the other was the helper. These roles were maintained throughout the intervention. During composition, thinking aloud helped expose the writing processes adopted by the writer to the helper's observation and comments. The students' writing processes were scaffolded by graphical organizers. While the pairs were planning their texts, the students had in front of them a sheet with spaces for the student's own notes in the margins laid out following the Vowels criteria (so students would not forget to do it). In the second session of this component (Session 4), the writer took the outline created in the previous session and translated it into text. This session then followed an identical pattern to Session 3, with a focus on the IDC mnemonic. In both sessions, during thinking aloud composing, the instructor also patrolled the class, listening to the thinking aloud and providing feedback and help for the writer about how to perform the thinking aloud and apply the strategies taught.

Finally, in the Modelling component the teacher demonstrated the correct writing process steps for the planning strategy (Session 5) and drafting strategy (Session 6) previously explained in the Direct Teaching component. Modelling involved thinking aloud while composing a compare-contrast text in front of the class. Thinking aloud was mainly scripted. The teacher emphasized explicit references to the strategy with a self-regulatory approach to the task (*The first thing I am going to do is, as the letter E -Esquema- says, a Plan of my text .... But first of all, it is important that everything in my plan is done thinking of the letter A*), and with self-statements about positive expectations (*If I make a last effort I will do it, I have good ideas, I will continue like this!*) producing a written plan (Session 5) and draft text (Session 6). The teacher explained to the students that during the Modelling they had to concentrate on all the teacher's steps and thoughts during the writing process. After Modelling, students made notes about the model's most important thoughts. Then, the instructor facilitated a whole-class discussion, drawing together the students' observations. At the end, each student individually wrote down reflections about the differences between their own writing practice and the processes that they had seen.

**The second experimental condition** included the sequence of components: Modelling, Peer-Practice and Direct Teaching. First, the two sessions of the Modelling component followed the same pattern as the Modelling component in experimental condition 1, with the exception that the explicit reference to the strategy and mnemonic rules associated was removed. In this case, the teacher did a Modelling following the steps of the planning process (session 1) and drafting process (Session 2) with thinking aloud in front of the class. The teacher emphasized explicit references to the process with a self-regulatory approach to the task (*The first thing that I am going to do is an outline of my text (...). But first of all, it is important that I do my whole outline thinking about who is going to read the text*), and self-statements about positive expectations to produce the written plan (Session 1) and draft text (Session 2). Again, after Modelling, students made notes about the model's most important thoughts to compare with their own writing practice and discussed it with the class-group. Second, the Peer-Practice component sessions had one difference compared to experimental

condition 1. In this case, students were able to watch the teacher Modelling without explicit mention of the strategy or mnemonic rules. Students had to emulate the writing processes that they had seen modeled. During emulation, as in experimental condition 1, students worked in writer-observer pairs with thinking aloud planning their compare-contrast texts (Session 3) and translated the plan to a complete text (Session 4). During emulation, the teacher provided direct input for the writer (including prompts to do the thinking aloud if they forgot) and provided a model for the observer. Finally, in the Direct Teaching, as in experimental condition 1, the students were introduced to the concepts underlying the POD + OAIUE mnemonic rule (Session 5) and the IDC + OAIUE mnemonic rules (Session 6).

**The control condition** was production-focused without any strategy instruction (i.e., without teaching explicit strategies for setting product goals) but with the same level of practice as the experimental conditions. This involved instruction relating to the structural and linguistic features of the compare-contrast text. The instructional program is described briefly below.

In the first sessions (1 to 3) the instructor focused on teaching different types of texts and their characteristics. The instructor started the first session with a brainstorm on the importance and value of a good writing. After that, the instructor taught the objectives of three text types (narrative, descriptive and compare-contrast text). Then, the instructor presented the students with two tasks in which they had to identify textual examples for each of the previously taught text types. In Sessions 2 and 3 the instructor presented the structure and production of each of the three types in detail. To make it easier for students to memorize content, the teacher gave the students different text types to identify and analyze the characteristics of.

In the last few Sessions (4 to 6) the instructor focused only on the compare-contrast type text. In the Sessions 4 and 5 students wrote a compare-contrast text in pairs (one text in each session) aiming to emulate the type-specific features of the example explained in the previous sessions. The instructor put students with similar abilities into pairs for the collaborative writing task. Students were assigned to writer or helper roles. The more extrovert student or the one more likely to do thinking aloud was the writer and the other was the helper. Once they had written the text, each of the pairs read their texts aloud so that the instructor and the class group could provide feedback on whether the text had the required characteristics of a compare-contrast text. In Session 6 the instructor gave students an incomplete compare-contrast text. The students were divided into groups and completed the missing parts of the text without the help of the instructor. After the task was finished, the instructor guided a group discussion about whether the text was completed properly.

## Instruments and measures

**Writing assessment task** In the writing assessment task, students had to write a compare-contrast text based on the similarities and differences between various topics that were selected beforehand based on students' interests: Pretest (computer games versus traditional games), Test 1 (summer holidays versus winter holidays), Test 2 (book versus film), Posttest (travel by car versus travel by plane), Follow-Up (superhero versus ordinary human).

**Evaluation session** The evaluation session lasted 1 h in which the researcher (an educational psychologist who was a member of the research group) gave the students small cards which included the title of the topic (for example traditional games versus computer games) with a

picture about the topic (e.g., computer and football). Then students were asked to write a compare-contrast text. The researcher provided students with two work sheets, one for planning and one for the final text. Students were told that using the first work sheet was optional. The researcher reminded the students that they had 1 h to write their text and encouraged them to produce the best essay that they could. The researcher did not provide any help during the evaluation writing task. Students were evaluated at 5 time points: at the beginning of the intervention (Pretest), after the end (Posttest), after each component (Test 1 and Test 2), and eight months after finishing the intervention (Follow-Up). Thus, each student wrote five texts. A total of 603 texts were collected and assessed.

**Product assessment** Texts were analyzed using the anchor text procedure from Rietdijk et al. (2017). In this procedure, texts were rated in separate rounds for each dimension or measure (Structure, Coherence and Quality), for which specific criteria and definitions were provided to the evaluators (independent researchers of writing instruction and assessment). First, they chose a sample of 50 texts from the total. From these 50 texts, five anchor texts were identified representing scores at the mean and 1 and 2 standard deviations above and below the average: two weaker texts (scores 70–85), an average text which was assigned the arbitrary score of 100, and two better anchor texts (scores 115–130) [see Appendix 1. Example of anchor texts (scores 70 weaker text, 100 average, and 130 better text) and transcribed student texts]. Then two evaluators independently rated all of the texts in 3 separate rounds, one round per variable (Structure, Coherence and Quality) with the five anchor texts providing benchmarks. The mean score was arbitrarily set at 100, and the remaining anchor texts at 70, 85, 115, and 130. This procedure was repeated for each evaluation time point (Pretest, Test 1, Test 2, Posttest and Follow-Up). For the development of the benchmark rating scale the evaluators took examples from other studies (Bouwer et al. 2018) in tandem with their prior knowledge of other types of evaluation procedures involving information about concepts and assessing aspects related to the Structure, Coherence and Quality of the texts.

The mean inter-rater consistency (Pearson's $r$) of each variable across time point (Pretest, Test1, Test 2, Posttest and Follow-Up) was greater than .90. In addition, the intraclass correlation coefficient (ICC) estimates and their 95% confidence intervals were calculated based on a mean-rating ($k = 6$), consistency, 2-way mixed-effects model. The ICC demonstrated moderate reliability for Structure (ICC = .576), Coherence (ICC = .687), and Quality (ICC = .659). Under these conditions, ICC values less than 0.5 are indicative of poor reliability, values between 0.5 and 0.75 indicate moderate reliability, values between 0.75 and 0.9 indicate good reliability, and values greater than 0.9 indicate excellent reliability (Koo and Li 2016).

