Strategy-focused writing instruction: just observing and reflecting on a model benefits 6th grade students

Fidalgo, R.; Torrance, M.; Rijlaarsdam, G.C.W.; van den Bergh, H.; Álvarez, M.L.

Published in:
Contemporary Educational Psychology

DOI:
10.1016/j.cedpsych.2014.11.004

Contemporary Educational Psychology 41 (2015) 37–50

Contents lists available at ScienceDirect

Contemporary Educational Psychology

journal homepage: www.elsevier.com/locate/cedpsych

Strategy-focused writing instruction: Just observing and reflecting on a model benefits 6th grade students

Raquel Fidalgo [a], Mark Torrance [b,*], Gert Rijlaarsdam [c], Huub van den Bergh [d], Mª Lourdes Álvarez [e]

[a] Departamento de Psicología, Sociología y Filosofía, University of León, Facultad de Educación Campus de Vegazana, s/n 24071, León, Spain
[b] Division of Psychology, Nottingham Trent University, Burton Street, Nottingham NG1 4BU, UK
[c] Research Institute for Child Development and Education, University of Amsterdam, Nieuwe Prinsengracht 130, Amsterdam, The Netherlands
[d] Utrecht Institute of Linguistics, University of Utrecht, NL-3508 TC Utrecht, The Netherlands
[e] Department of Online Education, International University of La Rioja, Portales, 71 y Gran Vía, 41, La Rioja, Spain

ARTICLE INFO

Article history:
Available online 15 November 2014

Keywords:
Writing
Strategy-focused instruction
Components analysis
Modelling
Observation

ABSTRACT

Three groups of typically-developing 6th grade students (total N = 62) each completed strategy-focused writing training. Using a combined lagged-group and cross-panel design we assessed the effectiveness of a sequence of four different instructional components: observation and group reflection on a mastery model, direct (declarative) instruction, peer feedback and solo practice. Cumulative effects on written product and writing process were assessed at baseline and after each component. Findings supported the effectiveness of strategy-focused intervention: All three groups showed gains, relative to controls, in the quality of their written products assessed by both holistic and text-analytic measures, and a more structured and goal-focused planning processes. These effects were associated almost exclusively with the modelling and reflection component. Improved performance was sustained through other instructional components but there was no strong evidence that they provided additional benefit. This finding was replicated in all three groups, and across two different text-types.

1. Introduction

Developing the ability to communicate clearly in writing is both an important educational focus in its own right, and necessary for demonstrating competence across the curriculum. From a psychological perspective, writing competence involves implementing, coordinating and monitoring processes for planning content, translating this content into sentences, and for reviewing what has been written (Hayes & Flower, 1980). While doing this, writers need to maintain, and have rapid access to, representations of what they want to communicate, of the structure of the emerging text, and of the characteristics of their audience (Kellogg, 2008). Effective writing requires that the student brings to the task knowledge and skills that are writing-specific. Communicating with an absent audience requires particular linguistic skills for maintaining coherence across the text and for guiding readers' focus and understanding. Writing also requires procedural skills for managing the demands of various writing sub-processes without overloading limited cognitive resources.

Mastery of word-level skills (spelling and handwriting) does not appear to be sufficient to ensure writing competence. Students must also develop text-specific linguistic and rhetorical knowledge, and processes that allow this knowledge to be brought to bear on specific writing tasks. Arguably, to be successful writers, students require appropriate strategies. "Strategies" in this context are understood as procedures that students deliberately and effortfully employ with a view to meeting specific goals (Alexander, Graham, & Harris, 1998). Strategy-focused writing instruction therefore teaches students a combination of explicit knowledge about the characteristics of good writing, and strategies for goal-setting and for organizing the writing process that allow this knowledge to be applied to the emerging text. The aim is that students emerge from instruction with the ability and motivation to regulate their own writing processes in a way that ensures that they set and work towards rhetorical goals, rather than just expressing whatever content comes to mind.

Strategy-focused writing instruction has been a major focus of recent research effort. Meta-analytic reviews suggest that it outperforms other approaches in both struggling and typically-developing students, and at both primary and secondary levels (Graham, McKeown, Kiuhara, & Harris, 2012; Graham & Perin, 2007; Rogers & Graham, 2008). Strategy-focused instruction can take a variety of forms. The most widely researched of these is an intervention called Self-Regulation Strategy Development (SRSD; Harris

* Corresponding author. Division of Psychology, Nottingham Trent University, Burton Street, Nottingham NG1 4BU, UK.
E-mail address: mark.torrance@ntu.ac.uk (M. Torrance).

http://dx.doi.org/10.1016/j.cedpsych.2014.11.004
0361-476X/

& Graham, 1996). This has proved successful in a North American school context (e.g., De La Paz & Graham, 2002) and has been adapted, and again has proved successful, for typically-developing students in schools in Germany (Brunstein & Glaser, 2011; Glaser & Brunstein, 2007), Portugal (Limpo & Alves, 2013), and Spain (Fidalgo, Torrance, & Garcia, 2008; Fidalgo, Torrance, & Robledo, 2011; Torrance, Fidalgo, & Garcia, 2007).

The theoretical basis for strategy-focused writing instruction lies in an understanding of writing as a thinking-and-reasoning (problem-solving) process (Bereiter, Burtis, & Scardamalia, 1988; Hayes, 1996; Hayes & Flower, 1980) and theories of learning that emphasize the importance of self-regulation (Schunk & Zimmerman, 1997; Zimmerman, 2000). Problem-solving accounts of writing see text composition, when performed successfully, as being goal-driven: Writers start by setting initial goals for what they want to communicate, and identify rhetorical constraints associated with intended audience and genre. These form the basis for mental or written plans for the content of the text to be produced. Writing, in the sense of producing linked sentences on the page, is then the act of translating these plans into full prose. Text production is possible without following this strict goal–plan–translate sequence, and this may prove successful for expert writers performing familiar tasks. Arguably, though, for developing writers the probability of success is maximized if composition involves deliberate, explicit, and appropriately sequenced decisions about what to say and how to say it. For this to happen students need to know both how to set goals for a particular composition and to have the knowledge to fulfil these.