## Procedure

**Training delivery** The study was conducted during the spring school term. The sessions took place in literacy lessons. The full implementation of the program was carried out by six teachers (educational professionals with master's degrees in primary education), one for each class. The training of the teachers was a fundamental element through which we ensured the program was carried out accurately by establishing what the teachers had to do and how they had to do it. Previous research has indicated that training teachers in writing practices has not only a positive relationship with the quality of student texts (De Smedt et al. 2016) but also more positive feelings about writing instruction (Koster et al. 2017).

**Teacher training** Prior to the start of the intervention, a specialist researcher who guided the study methodology presented the CSRI program to the teachers covering general information, background, and implementation schedule. Then, in order to facilitate the implementation of the CSRI program, all the teachers were given the complete set of materials for each student (individual portfolios) and a "teacher session manual" containing detailed descriptions of the 6 sessions. The manual contained: (a) Instructions for how to start, carry out, and finish the session (e.g., *This session starts with a Modelling of the planning process. An example Modelling will be shown to the class group about what steps students should take before starting to write a compare-contrast text*); (b) The specific materials to be used for each step of the session and how to address the students (e.g., *You should read the appendix about the Modelling technique with thinking aloud. An example video is provided to give you a better understanding of the Modelling technique. Use the appendix about implementing Modelling to facilitate*); (c) Instructions about how to talk to the students and activities for them (e.g., *Explain to the students that during the Modelling they have to be very attentive, they have to stay quiet and focus on all the steps and thoughts that you perform during planning the text; give the students the appendix which has an example of planning*).

The researcher asked teachers to read the session information carefully before the start of training to discuss and clarify any questions during the training sessions. There were a total of three training sessions (one training session for the Direct Teaching component, another to support Modelling and a final session to support the Peer-Practice component). Each training session took place a week before its content was implemented. Teachers were trained individually by the specialist researcher and all sessions lasted for approximately 60–80 min following the same two-part pattern. In the first part of the session, the researcher started with an explanation of the specific component (*what is Direct Teaching?*) and its goal [e.g., *the aim of session 1 is for students to develop declarative knowledge (what is textual planning), procedural knowledge (how a text is planned, through a specific strategy), and conditional knowledge (when and why a text is planned)*]. In the second part of the session the researcher explained and discussed the steps described in the teacher's portfolio. Specifically, in the Modelling training session in addition to reviewing the portfolio, the researcher showed an example video of how the Modelling component should be implemented. In addition, an example of thinking aloud was provided for use in the instructional session (see Appendix 2) and was trained during the session. In this second part of the session, teachers were able to ask questions and resolve any issues they may have had about implementing the sessions. The researcher also met with teachers on a weekly basis during the intervention period to go over the details of that week's program and share their experience of the intervention sessions.

**Treatment fidelity** We used the following measures to ensure that the teachers implemented the program appropriately: First, all teachers were given manuals including the elements and activities for each session. Second, a specialist researcher met with the teachers weekly to train them in applying the instructional procedures and to interview them about their experiences of the intervention sessions. Third, the student portfolios with the set of materials were reviewed following the sessions, allowing us to check whether the students had correctly completed the tasks, such as in the Modelling sessions, where students had to watch the teachers' thinking and actions and write them down. Data from the students' portfolios indicated completion of 92.4% of the tasks ($SD = 0.40$, range 80–90). Moreover, we made audio recordings of the intervention sessions from three of the six teachers. The first author listened to the intervention session recordings and noted whether each step or procedure was completed. The fidelity for these

teachers averaged 89.0% ($SD = 0.94$, range 80–100), 90.3% ($SD = 1.16$, range 80–100) 97.0% ($SD = 0.24$, range 80–100). For the other teachers (who did not consent to be recorded), we observed the implementation of three sessions (the first sessions for each component). Fidelity was evaluated with checklists that identified the specific steps of the session observed. Steps were coded as 1 (completed) or 0 (not completed). If teacher did complete the step, the performance was coded on a 10-point scale (a rating of 1 representing poor implementation and a score of 10 indicating high-quality implementation). Data from the observations indicated that teachers completed 92.9% of the steps ($SD = 0.40$, range 80–100). If teacher did not meet the study's strict fidelity requirements the data was excluded. To ensure a procedure that respects ethical standards, we sent a letter to the children's families in which they were informed of the objectives and nature of the study. We requested written informed consent from the families for their children's participation in the study. After the intervention, the strategy-focused instruction CSRI was delivered to each of the teachers in the control group. This procedure ensured that all participants had the opportunity to benefit from the strategic intervention.

## Data analysis

The data was analyzed using SPSS 24.0. The normal distribution of the three measures (Structure, Coherence and Quality) allowed us to conduct a parametric analysis. First, in preliminary analysis, the normal distribution and differences regarding Condition, Sex and Teacher-Class were analysed. The variable Teacher-Class was taken as a covariate. Considering the aims of the study, to determine the benefits of the CSRI program on every measure, different General Linear Models (GLMs) were used. Specifically, a one-way analysis of covariance (ANCOVA) was carried out for Structure, Coherence and Quality at each evaluation time point. The independent variable was the Condition (control condition, experimental condition 1 and experimental condition 2) and the dependent variables were the student performance in each measure (Structure, Coherence and Quality) at Pretest, Test1, Test2, Posttest and Follow-Up.

In order to assess the learning gain from the CSRI program more deeply, we analyzed the interaction between the condition by time point. We used one GLM (repeated measures ANCOVA model) for each measure at the time points Pretest-Posttest, Pretest-Test 1, Test 1-Test 2, Test 2-Posttest, and Posttest-Follow-Up. The independent variables were the evaluation time point and the condition, while the dependent variables were student performance in each measure (Structure, Coherence and Quality).

We used Bonferroni's multiple comparison test to determine the groups between which significant differences were found (post hoc Bonferroni comparison test, $p < .05/3 = .016$). Effect sizes were assessed using partial eta squared: $\eta_p^2 < 0.01$ = small effect, $\eta_p^2 \geq 0.59$ average effect; and $\eta_p^2 \geq 1.38$ = large effect (Cohen 1988).

## Results

### Preliminary compare-contrast text results

The asymmetry and kurtosis values of the variables were within the intervals that denote a normal distribution (Kline 2011). Table 4 summarizes the means and standard deviations for each variable by Condition and evaluation time.

Differences in Pretest variables were analyzed with regard to: Condition, sex, and Teacher-Class (given that the teachers varied across conditions). We carried out different Univariate Analysis of Variance (ANOVAs) taking the measure (Structure, Coherence, or Quality) as a dependent variable and the condition, sex, or Teacher-Class as independent variables. The results indicated that the differences were not significant for the condition groups in relation to Structure, $F(2, 123) = 1.076$, $p = .34$, $\eta_p^2 = 0.01$, Coherence, $F(2, 123) = 0.384$, $p = .68$, $\eta_p^2 = 0.00$, or Quality, $F(2, 123) = 1.277$, $p = .14$, $\eta_p^2 = 0.02$; nor by sex in Structure, $F(1, 124) = 3.091$, $p = .08$, $\eta_p^2 = 0.02$, Coherence, $F(1, 124) = 1.396$, $p = .24$, $\eta_p^2 = 0.01$, or Quality, $F(1, 124) = 2.142$, $p = .28$, $\eta_p^2 = 0.01$. However, at Pretest level we found differences regarding Teacher-Class in Structure, $F(5, 120) = 2.349$, $p = .04$, $\eta_p^2 = 0.089$; Coherence, $F(5, 120) = 4.274$, $p = .001$, $\eta_p^2 = 0.15$; and Quality, $F(5, 120) = 8.838$, $p < .001$, $\eta_p^2 = 0.26$. We included the Teacher-Class variable as covariate in all the analyses following.

## Condition effects

Table 4 shows the differences between the three conditions for each time point and in each variable (Structure, Coherence and Quality). The GLM consisted of ANCOVAs taking the measures in Structure, Coherence and Quality as dependent variables, and the Condition as an independent variable. With respect to the variables Structure and Coherence, the results indicated differences between the three conditions which were statistically significant at 3 time points, Test 1, Test 2 and Posttest, with small effect sizes (which were higher at Posttest than at Test 1 or Test 2). At Test 1, post hoc analysis did not show differences between the three conditions. However, in the case of Test 2 and Posttest, post hoc analyses indicated statistically significant differences between each of the experimental conditions compared to

Table 4 Means, standard deviations and results of the GLM (ANCOVAs) for Structure, Coherence and Quality variables by condition

|  | EC1 | EC2 | CC | $F(2, 122)$ | $\eta_p^2$ |
|---|---|---|---|---|---|
|  | *M(SD)* | *M(SD)* | *M(SD)* |  |  |
| Structure |  |  |  |  |  |
| Pretest | 80.80(6.75) | 80.88(7.42) | 78.93(6.73) | 1.03 | 0.01 |
| Test 1 | 89.97(11.92) | 85.41(11.52) | 86.11(10.58) | 7.82*** | 0.11 |
| Test 2 | 112.23(16.10) | 116.00(13.93) | 94.06(18.11) | 21.59*** | 0.26 |
| Posttest | 116.08(14.29) | 121.36(11.33) | 98.11(17.29) | 26.59*** | 0.30 |
| Follow up | 98.14(14.21) | 98.25(15.06) | 91.79(14.55) | 8.31 | 0.12 |
| Coherence |  |  |  |  |  |
| Pretest | 85.17(9.92) | 83.44(10.44) | 83.58(10.42) | 1.84 | 0.02 |
| Test 1 | 89.76(10.66) | 89.27(8.63) | 90.02 (10.42) | 5.92** | 0.08 |
| Test 2 | 106.63(13.52) | 112.16(16.76) | 93.72(14.88) | 14.35*** | 0.19 |
| Posttest | 109.00(12.91) | 114.08(12.53) | 93.09(15.66) | 23.30*** | 0.27 |
| Follow up | 96.53(12.84) | 100.33(12.44) | 93.00(13.82) | 2.97 | 0.04 |
| Quality |  |  |  |  |  |
| Pretest | 82.53(7.85) | 83.00(7.59) | 80.53(6.82) | 3.60* | 0.05 |
| Test 1 | 91.72(11.35) | 88.52(11.68) | 87.88(11.91) | 8.42*** | 0.12 |
| Test 2 | 109.00(13.95) | 114.38(13.61) | 98.79(16.40) | 9.611*** | 0.13 |
| Posttest | 112.93(13.05) | 116.97(11.67) | 96.16(16.74) | 24.07*** | 0.28 |
| Follow up | 97.17(12.86) | 99.77(13.12) | 93.00(13.82) | 3.268* | 0.05 |

EC1 = Experimental condition 1; EC2 = Experimental condition 2; CC = Control condition; $M$ = Mean; $SD$ = Standard Deviation; $\eta_p^2$ (eta-squared statistic) = estimates of effect size

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

the control condition for the three variables (experimental condition 1 vs. control condition $p < .001$; experimental condition 2 vs. control condition $p < .001$). In short, after four sessions of the CSRI program (Test 2) the three experimental conditions showed differences in the Coherence and Structure of their texts, reflecting better achievement from the two groups who received the intervention with no differences between them (experimental condition 1 and experimental condition 2). Effect sizes showed that differences were greater at Posttest (after six sessions of the CSRI program) followed by Test 2 (after four sessions of the CSRI program).

In terms of the variable Quality, the results of the ANCOVA showed statistically significant differences between the three conditions at each time point (see Table 4) with small effect sizes. Specifically, at Posttest the condition explained 28% of the variance; at Test 2, 13%; and at Test 1, 12%. Post hoc analyses showed that differences between the three conditions were not statistically significant at Pretest, Test 1 or Follow-Up. In the case of Test 2, post hoc analyses were statistically significant for the comparison experimental condition 1 and control condition ($p = .004$) and for the comparison experimental condition 2 and control condition ($p < .001$). Also, at Posttest, the post hoc analyses were significant for the comparison of each experimental condition with respect to the control condition (experimental condition 1 vs. control condition $p < .001$; experimental condition 2 vs. control condition $p < .001$). The effect of the covariate (Teacher-Class) was statistically significant in all the comparisons examined with the exception of Pretest for the variable Structure ($p = .490$).

Finally, Fig. 2 shows the change in performance over time in Structure, Coherence and Quality by condition.

## Condition effects by testing time point

GLM (repeated measures ANCOVA) for the variable Structure (Table 5), demonstrated that the main effect was significant for the comparisons between the time points Pretest-Posttest, Pretest-Test 1,Test 1-Test 2 and Posttest-Follow-Up (with small effect sizes). Considering the post hoc analysis, the differences between the two experimental conditions were nonsignificant. However, both of the experimental conditions exhibited significant differences compared to the control condition (with the exception of the time points Pretest-Test 1).

For the variable Coherence (Table 5), the main effect was significant for the comparison between three time points (Pretest-Posttest, Test 1-Test 2 and Posttest-Follow-Up) with small
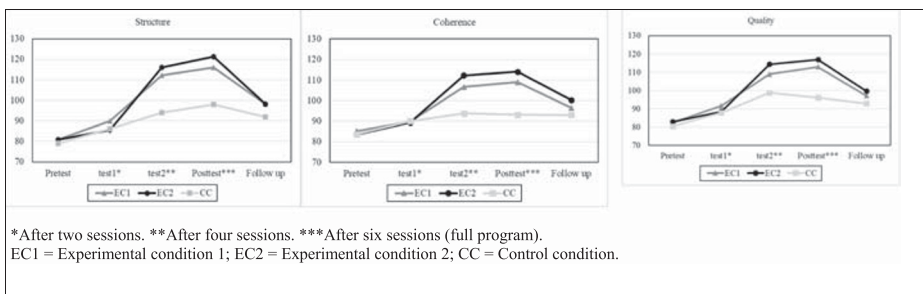


*After two sessions. **After four sessions. ***After six sessions (full program).
EC1 = Experimental condition 1; EC2 = Experimental condition 2; CC = Control condition.

**Fig. 2** Change in performance in each variable by condition

effect sizes. In the three cases, post hoc analysis revealed differences between each of the experimental conditions and the control condition.

Finally, for the variable Quality (Table 5), the main effect was significant for the five comparisons (Pretest-Posttest, Pretest-Test 1, Test 1-Test 2, Test 2-Posttest, Posttest-Follow-Up) with small effect sizes. The post hoc analysis indicated that both experimental conditions produced significantly better results in Quality than the control condition (with the exception of the comparison between Pretest-Test 1).

The effect of the covariate (Teacher-Class) was statistically significant in all the comparisons with $p$ values lower than .01 (Table 5).