Crucially, however, students must also choose to apply this knowledge to their writing processes, independently of teacher prompting. Schunk and Zimmerman (1997) presented a social cognitive model of how students develop sequential skills, such as those associated with specific procedures for planning text. Students initially observe the target skill being modelled by others, then deliberately and strategically emulate the behaviours that they have observed. This intermediate stage requires initial scaffolding, which is gradually decreased until students regulate their behaviour without needing regular external or internal monitoring. Social learning is central to this account. This occurs when students first see skills being modelled, and then process-focused comment and encouragement from their peers and teachers.

Strategy-focused writing instruction therefore aims to teach effective goal setting and planning skills using methods based in this social–cognitive model. It typically involves a combination of some or all of following instructional components: direct (declarative) teaching of writing strategies supported by mnemonics and graphic organizers, students observing mastery modelling of these strategies, practice of these strategies in pair or group writing tasks, and solo practice. As we have noted, taken together these components prove particularly effective in developing the writing skills of typically-developing students taught within full-range classes. However, little is known about the relative value of the different instructional components. Understanding this is important for both theoretical and applied reasons. From a theoretical perspective, understanding the effects of individual components gives insight into the psychological mechanisms by which the positive effects of strategy-focused instruction are achieved. For classroom practice, knowing the relative merits of different components allows teachers to incorporate strategy-focused instruction within existing timetables and curricula.

Graham, Harris and co-workers (Danoff, Harris, & Graham, 1993; Graham & Harris, 1989; Sawyer, Graham, & Harris, 1992) explored the relative effects of various decompositions of SRSD instruction. Instruction was individualized rather than whole-class, and with struggling writers. Graham and Harris (1989) contrasted strategy instruction with and without components explicitly aimed at developing self-regulation (goal setting, self-monitoring), finding similar benefits in both conditions. Sawyer et al. (1992) reproduced these conditions and added a third "direct instruction" condition that stripped away the social learning components—teacher modelling and collaborative practice—that have been specifically associated with developing self-regulation. Again, students in all three conditions showed benefit relative to practice-only controls, with no evidence of difference among conditions. Danoff et al. (1993) made similar comparisons by exposing students to a sequence of components starting with direct instruction, then teacher modelling of strategies followed by strategy memorization, (supported by mnemonics), and then collaborative and individual practice. Multiple single-case studies of 4th and 5th grade writers with writing-task probes after each component, suggested limited gains from declarative instruction—in contrast to Sawyer, Graham, and Harris—but gains from both modelling-plus-memorization and, particularly, after collaborative and solo practice. This is consistent with the findings of Zimmerman and Kitsantas (2002) who, in the context of a writing-related but much more constrained task, also demonstrated the benefits of observing models following direct strategy instruction.

The research that we report here also examined the role of modelling and collaborative practice in strategy-focused instruction. However, our aim was rather different. These previous studies aimed to manipulate the self-regulatory content of instruction by adding or removing these social learning components. As Sawyer et al. (1992) observe, however, self-regulation can be taught in many ways: Any learning of strategy, whether by observation of a model or by direct instruction can, in principle, result in an increased tendency for students to self-instruct and self-monitor. For present purposes we do not want to assume direct association between social learning and learning to self-regulate.

The main aims of the present study were as follows: (1) to determine whether observing and then group reflection on modelling that includes self-instruction and self-monitoring, in the absence of declarative instruction, results in improvement in student performance, and (2) to determine the extent to which direct instruction that explicitly formalizes and labels planning and drafting strategies provides additional benefits to student performance over and above those afforded by observation and group reflection (if any). We see the central difference between the modelling-and-shared-reflection and declarative components as whether or not strategies were made explicit, through labels and mnemonics, or inferred from observation and then discussion of a model which used these strategies but did not explicitly label them. For students to learn and apply effective writing strategies, both observation and direct instruction might be necessary: Modelling might be necessary to illustrate strategies taught through direct instruction and/or direct instruction might be necessary to provide a framework for understanding and retaining what has been observed. Alternatively declarative instruction may be essential, and modelling less important (as found by Sawyer et al., 1992, but contrary to Danoff et al., 1993). A third possibility, and the hypothesis that we test in the present study, is that, in certain populations at least, modelling and group discussion that does not explicitly label or directly teach strategies is alone capable of delivering substantial gains in students' writing performance. Thus, in contrast to the studies discussed above which all took direct teaching of strategies as a starting point, we tested the hypothesis that writing performance may improve just through observation and group discussion of effective writing processes in which specific strategies are not made explicit.

There is some reason to believe that this might be the case. Rijlaarsdam and co-workers have conducted a number of studies exploring the effects of observing peers performing composition tasks (reviewed in Rijlaarsdam et al., 2008). Observing peers has shown

benefits beyond solo practice in a number of studies (Couzijn, 1999; Raedts, Rijlaarsdam, Van Waes, & Daems, 2007). Effects are particularly clear when the task is new or complex (Braaksma, Rijlaarsdam, & Van den Bergh, 2002; Braaksma, Rijlaarsdam, Van den Bergh, & Van Hout-Wolters, 2004). They account for the value of these methods, in part, in terms of the "double challenge" that faces developing writers who have to recall and implement new strategies while struggling with an activity that they already find demanding (Rijlaarsdam et al., 2011; Rijlaarsdam & Couzijn, 2000). The benefits of direct strategy instruction may be partially offset by the fact that students have to maintain an explicit representation of the strategy in mind while writing.

Following Danoff et al. (1993), we evaluated the cumulative effects of different instructional components by implementing these in sequence, with assessment tasks administered after each component. Instructional components in our study followed this sequence: (1) Observation of a mastery model, followed by group reflection on what had been observed. (2) Declarative strategy instruction. (3) Peer modelling and feedback. (4) Solo practice. Across all four components participants therefore received a full strategy-focused intervention similar to SRSD. These components had previously been evaluated as a single package in a similar population, and found to be effective (Torrance et al., 2007). In addition to testing whether explicit, declarative strategy instruction provided benefit over modelling and reflection, this design also explored whether additional benefit was resulted from students observing and commenting on each other's attempts at adopting the strategies that they had been taught. Several studies have explored the benefits of peer feedback following strategy instruction (Graham, Harris, & Mason, 2005; Harris, Graham, & Mason, 2006; Yarrow & Topping, 2001). All three of these studies found greater improvements in writing performance as a result of peer support, relative to practice-only controls.