**Table 5** Results of GLMs (repeated measures ANCOVA) conducted for the three variables

|  | Interaction (evaluation time x condition) | | | Covariate (Teacher-Class) |
|---|---|---|---|---|
|  | $F(2, 122)$ | $\eta_p^2$ | Post hoc | $F(1, 122)$ |
| **Structure** |  |  |  |  |
| Pretest-Posttest | 19.84*** | 0.24 | EC1 > CC: $p < .001$; EC2 > CC: $p < .001$ | 10.44** |
| Pretest-Test 1 | 6.15** | 0.09 | EC1 > CC: $p = .20$; EC2 > CC: $p = .96$ | 12.53** |
| Test 1-Test 2 | 20.22*** | 0.24 | EC1 > CC: $p < .001$; EC2 > CC: $p < .001$ | 23.66*** |
| Test 2-Posttest | 0.43 | 0.00 | EC1 > CC: $p < .001$; EC2 > CC: $p < .001$ | 17.80*** |
| Posttest-Follow-Up | 14.25*** | 0.18 | EC1 > CC: $p < .001$; EC2 > CC: $p < .001$ | 36.75*** |
| **Coherence** |  |  |  |  |
| Pretest-Posttest | 19.75*** | 0.24 | EC1 > CC: $p < .001$; EC2 > CC: $p < .001$ | 10.24** |
| Pretest-Test 1 | 1.01 | 0.01 | EC1 > CC: $p = .97$; EC2 > CC: $p = .98$ | 23.67*** |
| Test 1-Test 2 | 23.80*** | 0.28 | EC1 > CC: $p = .027$; EC2 > CC: $p = .002$ | 28.46*** |
| Test 2-Posttest | 0.99 | 0.01 | EC1 > CC: $p < .001$; EC2 > CC: $p < .001$ | 14.00*** |
| Posttest-Follow-Up | 12.56*** | 0.17 | EC1 > CC: $p < .001$; EC2 > CC: $p < .001$ | 23.09*** |
| **Quality** |  |  |  |  |
| Pretest-Posttest | 18.51*** | 0.23 | EC1 > CC: $p < .001$; EC2 > CC: $p < .001$ | 20.97*** |
| Pretest-Test 1 | 3.42* | 0.05 | EC1 > CC: $p = .30$; EC2 > CC: $p = .98$ | 37.20*** |
| Test 1-Test 2 | 12.58*** | 0.17 | EC1 > CC: $p = .01$; EC2 > CC: $p = .008$ | 25.29*** |
| Test 2-Posttest | 4.36* | 0.06 | EC1 > CC: $p < .001$; EC2 > CC: $p < .001$ | 14.39*** |
| Posttest-Follow-Up | 13.43*** | 0.18 | EC1 > CC: $p < .001$; EC2 > CC: $p < .001$ | 22.28*** |

$\eta_p^2$ (eta-squared statistic) = estimates of effect size. From post hoc comparison of all students in each group the Bonferroni correction was applied ($p < .05/3 = .016$). EC1 = Experimental condition 1; EC2 = Experimental condition 2; CC = Control condition

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

## Discussion

The aim of this study was to explore the short- and long-term effects of different sequences of strategy-focused instruction (CSRI) on the writing skills of 4th grade students. Our findings provide empirical support for strategy-focused instruction, as we initially hypothesized, and in accordance with previous findings (e.g., Brunstein and Glaser 2011; Palermo and Thomson 2018; Rosário et al. 2019; Saddler et al. 2019).

Strategy-focused instruction provides a complex package of content and instructional methods that gives students strategic knowledge that they can use to regulate what and how they write (MacArthur 2017). In this study we compared the benefits of two experimental conditions based on strategy-focused instruction which differed in the order of implementation of the components, to assess which sequence produced better improvement in writing skills (Structure, Coherence and Quality). The results suggested that after six sessions of training, both sequences of the CSRI program were effective for the improvement of writing skills in 4th grade students. Both experimental conditions exhibited improvements in the quality of their writing compared to the control, which received traditional instruction based on text analysis. As we predicted, and in line with previous studies analyzing strategy-focused instruction (e.g., Bouwer et al. 2018; Cer 2019; Koster et al. 2017; Shen and Troia 2018), students in the experimental conditions produced texts that were assessed as being more coherent, better structured, and of higher quality.

Additionally, we focused on the individual effects of each component (Direct Teaching, Modelling, or Peer-Practice) to determine which of these components were necessary for substantial positive learning in writing (De la Paz 2007; Fidalgo et al., 2017a) when taught to full-range classes. Specifically, we assessed students during the treatment (after each instructional component) to explore changes in writing variables, which allowed us to analyze which components were more effective. Results suggested significant benefits of CSRI after four intervention sessions. Before a firm conclusion can be drawn about the benefits of the CSRI the results need to be explained in more detail. Contrary to Fidalgo et al. (2015) and López et al. (2017) who found immediate benefits of Modelling or Direct Teaching components, after two sessions in our study, we did not find significant effects between CSRI program students and the control condition in the three variables (Structure, Coherene and Quality). There may be several reasons for this lack of differences between the three conditions. One possible explanation is the combination of the nature of the sample, given that they were younger than those in previous studies of CSRI and the novel instructional content (typical writing instruction in Spanish primary schools is the analysis of features of different text types, sentence construction and spelling without any strategic-focused instruction). In experimental condition 1, students had to remember the metacognitive knowledge of the explicit strategy, remembering a mnemonic that represented the planning or drafting steps during the writing process. This could be an excessive cognitive load for developing students who have not yet automated sentence production (Rijlaarsdam et al. 2011). On the contrary, experimental condition 2, with students learning through observation, avoided the need for students to maintain explicit strategic representations and could reduce the dual challenge of learning-to-write and task execution (Kellogg 2018; Rijlaarsdam and Couzijn 2000). However, the content delivery method was novel (Modelling) and could have diverted their attention to the ideas the model gave about the topic *("the similarity between cars and airplanes is that both are transport")* not on the process *("the first thing I have to do is plan my text about similarities and differences between cars and airplanes").* The ability to differentiate a thought- process from

an idea per se, is a complex task for young students if they have no previous experience and their metacognitive skills are in development. As Pennequin et al. (2010) stated, metacognitive knowledge begins to develop at age of 6, however, this skill does not seem to reach maturity until early adolescence, at 11–12 years old. Another possible explanation is that in both experimental conditions instructional content was taught in a single, short session each week. It has been well demonstrated that implementing strategy instruction can be a challenge for teachers, especially Modelling, which many teachers are not familiar with (Harris and Graham 2017).

After two more sessions, adding the Peer-Practice component, we found that this component provided the greatest learning gains. Both experimental conditions showed significant gains in all variables compared to the control. This is consistent with previous research that emphasized the need for students to practice their writing in a supportive environment that includes peer-assisted writing (e.g., De Smedt et al. 2020; Graham et al. 2012; Koster et al. 2017; Yarrow and Topping 2001). Again, we found no differences between the two experimental conditions. During the Peer-Practice sessions, students rehearsed a strategy mnemonic (experimental condition 1) or specific process (experimental condition 2) with pair work clearly established (writer-helper). However, the way in which the students participated and gave feedback to each other could have influenced the quality of their learning (Graham et al. 2015; Wigglesworth and Storch 2012). During the collaborative writing in experimental condition 1, students had to remember the mnemonic rules about the planning and drafting steps. However, students had not had the opportunity to watch a model showing them how to implement the steps. In contrast, in the collaborative writing in experimental condition 2, the students had to remember the Modelling the teacher had done in the previous sessions ["*the model first thought of making an outline with three parts: Introduction, development and conclusion*"]. This probably explains why this condition had the highest average score. In both cases, the complexity of the strategies and processes practiced in a short period of time may be the reason that the two experimental conditions had a similar performance.