The study implemented this sequence of components in three, mixed-ability, 6th grade classes, using a combined lagged-group and cross-panel design. If writing strategies are best developed through social learning, then we predicted substantial increase in students' use of the target strategies following observation and subsequent group reflection on a mastery model, and following peer feedback, but relatively modest gains from declarative instruction and solo practice. And if, consistent with a problem-solving understanding written composition, explicit goal setting and planning strategies are important in developing good text, then these changes in strategy would be associated with improved text quality. The lagged-group, cross-panel design provided a particularly robust test of the cumulative benefits of the different instructional components with each hypothesized effect being tested in four different between-group comparisons and for instruction in two different genres of expository text.

2. Method

2.1. Design

Three intact classes of students were randomly designated as Groups A, B and C. Each received strategy-focused writing instruction that aimed to teach both writing procedures (specifically, effective planning strategies) and appropriate discourse knowledge and goal setting for creating coherent text. In all three groups instruction had four components with each component delivered in two consecutive one-hour training sessions. Training content is described in detail below. The first *Modelling-and-Shared-Reflection* component involved students observing an instructor modelling effective writing strategies, and then reflecting, individually and as a whole class, on what they had observed. In the second *Declarative Instruction* component students were given explicit, declarative instruction about these strategies. The third, *Peer Feedback* component involved practice of these strategies, with students thinking-aloud while emulating the strategies that they had been taught, and receiving comments on their strategy use from a partner. The final, *Solo Practice* component involved individual practice. Components were delivered in this order in all three groups.

The design of the evaluation is outlined in Fig. 1. It was divided into two phases. In Phase 1 Groups A and B received training, and Group C served as a control. Group A started training at the beginning of Phase 1. Training in Group B was identical to training in Group A, but was delivered with a time lag of one component. So Group B started the Modelling-and-Shared-Reflection component as Group A started the Declarative component, and so forth. This meant that during Phase 1 effects of a specific component were evaluated at two points in time, in two different groups, with two different test prompts. In Phase 2 Group C received training, following the same pattern as Groups A and B in Phase 1, with Groups A and B acting as post-intervention controls. Training and assessment in Phase 1 focused on writing compare-and-contrast essays. Training and assessment in Phase 2 focused on opinion essays. Baseline writing ability was assessed in all groups immediately prior to the start of Phase 1 and of Phase 2 and then at times corresponding to the end of instruction components. The effect of a component was therefore assessed by the change in performance across the probes immediately prior to and immediately following a component relative to normal-curriculum controls.

The design of our study was such that the effect of each instructional component was assessed by four comparisons. In Phase 1, this was by separate comparisons of group A and group B with group C. In Phase 2 the effect of each component was assessed by comparison of Group C with Group A and with Group B. This second phase of the study served both to indicate whether effects found in Phase 1 were present for a different genre, and whether

Group	Phase 1 (Compare-contrast essay)											Phase 2 (Opinion essay)								
A	P_1	M&SR	P_2	Declarative	P_3	Peer Feedback	P_4	Solo Practice	P_5	Control	P_6	P_7	Control	P_8	Control	P_9	Control	P_{10}	Control	P_{11}
B	P_1	NC	P_2	M&SR	P_3	Declarative	P_4	Peer Feedback	P_5	Solo Practice	P_6	P_7	Control	P_8	Control	P_9	Control	P_{10}	Control	P_{11}
C	P_1	Control	P_2	Control	P_3	Control	P_4	Control	P_5	Control	P_6	P_7	M&SR	P_8	Declarative	P_9	Peer Feedback	P_{10}	Solo Practice	P_{11}

Time →

Note. M&SR = Modelling-and-Shared-Reflection component; NC = normal curriculum; P = writing assessment probes

Fig. 1. Research design showing instructional components and writing assessment probes for each group.

benefits of the whole package of instruction delivered in Phase 1 were persistent and transferred to a different genre. For us to conclude that an instructional component resulted in improvement in participants' writing performance then we would need to see statistically significant positive effects for *all four* of these comparisons. This "internal replication" feature of our design therefore provided a particularly strong test of whether or not a component gave benefit over and above that afforded by those that preceded it.

2.2. Participants

Participants (N = 62) comprised all of the students in three intact, mixed ability, 6th grade classes at a *colegio concertado* (religious foundation) school serving a middle-class, suburban, native-Spanish population. Groups were of similar age and did not differ significantly on a measure of general verbal ability (Thurstone & Thurstone, 2004) that includes tasks about vocabulary (identification of the picture named into a set and identification of a word with the same meaning of other into a set) and verbal reasoning (choosing of the different word in a set of words). Scores on the verbal ability task placed students just above national norms (means of between 55th and 63rd centile). There was no statistically significant difference among groups. Participant details are summarized in Table 1.

Prior to training and in the control conditions, students received writing instruction that followed patterns typical of those adopted in Spanish primary schools (summarized in García, de Caso-Fuertes, Fidalgo-Redondo, Arias-Gundín, & Torrance, 2010). Baseline measures in this and in previous studies suggest that students taught within this system do not spontaneously outline in advance of producing full text (Torrance et al., 2007).

Student absences meant a small amount of missing data, with 1.3% missing observations overall. All participants, including those with missing observations, were included in statistical analyses. The multilevel modelling method that we adopted permitted robust parameter estimates when small numbers of observations are missing.

2.3. Training

The four training components are described below, followed by a description of the instruction received in the control condition. In common with other omnibus strategy-focused writing interventions, training aimed to develop in students both knowledge and understanding of product, in the present case focusing on expository compare–contrast essays (Phase 1) and opinion essays (Phase 2), and specific process strategies. With regards written product, training aimed to teach both appropriate meta-linguistic knowledge—specifically the genre-specific characteristics of the target text-type—and more general product goals—the need for text to accommodate to reader needs, the need for global cohesion, and so forth. Training related to process focused on strategies for outlining, as a distinct pre-planning stage, and for translating this plan into full text.