Finally, by adding the last two training sessions, both experimental conditions produced slight increases in writing performance with no differences between them. The small effect size demonstrated that the last two sessions reflected a maintained gain in performance but did not provide additional benefits over and above the Peer-Practice component. In experimental condition 1, students watched the model, which helped them to understand that using the strategies worked and was beneficial in the context of a real writing task. On the contrary, but equally effectively, in experimental condition 2, the teacher explicitly gave declarative knowledge about planning and drafting, and the mnemonics associated with helping students remember the planning and drafting steps.

One of the gaps in the literature that we aimed to fill was the absence of studies examining the long-term effects of strategy-focused instruction (e.g., Shen and Troia 2018; Torrance et al., 2015; Tracy et al. 2009). Our results showed that eight months after finishing the intervention both experimental conditions exhibited a significant decrease in all variables, reflected in the means for Structure, Coherence and Quality. At the Follow-Up evaluation, although we did not find significant differences between the three conditions, the two experimental conditions produced texts with more structure, coherence and quality than the control group, above their Pretest scores. These results are somewhat similar to Hacker et al. (2015) and Brunstein and Glaser (2011) who carried out an intervention based on SRSD with elementary grade students. In both cases their results showed that SRSD students wrote better compositions compared to the control at a Follow-Up test. However, Follow-Up probes were

administered in the first study two months after the intervention and, in the second study after only six weeks.

Finally, following the suggestions of Bouwer et al. (2015) about the validity of inferences made based on writing performance, in our study we used multiple tests and different panels of raters in order to obtain valid genre-specific writing conclusions. In addition, the significant effects of the covariate (teacher-cass) reflects the importance of the role of the teacher in classroom interventions and the need to control this variable more than just as a covariate (Murnane and Willett 2011). The aim is to ensure that the changes in students' performance are due to the CSRI program and not to the particular teacher's practices. Future studies might consider running multilevel analyses.

In summary, the present study suggests that for 4th grade students, a short strategy-focused instruction is effective in improving writing skills in a short-term period. In particular, the Peer-Practice component may be a useful practice for promoting improvements in writing performance.

## Limitations and future directions

The implications of our results should be considered in light of the following limitations. This research was performed in a specific school context and population. Therefore, to confirm these findings, the present study needs replication using a larger sample and more homogeneous groups in a different school context with different students (for example, other socio-economic levels or students with learning difficulties). Another limitation was the lack of a specific assessment of writing strategies, metacognitive strategies, and self-regulated learning. This limitation could be overcome through analyzing student text per se (re-writing, errors or symbols and abbreviations that represent the strategy used by each student). A further limitation was related to the lack of a specific assessment of social validity (i.e., the acceptability of and satisfaction with the intervention procedures; Koster et al. 2017). This limitation could be overcome through interviews or questionnaires (e.g., Kiuhara et al. 2012) at different times during the intervention.

In addition, we also want to draw attention to the limitations related to the use of the anchor text procedure. Recently, the value of the anchor text procedure has been noted, given that it allows raters to evaluate text more reliably (they can compare student compositions with fixed example texts as benchmarks that represent the range in text variables) (De Smedt et al. 2020). However, the text scores represent only an approximation to the fixed text (Tillema et al. 2013). Nevertheless, since we chose a large number of anchor texts, the error due to interpolation was reduced to a minimum.

Finally, we suggest that it may be fruitful to study the linking of thinking aloud records with time spent on the text or changes produced on the text. Also, with the use of strategies as revealed in the traces of the written text (López et al. 2019). Moreover, a key aspect of Peer-Practice is to explore the feedback between students during writing. The form of feedback between students could have had an impact on learning (providing direct positive comments that suggest how to correct an error is not the same as providing indirect feedback such as general coments). The literature suggests direct and constructive feedback is a crucial factor that significantly improves writing performance (Duijnhouwer et al. 2012; Graham et al. 2015). The recording and analysis of Peer-Practice during writing may provide interesting information about students' interactions. At the same time, it may be fruitful to study whether

the process transfers to another text type that is not the focus of instruction, or compare the effects of this instruction to other types of writing programs that place greater emphasis on giving students only strategies for setting appropriate production goals (Torrance et al. 2015).

## Educational implications

In terms of educational implications, this study underlines not only the importance of promoting a supportive writing environment in which students practice, test and apply the knowledge learned, but also the message that writing is not a complementary task that is learned automatically. Writing needs to be taught and shown effectively to students for them to learn. CSRI has been shown to be an option as a tool for promoting writing skills in 4th grade students. The CSRI program was developed considering students in private and public schools (with and without learning difficulties, from different socioeconomic contexts) and considering the normal duration of classes in public and private schools. In this way, the CSRI program can be included in the annual classroom program, which would allow teachers to implement it within the curriculum from the beginning of the school year.

## Compliance with ethical standards

**Conflict of interest** The authors declare that they have no conflict of interest.

**Ethical Approval** All procedures performed in this study were in accordance with the ethical standards of the institutional research committee (Ethical Committee of the Principality of Asturias (reference: CPMP/ICH/135/95, code: TDAH-Oviedo) and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

**Informed consent** Informed consent was obtained from all individual participants and/or families included in the study.

## Appendix 1. Example of anchor texts scores (70 weaker text; 100 average; and 130 greater text) and transcribed student texts. Translated from Spanish

### Example anchor text (score 100) for Structure measure

a) Presents an introduction with a meta-structure at the beginning (e.g. *in this text, I am going to present...)* of the text.
b) Presents ideas collected together according to similarities and differences in a consistent manner
c) Uses standard connection conjunction words, albeit mostly simple ones, when linking ideas (e.g. *class textbooks are for studying, but adventure novels serve for entertainment)*

d) Separates the text by paragraphs (without using links between them)
e) Does not present a conclusion section at the end of the text.

## Example transcription (anchor text reference score 100).

*In this text, I am going to compare and contrast class textbooks and adventure novels. Among their similarities, both are books, both have a topic, both have written letters, and both have a cover.*

*Among their differences, one is for studying and the other for entertainment.*

## Example Anchor text (score 130) for Quality measure

a) Explicitly presents the topic and objective of the text through an introduction, which is differentiated from the body of the text *(e.g. in this text, I will compare and contrast two ways of travelling, these are by plane and by car).*
b) Presents and keeps a clear sequence of ideas throughout the text (organized according to similarities and differences). In addition, presents several ideas for similarities and differences.
c) The ideas are coherent, detailed and use examples *(e.g. another similarity is that both ways of travelling cost some money, because when flying you pay for the flight ticket, while when going by car, you have to pay for the fuel)*
d) Presents correct sentence structure, with most of them being complex
f) Presents good general organization of the text, an introduction, a well-written body and a well worked conclusion *(e.g. in this text we are going to compare and to contrast two different ways of travelling: the main similarity is… the main difference is… /To summarize…)*
g) Correct use of different, mainly complex, links between the ideas *(e.g. the main similarity, another important similarity, either travelling by car or by plane, the main difference, …)*
h) Use of complex links or discursive markers both at the beginning and at the end of the text *(e.g. In this text, we are going to compare and to contrast two different forms of travelling… In conclusion….)*
i) Accurate vocabulary, whether simple or standard *(e.g. vehicles, fuel, consume)*
j) Does not make spelling mistakes. Correct use of accents throughout the whole text

## Example transcription (anchor text reference score 130).

*In this text we are going to compare and contrast two ways of traveling, which are traveling by plane and traveling by car.*

*The main similarity is that both are vehicles and need someone to drive them. Another similarity is that both cost money.*

*The main difference is that traveling by plane takes less time but traveling by car it takes longer. Another difference is that in the plane you travel by air and in the car you travel by land.*