The four training components focused on the same content (i.e. there was no progression in content across components) but varied in terms of how this content was delivered. So by the end of the two sessions in the Modelling-and-Shared-Reflection component students had observed and reflected upon a writing process that exhibited the full range of product knowledge and goals and process strategies (but without these strategies being labelled and made explicit). This same content was then delivered explicitly in the Declarative component, with learning supported through mnemonics and graphics. Students then applied this content to their own writing, and this application was monitored and shaped, in the Peer Feedback and Solo Practice components.

2.3.1. Modelling and shared reflection

The Modelling-and-Shared-Reflection component involved students observing a "mastery model" (Kitsantas, Zimmerman, & Cleary, 2000) first of effective pre-planning processes (Session 1), and then of translation of the resulting written outline into full text (Session 2). Modelling included self-talk and self-monitoring, but did not make explicit reference to the strategies (and associated mnemonics) taught in the direct-instruction component that followed. This was followed by activities intended to ensure that students engaged with and processed what they had observed.

Session 1 started with the instructor briefly introducing the aim of the training, and a discussion within the class about the importance and relevance of writing and being able to write well. The teacher then modelled planning an essay (compare–contrast in Phase 1, opinion in Phase 2). This involved "thinking aloud" in front of the class. The content of the think-aloud was closely prescribed by the researchers, with the instructor following a script (with scope for minor, non-substantive improvisation). An example is given in the Appendix. In producing her text, the instructor referenced each of the following: Objectives and goals for the text; audience—considering who might read the text and what their needs and expectations might be; generation of ideas to include in the text; consideration of how these ideas might be coordinated into a coherent argument; and consideration of the specific genre conventions associated with the type of text to be produced. The instructor's think-aloud included self-questioning and self-instruction, focused both on the content of the task (e.g., "Now, what do I need to do first?", "I must check that I've structured this well") and on motivation ("I'm sure I can find lots more ideas about this" and so forth). In all cases these were bound to the writing of the particular text that the instructor was working on, rather than being abstracted as general principles, strategies, or rules. In particular, the instructor did not make any direct reference to the strategies taught in the second, declarative component and their associated mnemonics.

After modelling students reflected individually on what they had observed, writing down their observations. The instructor then led a whole class discussion in which students shared their observations with the rest of the class. After this discussion, each student individually tabulated differences between their own writing processes and the process that the instructor had modelled. Finally, students studied a copy of the written plan that was the output of the instructors' planning process.

Session 2 followed a similar pattern to Session 1, but with a focus on translating the outline created in the previous session into full text. Students were first reminded of the content of the previous session, reading through their reflections on the planning process that they had seen modelled, and the resulting written plan. The instructor then modelled drafting the text that had been planned in the previous session. This involved "thinking aloud" (again

Table 1
Participants.

	Phase1	Phase 2	N (N female)	Mean age months (sd)	Mean verbal ability (sd)	Centile verbal ability
Group A	Training	Control	21 (6)	134 (3.7)	35.6 (5.9)	63
Group B	Training (lag)	Control	21 (9)	135 (4.5)	34.0 (6.2)	55
Group C	Control	Training	20 (6)	136 (5.7)	35.1 (6.6)	60

scripted) about decisions concerning text structure, choice of appropriate devices for maintaining coherence across the text, and cohesion across sentences and paragraphs, and decisions about word choice. Again the instructor used self-instruction and self-questioning to regulate her writing (e.g., "Readers need to know where I'm heading, so maybe I should say what the structure is first", "Will they [readers] understand that word? I should explain it to them!").

When the instructor had finished students made notes about the procedure that they had just seen modelled. As in the previous session, the instructor then facilitated a whole-class discussion drawing together the students' observations. Finally, each student individually wrote down reflections about the differences between they do when they write a text, and the processes that they had observed. Finally, the teacher passed all students a copy of the text that she had produced.

2.3.2. Direct (declarative) instruction

In this component the product knowledge and process strategies that were displayed during modelling were made explicit through direct instruction, supported by the use of mnemonics and graphic organizers.

The first session of this component (Session 3) focused on planning. First the instructor reminded students of the main features of the planning processes modelled in Session 1, and specifically the importance of setting goals, generating ideas, organizing ideas into a coherent structure, and considering audience. This instruction was supported by a planning–metacognition matrix, given to the students, which gave a bullet-point summary of what planning is, what function it serves, and what it might involve.

The instructor then introduced two mnemonics, which the students memorized: A general exhortation to think before writing, represented by POD—*Piensa* (stop and think before writing; *Organiza tu pensamiento* (organize your thoughts); *Desarrolla tu texto* (develop your text), and specific self-questions to regulate students' thinking when Organizing and Developing, represented by The Vowels—O*bjetivo*/objective (what is the purpose of the text?); A*udiencia*/audience (for whom is it intended?); *Ideas*/ideas (what ideas might be included?); *Unir ideas*/unite ideas (how might my ideas be unified and organized into a coherent whole?); *Esquema*/structure—fit the text into a genre-appropriate structure.

The second session in this component (Session 4) focused on the process of translating a plan into full text. Students learned a generic structure for expository text represented by the mnemonic IDC—*Introducción*/introduction,*Desarrollo*/development, C*onclusión*/conclusion as a means for structuring the process by which the text is produced. The instructor reinforced the need to maintain focus on the criteria identified by the Vowels during all three of the IDC production phases. As in the previous session, students were given a metacognitive matrix identifying the nature, purpose and central features of effective translation processes.

The instructor then presented specific ways in which ideas might be structured and linked within the text, focusing on the value of using paragraphs to structure text, and on linguistic strategies for making links within and between paragraphs. Students then analyzed and contrasted two texts: The text that was produced by the teacher when modelling and a less competent example.

2.3.3. Peer feedback

In this and the following component students aimed to emulate the writing processes that they had seen modelled, and then were explicitly taught. During emulation students thought aloud. This made students aware of their own writing processes, which they were then able to monitor.

In the peer feedback component, think-aloud served the additional function of laying the writing processes adopted by the students open to monitoring and comment by peers. Students worked in writer–observer pairs, planning a text and then, after swapping roles, translating these plans into full text. The first session (Session 5) started with the instructor reminding students of the content of previous sessions and giving students some practice in thinking aloud while writing. Students were then paired. One student constructed a plan for an expository essay, aiming to emulate the writing processes that they had seen modelled and had been explicitly taught, structuring writing processes and decisions around the POD, Vowels, and IDC strategies. This student thought aloud throughout. The students' writing processes were scaffolded by graphic organizers. For example, when planning their text, students had in front of them a table that laid out the Vowels criteria, with spaces for the student's own notes. The second student observed and supported this process, reminding their partner when they skipped consideration of one of the Vowels criteria, when they failed to make clear links between paragraphs, and so forth. The instructor also patrolled the class, listening to the think-aloud and commenting when necessary, both to provide direct input for the writer (including prompts to think aloud if they forgot to do so) and to provide a model for the observer.

In the second session of this component (Session 6) students swapped roles. The new writer then took the outline created in the previous sessions and translated it into text. This session then followed an identical pattern to the previous one, with a focus on the IDC mnemonic, supported again by a graphic organizer.

2.3.4. Solo practice

The two sessions of this component (Sessions 7 and 8) followed exactly the same pattern as the previous component, with the exception that the scaffolding provided by having peer feedback was removed. Students therefore worked alone on a different essay, planning it in the first session and then translating this plan into full text in the second. Students again thought aloud throughout, to support self-monitoring of process. The teacher patrolled the classroom reminding students to continue thinking aloud if they stopped.

2.4. Control

In the control conditions (Group C in Phase 1, Groups A and B in Phase 2) students received an intervention that gave the same level of writing practice as the training condition, focusing on the same text-types. Instruction in the control condition, however, lacked the strategy focus of the training. Specifically instruction did not aim to provide specific self-regulatory metaknowledge of process and product.

As in the first training session, the first Control session started with discussion about the importance and value of writing and learning to write well. The instructor then introduced the specific text-type that would be the focus of the instruction, and discussed its relevance to the students' curriculum. The instructor then gave direct teaching about the particular features of this kind of text—typical structure, use of paragraphs, and so forth. In the second session students then analyzed in detail the linguistic and content features of a good example, using the text created during modelling in the training conditions as the example.

Sessions 3 and 4 continued a focus of text-type, with students reading a range of different texts, with different genres, and learning to identify the particular features that made the text-type that was the focus of the instruction (compare–contrast essay in Phase 1, opinion essay in Phase 2) distinct. Both session involved identify-the-text-type exercises, and exercises and discussion of the how features taught in the first two sessions mapped onto examples of the target text-type.

During the remaining sessions students performed the same writing tasks as in the training condition, but without scaffolding

of process or explicit reference to strategy. In Sessions 5 and 6 students produced text in pairs, without intervention from the instructor and without strategy-related peer feedback. They then read out their finished texts to the class, and the instructor provided spoken feedback, commenting specifically on the extent to which the text conformed to the intended text-type. This was repeated in the remaining sessions but with students writing alone.

2.5. Training delivery and treatment fidelity

All sessions in all conditions were delivered by the same instructor, who was also the students' regular literacy teacher. She had previous experience of delivering the instructional content and methods evaluated. The teacher was also part of the research team, with extensive previous experience of controlled, whole-class evaluation studies. She had a clear understanding of the design and purpose of the present research and particularly of the importance of avoiding bleed both between instructional components within groups, and across groups. Session content was closely prescribed, with the instructor following detailed scripts. The instructor met with another member of the research team between each instructional session to discuss whether in the previous session it had been possible to run the previous session as scripted (in all cases the instructor reported that this was the case) and to discuss the upcoming session.

All sessions prescribed one or more written tasks for the students to complete (e.g., table comparing model with own writing processes in the Modelling-and-Shared-Reflection component, graphic planning organizers in the Declarative component). These were collected after each session and analyzed for (a) successful completion, indicating that students had engaged with and understood instruction within the session, and (b) evidence that instructional content had gone beyond what was prescribed for the session. There was, as might be expected, some variation among students in the extent to which their written output showed engagement with session content. However, we found no evidence, for any of the four components in any of the three groups in which they were delivered, that tasks had not been completed correctly by anything but a small minority of students. We also did not find evidence of bleed from later sessions. We looked, in particular, for evidence that the explicit strategy instruction, and associated mnemonics, intended for delivery in the second, Declarative component, had crept into the instructor's think aloud in the first, Modelling-and-Shared-Reflection component. There was no evidence that this had occurred.

2.6. Evaluation

The writing tests were administered in each group in a lesson immediately following each instructional component. These were administered by one of the researchers.

2.6.1. Probe writing tasks

The assessment writing tasks were based on topics that drew on content from elsewhere in the students' curriculum. During Phase 1 students wrote essays making comparisons between: holidays in the beach and in the mountain; living in a country and in a city; traditional and electronic games; landline phones and cell phones; theatre and cinema; and Christmas and summer holidays. In Phase 2, students wrote opinion essays about: fast food, recycling; doing exercise; tobacco; vegetarianism; keeping wild animals in captivity; and environmental pollution.

2.6.2. Product assessment

We assessed the quality of the texts produced by participants with both holistic (reader-based) ratings, and with a text-analytic approach based on counts of specific features associated with maintaining textual cohesions.

Holistic measures were adapted from methods described by Spencer and Fitzgerald (1993). All texts were independently scored by two trained raters on three non-orthogonal dimensions. *Structure* was assessed on a four point scale based on whether the text included introductory sentences, cues signposting text structure, a topic or thesis sentence, organization of ideas based around a clear and definite scheme, thematic unity within paragraphs, thematic unity across the whole essay, and a conclusion that reiterated the purpose of the paper. *Coherence* was also assessed on a four point scale and was based on whether it was possible to identify a topic or thesis that then provided a consistent focus for the essay, whether the text included sufficient context to orientate the reader to this theme, whether there was a clear thematic and linguistic flow between sentences and paragraphs. *Global Quality* was assessed on a six point scale and gave an evaluation of the extent to which the text had a clear sequence of ideas with little or no irrelevant detail, clear organization, fresh and vigorous word choice, varied and interesting detail, correct sentence structure, and accurate punctuation, capitalization and spelling.