*In conclusion, both traveling by car and traveling by plane are two good ways to travel and both cost money.*

**Example Anchor text (score 70) for Coherence measure**

a. The topic of the text is not clearly identified
b. In the body of the text, the sequence of ideas follows a distinguishable and coherent plan (e.g. firstly discusses similarities, and later differences)
c. In the body of the text, the ideas (e.g. similarities and differences) are presented in a relatively simple manner, i.e., simply resorting to enumeration without using proper connectors (and/also…)
d. Reference markers (e.g. one, the other,) are incorrectly used, since they do not specify which objects are referred to
e. There are no digressions and/or loops back in the content. There is no repetition or mixing up of ideas
f. There is no kind of conclusion/sentence that brings sense to the text

**Example transcription (anchor text reference score 70).**

*Similarities: the two are entertaining (referring to textbooks and adventure novels), the two can be found in a library, the two are of different types. Differences: one is to study and another to read at night.*

## Appendix 2. Example of script that tormed the basis for the modelling component (adapted from Fidalgo and Torrance 2017) translated from Spanish

### General Instructions.

During thinking aloud, you should remember that you have to be explicit about what thoughts you should follow to correctly plan a text. In this case you should explicitly reference the POD + vowels strategy. During the modeling, you should use self-questions and explicit self-instructions responding to those self-questions in relation to:

– Regulation of what you are doing. For example: self-questions: *What is the first thing that I have to do?* Self- instructions: *the first thing that I have to do is Plan my text.*
– Regulation during the writing process of what you have done so far. For example: self-questions: *Have I thought of enough ideas?* Self-instructions: *Yes, I've done that right, I have followed my plan well.*
– Regulation at the end of the process about what you did. For example: self-questions: *Have I organized my ideas well?* Self-instructions: *Yes, I have enough differences and similarities between the ideas, so I can continue.*
– Regulation of interests/motivations/self-efficacy beliefs, such as: *I am doing it very well; I am going to write a great text…*

Once these general instructions are clarified, an example of how thinking aloud for the comparison-contrast text of the topic "letter versus e-mail" is shown below.

*I have to write a compare-contrast essay that explains the similarities and the differences between letters and e-mails. I must focus on this task and on the text that I have to write. I am confident that I can write this well. The first thing I have to do is Plan my text. My teacher has told me that it is the most important thing and also, they have given me a tip for planning, POD + the vowels. The first letter of POD is P, (Piensa- think), I have to stop and think about what I'm going to write. Ohh!! I have so many thoughts ... I have to Organize them, as the O in POD says ... but ... How do I do it? Oh yeah! With the vowels!!! They will help me! Mmmm it was the O ... the A .... Well, the first thing I'm going to do is, as letter E (Esquema-Outline) says, an outline of my text in which I will write down my ideas about the introduction of the text, how I develop it, and the conclusion* [pretending to write a plan in the sheet]. *But first of all, it is important when I'm planning to always think of A, which means audience, that is, who will read my text. My audience is the teacher so I'm going to write the best text in my class!! I will continue with the vowels. They are helping me a lot!! The meaning of O ... is the Objective! The purpose of my text is to explain the differences and similarities between letters and emails. I'll write the objective in my introduction* [pretending to write in the sheet]. *Well, now, I am going to continue with the development paragraph (letter D in POD). What ideas can I write in my text? First, I'm going to think about general ideas that I know. So, first, I am going to think...what do e-mails and letters look like? Or what do they have in common? What are the main ideas here? Well, between similarities, the main ideas are (...). But first of all, it is important to remember what U means (Unir-link), the ideas about similarities should go in one paragraph, and the differences in another paragraph.*

[write notes regularly asking: Is that suitable? Have I already got that one? After writing down several ideas …] *I don't have any more ideas. I am going to check to see if I have enough.* [Read aloud the list of similarities. Ask whether ideas are main or secondary. Possibly generate more ideas then …]. *Now what about the other differences between e-mails and letters?* [repeat the same process then …] *I'll read all of these again* [read list]. *Are the ideas correctly organized as letter O in POD says?* [Check the sheet] *Yes! ideas which explained the similarities are in one paragraph and the differences are in another paragraph. I am doing this very well! So, have I finished? Ah, no. As I wrote in my outline*, I must develop the last part of my text, the conclusion. *I need to sum up ideas and write my thoughts about them. Something like ...* [pretending to write a concluding sentence] *Great! I have already finished my planning and I think that I am doing everything correctly, I am sure, the POD + vowels tip has helped me a lot! I'll just read it again* [read through notes, with self-praise and other comments, maintaining student's' interest].

# References

Albertson, L. R., & Billingsley, F. F. (2000). Using strategy instruction and self-regulation to improve gifted students' creative writing. *Journal of Secondary Gifted Education, 12*(2), 90–101. https://doi.org/10.4219/jsge-2000-648.

Arrimada, M., Torrance, M., & Fidalgo, R. (2019). Effects of teaching planning strategies to first-grade writers. *British Journal of Educational Psychology, 89*(4), 670–688. https://doi.org/10.1007/s11145-018-9817-x.

Bai, B. (2018). Understanding primary school students' use of self-regulated writing strategies through think-aloud protocols. *System, 78*, 15–26. https://doi.org/10.1016/j.system.2018.07.003.

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*(2), 415–428. https://doi.org/10.1037/a0022545.

Boon, S. I. (2016). Increasing the uptake of peer feedback in primary school writing: Findings from an action research enquiry. *Education 3–13, 44*(2), 212–225. https://doi.org/10.1080/03004279.2014.901984.

Bouwer, R., Béguin, A., Sanders, T., & van den Bergh, H. (2015). Effect of genre on the generalizability of writing scores. *Language Testing, 32*(1), 83–100. https://doi.org/10.1177/0265532214542994.

Bouwer, R., Koster, M., & van den Bergh, H. (2018). Effects of a strategy-focused instructional program on the writing quality of upper elementary students in the Netherlands. *Journal of Educational Psychology, 110*(1), 58–71. https://doi.org/10.1037/edu0000206.

Braaksma, M. A. H., Rijlaarsdam, G., & van den Bergh, H. (2002). Observational learning and the effects of model-observer similarity. *Journal of Educational Psychology, 94*(2), 405–415. https://doi.org/10.1037/0022-0663.94.2.405.

Braaksma, M. A. H., Rijlaarsdam, G., van den Bergh, H., & Van Hout-Wolters, B. H. A. M. (2010). Observational learning and its effect on the orchestration of writing processes. *Cognition and Instruction, 22*(1), 1–36. https://doi.org/10.1207/s1532690Xci2201_1.

Brunstein, J. C., & Glaser, C. (2011). Testing a path-analytic mediation model of how self-regulated writing strategies improve fourth graders' composition skills: A randomized controlled trial. *Journal of Educational Psychology, 103*(4), 922–938. https://doi.org/10.1037/a0024622.

Cer, E. (2019). The instruction of writing strategies: The effect of the metacognitive strategy on the writing skills of pupils in secondary education. *SAGE Open, 9*(2), 1–17. https://doi.org/10.1177/2158244019842681.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences (2nd ed.)*. Hillsdale: Erlbaum.

de Boer, H., Donker, A. S., Kostons, D. D., & van der Werf, G. P. (2018). Long-term effects of metacognitive strategy instruction on student academic performance: A meta-analysis. *Educational Research Review, 24*, 98–115. https://doi.org/10.1016/j.edurev.2018.03.002.

De La Paz, S. (2007). Managing cognitive demands for writing: Comparing the effects of instructional components in strategy instruction. *Reading and Writing Quarterly, 23*(3), 249–266. https://doi.org/10.1080/10573560701277609.

De Smedt, F., Graham, S., & Van Keer, H. (2020). "It takes two": The added value of structured peer-assisted writing in explicit writing instruction. *Contemporary Educational Psychology, 60*, e101835. https://doi.org/10.1016/j.cedpsych.2019.101835.

De Smedt, F., & Van Keer, H. (2018). Fostering writing in upper primary grades: A study into the distinct and combined impact of explicit instruction and peer assistance. *Reading and Writing, 31*(2), 325–354. https://doi.org/10.1007/s11145-017-9787-4.

De Smedt, F., Van Keer, H., & Merchie, E. (2016). Student, teacher and class-level correlates of Flemish late elementary school children's writing performance. *Reading and Writing, 29*(5), 833–868. https://doi.org/10.1007/s11145-015-9590-z.

Duijnhouwer, H., Prins, F. J., & Stokking, K. M. (2012). Feedback providing improvement strategies and reflection on feedback use: Effects on students' writing motivation, process, and performance. *Learning and Instruction, 22*(3), 171–184. https://doi.org/10.1016/j.learninstruc.2011.10.003.

Festas, I., Oliveira, A. L., Rebelo, J. A., Damião, M. H., Harris, K., & Graham, S. (2015). Professional development in self-regulated strategy development: Effects on the writing performance of eighth grade Portuguese students. *Contemporary Educational Psychology, 40*, 17–27. https://doi.org/10.1016/j.cedpsych.2014.05.004.

Fidalgo, R., Harris, K. R., & Braaksma, M. (2017a). Design principles for teaching effective writing: An introduction. In R. Fidalgo, K. R. Harris, & M. Braaksma (Eds.), *Studies in writing series: Design principles for teaching effective writing* (Vol. 34, pp. 3–10). Leiden: Brill.

Fidalgo, R., & Torrance, M. (2017). Developing writing skills through cognitive self-regulation instruction. In R. Fidalgo, K. R. Harris, & M. Braaksma (Eds.), *Studies in writing series: Design principles for teaching effective writing* (Vol. 34, pp. 89–118). Leiden: Brill.

Fidalgo, R., Torrance, M., & López-Campelo, B. (2017b). CSRI program on planning and drafting strategies. Sessions and supportive instructional materials. In R. Fidalgo, K. R. Harris & M. Braaksma (Eds.), *Studies in writing series: Design principles for teaching effective writing* (Vol. 34, pp. 17–51). Leiden: Brill.

Fidalgo, R., Torrance, M., Rijlaarsdam, G., van den Bergh, H., & Álvarez, M. L. (2015). Strategy-focused writing instruction: Just observing and reflecting on a model benefits 6th grade students. *Contemporary Educational Psychology, 41*, 37–50. https://doi.org/10.1016/j.cedpsych.2014.11.004.

Gillespie, A., & Kiuhara, S. A. (2017). SRSD in writing and professional development for teachers: Practice and promise for elementary and middle school students with learning disabilities. *Learning Disabilities Research & Practice, 32*(3), 180–188. https://doi.org/10.1111/ldrp.12140.

Graham, S. (2018). Introduction to conceptualizing writing. *Educational Psychologist, 53*(4), 217–219. https://doi.org/10.1080/00461520.2018.1514303.

Graham, S., Gillespie, A., & McKeown, D. (2013). Writing: Importance, development, and instruction. *Reading and Writing, 26*(1), 1–15. https://doi.org/10.1007/s11145-012-9395-2.

Graham, S., & Harris, K. R. (2017). Evidence-based writing practices: A meta-analysis of existing meta-analyses. In R. Fidalgo, K. R. Harris, & M. Braaksma (Eds.), *Studies in writing series: Design principles for teaching effective writing* (Vol. 34, pp. 13–37). Leiden: Brill.

Graham, S., Hebert, M., & Harris, K. R. (2015). Formative assessment and writing: A meta-analysis. *The Elementary School Journal, 115*(4), 523–547. https://doi.org/10.1086/681947.

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*(4), 879–896. https://doi.org/10.1037/a0029185.

Graham, S., & Perin, D. (2007). A meta-analysis of writing instruction for adolescent students. *Journal of Educational Psychology, 99*(3), 445–476. https://doi.org/10.1037/0022-0663.99.3.445.

Grenner, E., Åkerlund, V., Asker-Árnason, L., van de Weijer, J., Johansson, V., & Sahlén, B. (2018). Improving narrative writing skills through observational learning: A study of Swedish 5th-grade students. *Educational Review*, 1–20. https://doi.org/10.1080/00131911.2018.1536035.

Grünke, M., Saddler, B., Asaro-Saddler, K., & Moeyaert, M. (2019). The effects of a peer-tutoring intervention on the text productivity and completeness of narratives written by eighth graders with learning disabilities. *International Journal for Research in Learning Disabilities, 4*(1), 41–58. https://doi.org/10.28987/ijrld.4.1.4.

Hacker, D. J., Dole, J. A., Ferguson, M., Adamson, S., Roundy, L., & Scarpulla, L. (2015). The short-term and maintenance effects of self-regulated strategy development in writing for middle school students. *Reading and Writing Quarterly, 31*(4), 351–372. https://doi.org/10.1080/10573569.2013.869775.

Harris, K. R., & Graham, S. (2017). Self-regulated strategy development: Theoretical bases, critical instructional elements, and future research. In R. Fidalgo, K. R. Harris, & M. Braaksma (Eds.), *Studies in writing series: Design principles for teaching effective writing* (Vol. 34, pp. 119–151). Leiden: Brill.

Hartman, H. J. (2001). Developing students' metacognitive knowledge and skills. In H. J. Hartman (Ed.), *Metacognition in learning and instruction. Theory, research and practice* (pp. 33–68). Dordrecht: Kluwer Academic Publishers.

Hayes, J. R. (2012). Modeling and remodeling writing. *Written Communication, 29*(3), 369–388. https://doi.org/10.1177/0741088312451260.

Kellogg, R. T. (2018). Professional writing expertise. In K. A. Ericsson, R. R. Hoffman, A. Kozbelt, & A. M. Williams (Eds.), *The Cambridge handbook of expertise and expert performance* (pp. 389–399). New York: Cambridge University Press.

Kistner, S., Rakoczy, K., Otto, B., Dignath-van Ewijk, C., Buttner, G., & Klieme, E. (2010). Promotion of self-regulated learning in classrooms: Investigating frequency, quality, and consequences for student performance. *Metacognition and Learning, 5*(2), 157–171. https://doi.org/10.1007/s11409-010-9055-3.

Kiuhara, S. A., O'Neill, R. E., Hawken, L. S., & Graham, S. (2012). The effectiveness of teaching 10th-grade students STOP, AIMS, and DARE for planning and drafting persuasive text. *Exceptional Children, 78*(3), 335–355. https://doi.org/10.1177/001440291207800305.

Kline, R. B. (2011). *Principles and practice of structural equation modeling*. New York: Guilford Press.

Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine, 15*(2), 155–163. https://doi.org/10.1016/j.jcm.2016.02.012.

Koster, M., Bouwer, R., & van den Bergh, H. (2017). Professional development of teachers in the implementation of a strategy-focused writing intervention program for elementary students. *Contemporary Educational Psychology, 49*, 1–20. https://doi.org/10.1016/j.cedpsych.2016.10.002.

Koster, M., Tribushinina, E., de Jong, P., & van den Bergh, H. (2015). Teaching children to write: A meta-analysis of writing intervention research. *Journal of Writing Research, 7*(2), 299–324. https://doi.org/10.17239/jowr-2015.07.02.2.