This strategy for generating reader-based assessments of text quality has been used in several previous studies (Fidalgo et al., 2008, 2011; Torrance et al., 2007) with good inter-rater reliability. In the present study, all texts were independently scored on all three dimensions by two trained raters. Inter-rater correlations, averaging across the 11 writing tasks, indicated reasonable reliability (Structure, .80; Cohesion, .73; Global Quality, .80).

Texts were also submitted to text-analytic analysis aimed at identifying linguistic features that mark the use of specific coherence-maintaining devices, drawing on Halliday and Hassan (1976) and more recent developments of their work (Bosque & Demonte, 1999; Sanders, Spooren, & Noordman, 1992). We coded for two general types of device: *Basic Cohesion Ties*—genre-independent features that might be expected to be present in children's texts from early stages of writing, and *Advance Cohesion Ties*—specific linguistic devices associated with expository text that, we hypothesized, will not appear in students' texts unless students deliberately and strategically adopt their use. That is, these devices will not be present unless students possess appropriate meta-linguistic knowledge, and have the necessary self-regulation to apply this during the production of their text. We counted the following Basic Ties: lexical repetition (e.g., *Peter is a young man. Peter likes playing football.*); use of coordinating conjunctions (e.g., *and, or, but*); and anaphoric reference using pronouns (e.g., *Peter is a young man. He likes playing football.*) We counted the following Advanced Ties: Use of structural ties, marked, for example, by structures such as *first…, second…, finally*; reformulation, flagged by markers such as *in conclusion…, that is to say…, in other words*; argumentation, marked by, for example, *for example, however, despite this*; and use of meta-structural markers such as *Now I will describe…, The following paragraph talks about…* and so forth.

Texts were coded independently by two trained raters. Correlations between raters, averaging across writing tasks, were .83 for Advanced Ties and .88 for Basic Ties. Measures are reported as number of occurrences per 100 words to give an index of tie density, independent of text length.

2.6.3. Process assessment

Writing process was assessed in two ways. Notes and outlines made on the "planning" sheet, given to the students at the start of the writing task, were coded for features associated with strategic planning activity, and times spent in the planning, writing, and revising phases of task were recorded.

Notes produced during planning were coded for structure, for evidence of goal setting, and for explicit reference to strategy. Scores

Table 2
F, *p*, and partial η^2 from tests of the hypothesis that the intervention as a whole affected text quality.

Between groups factor	Within groups factor	Holistic quality	Coherence	Structure	Advanced cohesion-tie density
Phase 1: Group A vs. Group C (control)	Probe 1 (baseline) vs. Probe 5 (post-test)	26.7, <.001, .44	26.2, <.001, .43	26.9, <.001, .44	26.3, <.001, .43
	Probe 1 (baseline) vs. Probe 7 (transfer post-test)	31.7, <.001, .47	16.7, <.001, .32	18.0, <.001, .34	6.4, .016, .15
Phase 1: Group B vs. Group C (control)	Probe 2 (baseline) vs. Probe 6 (post-test)	48.2, <.001, .56	60.3, <.001, .62	44.9, <.001, .54	35.9, <.001, .49
	Probe 2 (baseline) vs. Probe 7 (transfer post-test)	22.4, <.001, .36	47.0, <.001, .54	20.9, <.001, .34	19.3, <.001, .33
Phase 2: Group C vs. Groups A & B (control)	Probe 7 (baseline) vs. Probe 11 (post-test)	39.8, <.001, .41	36.7, <.001, .39	16.2, <.001, .22	22.2, <.001, .28

Note: Reported effects are for the interaction between Group and Test for two-way mixed-design ANOVAs. Phase 1 $df = 1$ and 39. Phase 2 $df = 1$ and 60.

in all cases ranged from 0 (when there was an absence of structure, goal setting and strategy in planning) to 2. *Structure* scoring involved looking for any direct and clear reference to introduction, body, and conclusion components in the text. The introduction–body–conclusion structure was a specific feature of the training condition. Plans scored 1 where there was explicit evidence of content being planned under one of these headings, in the form of lists of keywords, and scored 2 when the content was more elaborated text. *Evidence of goal setting* was scored as 1 if there was any reference to the need to set a goal for the text within the student's notes, and 2 if this was elaborated into a specific goal (i.e. statements of the form *The purpose of this text is to argue…*). *Explicit reference to strategy* measured whether or not the student's notes showed evidence of explicitly recalling strategic knowledge about how to plan. If notes reproduced, or partially reproduced, any of the planning mnemonics that featured during training, or showed explicit reference to any other well-developed planning schemas, then this was given a score of 2. A score of 1 was given for any evidence of a deliberate planning strategy, over and above just listing ideas, but without explicit reference to mnemonics or other well-developed schemas. This included the use of boxes and arrows, drawing clouds around thesis statements, and so forth.

We estimated times spent in the initial planning phase (if any), time spent writing full text, and time spent revising (if any) by asking students to report the time at which they switched from writing notes to writing full text, when they finished writing full text, and when they felt they had completed the task. The paper on which they wrote full text had a box marked "started writing" at the top, and "finished writing" and "finished task" at the bottom. A clock was visible to all students, and they were instructed to take times from this.

3. Results

3.1. Whole-intervention effects

The main purpose of this study was to explore incremental effects of each of the four components. However, it is worth first asking whether the four components, taken together, resulted in an overall improvement in the quality of students' text. We therefore conducted, for each the three holistic quality measures and for advanced coherence tie density, probe (pre, post) by group (control, intervention) mixed ANOVAs as follows. For Group A vs. Group C (control): Probe 1 vs. Probe 5 and Probe 1 vs. Probe 7. For Group B vs. Group C (control): Probe 2 vs. Probe 6 and Probe 2 vs. Probe 7. For Group C vs. Groups A and B (controls): Probe 7 vs. Probe 11. Analyses with Probe 7 as the post-test tested for transfer to a different genre. For effects in Phase 2, groups A and B were combined to create a single control condition. We found significant interactions between test and group for each of these analyses. In all cases this indicated improved performance in the intervention group relative to control and in most cases this effect was large. Descriptive statistics can be found in Tables 4 and 5, and test statistics can be found in Table 2.

3.2. Effects of individual components

3.2.1. Statistical analyses

To claim evidence that a particular instructional component has been effective we would need to observe a greater improvement in performance across the probes prior to and following that component relative to performance by control students tested on the same probes (i.e. a 2 (Probe: pre vs. post) by 2 (Group: intervention vs. control) interaction). The design of our research (Fig. 1) was such that each component was repeated on three different occasions, each in a different group. If a component was effective we therefore expected to see this interaction at *all* of the following points: Phase 1 for Group A relative to Group C, and in Group B relative to Group C. Phase 2 for Group C relative to Group A, and for Group C relative to Group B. So, for example, to make a strong claim that the Modelling and Shared Revision component had a positive effect on students' performance we would need to find that all four of the following Probe by Group interactions were statistically reliable: In Phase 1, Probe (Probe 1 vs. Probe 2) by Group (A vs. C), and Probe (Probe 2 vs. Probe 3) by Group (B vs. C). In Phase 2, Probe (Probe 7 vs. Probe 8) by Group (C vs. A), and Probe (Probe 7 vs. Probe 8) by Group (C vs. B).

We therefore adopted a linear mixed-effects approach to interpreting our data, based on methods described by Quené and Van den Bergh (2004).[1] We tested the following model

$$y_{ij} = \beta_{0ij}\mathrm{CP1} + \beta_1\mathrm{C}\Delta\mathrm{P1P2}_{ij} + \cdots + \beta_{10}\mathrm{C}\Delta\mathrm{P10P11}_{ij} + \beta_{11}\mathrm{P1}\Delta\mathrm{AC}_{ij} + \cdots + \beta_{21}\mathrm{P11}\Delta\mathrm{AC}_{ij} + \beta_{22}\mathrm{P1}\Delta\mathrm{BC}_{ij} + \cdots + \beta_{32}\mathrm{P11}\Delta\mathrm{BC}_{ij} + u_{0j} + u_{33j}\mathrm{probe}_{ij} + e_{0ij}$$

where y is the score for a dependent variable for participant j at probe i. P1 to P11 represent Probe 1 to Probe 11, and A, B, and C represent Groups A, B and C. Fixed parameters are weights on dummy variables which, taken together, capture the score for Group C at baseline (β_0), the change in score for Group C between adjacent probes (β_1 to β_{10}), the deviation of the score for Group A from the score for Group C at each probe (β_{11} to β_{21}), and the deviation of the score for Group B from the score for Group C at each probe (β_{22} to β_{32}).

Modelled in this way, significance tests on the parameters for the deviation between control and intervention groups on the post-component probe exactly capture the interaction effect that provides evidence for the efficacy of that component. So, for example, evidence for the effectiveness of the Modelling-and-Shared-Reflection

[1] One approach to exploring these interactions would be to conduct four separate mixed two-way ANOVAs and then follow up with pairwise comparisons. This approach is suboptimal. It does not adequately capture the specific interaction effects that we are looking for—greater improvement in intervention relative to control is associated with more than one pattern of pairwise effects. A second disadvantage of ANOVA in that it requires the (typically unwarranted) assumption that between-participant variance is homoscedastic. Multilevel (linear mixed effects) modelling offers a more flexible approach to statistical analysis in this context, avoiding these problems.

Table 3
Standardized effect of each teaching component for all dependent variables.

	Modelling and shared revision			Declarative			Peer feedback			Solo practice		
	Phase 1 A vs. C	Phase 1 B vs. C	Phase 2 C vs. A, B	Phase 1 A vs. C	Phase 1 B vs. C	Phase 2 C vs. A, B	Phase 1 A vs. C	Phase 1 B vs. C	Phase 2 C vs. A, B	Phase 1 A vs. C	Phase 1 B vs. C	Phase 2 C vs. A, B
Final Text												
Holistic quality	.74*	1.00**	1.21** 1.06**	.69*	.16	–.34 –.27	.45	–.12	.23 –.08	–.64*	.61*	.07.25
Coherence	1.22**	1.32**	1.92** 1.43**	.71*	.41	–.20 –.03	.54	.02	.46.19	–.76*	.02	.27.07
Structure	.79**	.93**	1.51** 1.10**	.83**	.54	–.35 –.26	.20	–.20	.11.08	–.53	.01	.20.29
Simple coherence	1.79**	–.51	–.22 –.48	–.75*	–.57	–.14.14	–.39	.12	–.11 –.02	.20	.19	.09.21
Advanced coherence	.90*	1.09**	1.17** 1.89**	.66	.37	.89**.80*	–.34	.60	.22.35	.62	.48	.62*.64*
Paragraph count	.71*	.82*	1.35** 1.02**	.50	.58	.28.76**	.32	.01	–.63* –.82**	–.43	.07	.58*.33
Word count	.78**	.18	1.08**.26	.49*	–.14	–.26 –.11	–.38	.16	–.45 –.47	.04	.51	–.26 –.47
Written plan												
Structure	1.01**	.76*	1.31**.74**	.16	.34	.17.81**	.07	–.07	.13 –.47	.14	.03	–.27 –.27
Strategy	.66*	.43	.33 –.12	.72*	–.13	.02.24	.03	–.08	–.27 –.00	–.36	–.34	.37.11
Objective	.87*	.96**	1.40**.78**	.39	.14	–.12.31	.32	.44	.81**.29	–.05	.29	–.50* –.33
Writing process												
Planning	.53*	.36	.99**.46	.54*	–.20	–.47.14	–.20	–.23	.03 –.64*	–.47	–.01	.07.55
Drafting	.74*	.52	.38.08	.11	–.53	.18 –.13	.14	.29	–.64 –1.04**	–.02	.13	.65 1.14**
Revising	–.61	.97**	–.12 –.16	.60	.08	–.01.51	.08	–.49	–.39 –.31	–.26	.01	.10.14
Total time	.78*	.86**	1.15**.48	.65*	–.52	–.34.09	–.05	–.09	–.59 –1.40**	–.46	.08	.61 1.35**