Limpo, T., & Alves, R. A. (2013a). Modeling writing development: Contribution of transcription and self-regulation to Portuguese students' text generation quality. *Journal of Educational Psychology, 105*(2), 401–413. https://doi.org/10.1037/a0031391.

Limpo, T., & Alves, R. A. (2013b). Teaching planning or sentence-combining strategies: Effective SRSD interventions at different levels of written composition. *Contemporary Educational Psychology, 38*(4), 328–341. https://doi.org/10.1016/j.cedpsych.2013.07.004.

Limpo, T., & Alves, R. A. (2018). Tailoring multicomponent writing interventions: Effects of coupling self-regulation and transcription training. *Journal of Learning Disabilities, 51*(4), 381–398. https://doi.org/10.1177/0022219417708170.

Limpo, T., Alves, R. A., & Fidalgo, R. (2014). Children's high-level writing skills: Development of planning and revising and their contribution to writing quality. *British Journal of Educational Psychology, 84*(2), 177–193. https://doi.org/10.1111/bjep.12020.

López, P., Torrance, M., & Fidalgo, R. (2019). The online management of writing processes and their contribution to text quality in upper-primary students. *Psicothema, 31*(3), 311–318. https://doi.org/10.7334/psicothema2018.326.

López, P., Torrance, M., Rijlaarsdam, G., & Fidalgo, R. (2017). Effects of direct instruction and strategy modeling on upper-primary students' writing development. *Frontiers in Psychology, 8*, e1504. https://doi.org/10.3389/fpsyg.2017.01054.

MacArthur, C. A. (2017). Thoughts on what makes strategy instruction work and how it can be enhanced and extended. In R. Fidalgo, K. R. Harris, & M. Braaksma (Eds.), *Studies in writing series: Design principles for teaching effective writing* (Vol. 34, pp. 235–252). Leiden: Brill.

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*(1), 51–68. https://doi.org/10.17239/jowr-2011.03.01.3.

McKeown, D., FitzPatrick, E., Brown, M., Brindle, M., Owens, J., & Hendrick, R. (2019). Urban teachers' implementation of SRSD for persuasive writing following practice-based professional development: Positive effects mediated by compromised fidelity. *Reading and Writing, 32*(6), 1483–1506. https://doi.org/10.1007/s11145-018-9864-3.

Montague, M., Enders, G., & Dietz, S. (2011). Effects of cognitive strategy instruction on math problem-solving of middle school students with learning disabilities. *Learning Disability Quarterly, 34*(4), 262–272. https://doi.org/10.1177/073i9487M421762.

Murnane, R. J., & Willett, J. B. (2011). *Methods matter: Improving causal inference in educational and social science research*. Oxford: University Press.

Olive, T., & Kellogg, R. T. (2002). Concurrent activation of high-and low-level production processes in written composition. *Memory & Cognition, 30*(4), 594–600. https://doi.org/10.3758/BF03194960.

Palermo, C., & Thomson, M. M. (2018). Teacher implementation of self-regulated strategy development with an automated writing evaluation system: Effects on the argumentative writing performance of middle school students. *Contemporary Educational Psychology, 54*, 255–270. https://doi.org/10.1016/j.cedpsych.2018.07.002.

Pennequin, V., Sorel, O., Nanty, I., & Fontaine, R. (2010). Metacognition and low achievement in mathematics: The effect of training in the use of metacognitive skills to solve mathematical word problems. *Thinking & Reasoning, 16*(3), 198–220. https://doi.org/10.1080/13546783.2010.509052.

Rietdijk, S., Janssen, T., van Weijen, D., van den Bergh, H., & Rijlaarsdam, G. (2017). Improving writing in primary schools through a comprehensive writing program. *The Journal of Writing Research, 9*(2), 173–225. https://doi.org/10.17239/jowr-2017.09.02.04.

Rijlaarsdam, G., & Couzijn, M. (2000). Writing and learning to write: A double challenge. In R. J. Simons (Ed.), *New Learning* (pp. 157–190). Dordrecht: Kluwer Academic.

Rijlaarsdam, G., van den Bergh, H., Couzijn, M., Janssen, T., Braaksma, M., Tillema, M., et al. (2011). Writing. In S. Graham, A. Bus, S. Major, & L. Swanson (Eds.), *Handbook of educational psychology: Application of educational psychology to learning and teaching* (Vol. 3, pp. 189–228). Washington: American Psychological Society.

Rosário, P., Högemann, J., Núñez, J. C., Vallejo, G., Cunha, J., Rodríguez, C., & Fuentes, S. (2019). The impact of three types of writing intervention on students' writing quality. *PLoS One, 14*(7), e0218099. https://doi.org/10.1371/journal.pone.0218099.

Saddler, B., Asaro-Saddler, K., Moeyaert, M., & Cuccio-Slichko, J. (2019). Teaching summary writing to students with learning disabilities via strategy instruction. *Reading & Writing Quarterly, 35*(6), 572–586. https://doi.org/10.1080/10573569.2019.1600085.

Shen, M., & Troia, G. A. (2018). Teaching children with language-learning disabilities to plan and revise compare–contrast texts. *Learning Disability Quarterly, 41*(1), 44–61. https://doi.org/10.1177/0731948717701260.

Tillema, M., van den Bergh, H., Rijlaarsdam, G., & Sanders, T. (2013). Quantifying the quality difference between L1 and L2 essays: A rating procedure with bilingual raters and L1 and L2 benchmark essays. *Language Testing, 30*(1), 71–97. https://doi.org/10.1177/0265532212442647.

Torrance, M., Fidalgo, R., & Robledo, P. (2015). Do sixth-grade writers need process strategies? *British Journal of Educational Psychology, 85*(1), 91–112. https://doi.org/10.1111/bjep.12065.

Tracy, B., Reid, R., & Graham, S. (2009). Teaching young students strategies for planning and drafting stories: The impact of self-regulated strategy development. *The Journal of Educational Research, 102*(5), 323–332. https://doi.org/10.3200/JOER.102.5.323-33.

van de Weijer, J., Åkerlund, V., Johansson, V., & Sahlén, B. (2019). Writing intervention in university students with normal hearing and in those with hearing impairment: Can observational learning improve argumentative text writing? *Logopedics Phoniatrics Vocology, 44*(3), 115–123. https://doi.org/10.1080/14015439.2017.1418427.

van den Bergh, H., Rijlaarsdam, G., & van Steendam, E. (2016). Writing process theory: A functional dynamic approach. In C. A. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (pp. 57–71). New York: The Guilford Press.

Wigglesworth, G., & Storch, N. (2012). What role for collaboration in writing and writing feedback. *Journal of Second Language Writing, 21*(4), 364–374. https://doi.org/10.1016/j.jslw.2012.09.005.

Yarrow, F., & Topping, K. J. (2001). Collaborative writing: The effects of metacognitive prompting and structured peer interaction. *British Journal of Educational Psychology, 71*(2), 261–282. https://doi.org/10.1348/000709901158514.

Zimmerman, B. J. (2000). Attaining self-regulation: A social cognitive perspective. In M. Boekaerts, P. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 13–39). San Diego: Academic Press.

Zimmerman, B. J. (2013). From cognitive modeling to self-regulation: A social cognitive career path. *Educational Psychologist, 48*(3), 135–147. https://doi.org/10.1080/00461520.2013.794676.

Zimmerman, B. J., & Kitsantas, A. (2002). Acquiring writing revision and self-regulatory skill through observation and emulation. *Journal of Educational Psychology, 94*(4), 660–668. https://doi.org/10.1037/0022-0663.94.4.660.