Note: Effect sizes are for the interaction between probe (pre-component vs. and post-component) and group (control vs. training, as detailed in column header).
* $p < .05$, ** $p < .01$.

component would be found in a statistically significant effect of β_{12}. β_{12} is the weight on the parameter representing the difference between Groups A and C in the test immediately after delivery of the Modelling-and-Shared-Reflection component in Phase 1, controlling for differences between pre- and post-observation scores for Group C, and pre-observation score for Group A. We would also want to see this effect replicated in β_{24} (deviation of Group B from Group C on the probe immediately after Group B has received the Modelling-and-Shared-Reflection component) and in β_{19} and β_{30} (deviations of Group C from Group A and from Group B on the probe immediately after Group C has received the Modelling-and-Shared-Reflection component). The statistical significance of these parameters was established by evaluating $\beta/standard\ error\,(\beta)$ against the normal distribution (a Z test). Alpha for individual effects was set at .05.[2]

The model therefore estimated 33 fixed parameters, four random parameters, and had 681 degrees of freedom. The random terms were as follows: variance at intercept for participant and for probes (variance of u_{0j} and of e_{0ij}), participant variance associated with a dummy variable representing time-of-probe as a linear sequence (Probe 1 = –5 to Probe 11 = 5; variance of u_{33j}), and covariance between u_{33j} and u_{0j}.

To permit effect size calculations variance between students at each measurement occasion was calculated from the estimated parameters as follows:

$$\mathrm{VAR}\,(\mathrm{students}\,|\,\mathrm{Probe}) = \mathrm{Var}\,(\mathrm{u0j}) + 2 * \mathrm{Probe} * \mathrm{Cov}\,(\mathrm{u0j}, \mathrm{u1j}) + \mathrm{Probe2} * \mathrm{Var}\,(\mathrm{u1j})$$

Effect sizes were calculated by dividing the relevant parameter estimate by the sum of the square roots of this between-students variance estimate and the residual within-students.

Standardized effect sizes, and statistical significance, for all dependent variables, are reported in Table 3.

3.2.2. Text quality

Observed means for text quality measures are reported in Table 4. Across the Modelling-and-Shared-Reflection component structure, coherence and holistic quality all improved in all three groups, relative to the controls. This was true in both Phase 1 (compare–contrast essays with Group C as control, A vs. C and B vs. C) and in Phase 2 (opinion essays with Groups A and B as controls, C vs. A and C vs. B). As can be seen from Table 3, effects were all statistically significant and effect sizes were greater than .7 in all cases, and in most cases greater than 1. We did not find similarly clear effects (replicated across all four relevant contrasts) for any of the other three instructional components. Group A showed statistically significant improvement relative to control in all three measures as a result of the Declarative component, no statistically significant effect of Peer feedback and a statistically significant decline in performance (for structure and coherence, but not for holistic quality) following Solo Practice. Group B showed no further statistically significant improvement or decline in performance after the Modelling-and-Shared-Reflection component, with the exception of a significant improvement following Solo Practice (the opposite effect to that observed in Group A). In Phase 2 Group C showed no further statistically significant change in performance, relative to either of the other groups, across the Declarative, Peer Feedback, and Solo Practice components.

We predicted an increase in the use of advanced cohesion-maintaining devices, including paragraphing, as a result of training, indicating greater meta-linguistic awareness and use of argumentation. We did not make any prediction about changes in the use of simple coherence-maintaining devices (lexical repetition, use of connectives, and so forth) or in text length as a result of intervention, and found no consistent pattern of effects. Use of advanced cohesion devices and use of paragraphing increased following the Modelling-and-Shared-Reflection component for all four groups. Effects were statistically significant, and, with one exception, large (>.8). Use of advanced cohesion devices remained high, relative to baseline, on all subsequent probes in all three groups, but only showed further statistically significant increases in Group C (following Declarative instruction, and following Solo Practice). There was no further statistically reliable change across components in

[2] Some readers may question our decision not to apply a Bonferroni or similar correction for familywise error rate. Note, however, that the "internal replication" achieved by the present design makes it very improbable that we claim effects where none exist. We claim that an instructional component is effective only if all four associated comparisons give effects that are significant at $p < .05$. The combined probability of falsely claiming an effect for the component is therefore substantially less than .0001. Taking into account the fact that there are four components, any one or more of which might show an effect, raises the probability of Type 1 error, but it still remains below .0001. Our research design therefore makes Type 1 errors very improbable.

Table 4
Observed means for reader-based quality ratings. Standard deviations in parenthesis.

	Group A	Phase 1 Baseline	M&SR	Declarative	Peer feedback	Solo		Phase 2 Baseline	M&SR	Declarative	Peer feedback	Solo
	Group B		Baseline	M&SR	Declarative	Peer feedback	Solo					
	Group C							Baseline	M&SR	Declarative	Peer feedback	Solo
Holistic quality	Group A	2.40 (.60)	3.00 (.97)	3.07 (1.20)	3.58 (.85)	3.14 (1.03)		3.45 (1.00)	2.86 (.92)	3.60 (1.16)	3.21 (.98)	3.62 (.92)
	Group B		2.14 (.67)	2.64 (1.20)	2.71 (.77)	2.93 (1.14)	3.63 (1.22)	3.41 (.84)	2.95 (.91)	3.61 (.92)	3.52 (.97)	3.98 (.80)
	Group C	2.68 (.63)	2.35 (.67)	1.95 (.46)	2.08 (.55)	2.20 (.66)	2.15 (.54)	2.15 (.54)	2.63 (.53)	3.20 (1.02)	3.79 (.98)	3.63 (1.04)
Coherence	Group A	1.10 (.30)	2.13 (.74)	2.50 (.85)	2.85 (.76)	2.31 (.83)		2.64 (.98)	1.88 (.69)	2.43 (.88)	2.24 (.82)	2.52 (.86)
	Group B		1.19 (.33)	2.10 (1.00)	2.21 (1.01)	2.38 (.97)	2.65 (.78)	2.80 (.68)	2.45 (.83)	2.84 (1.03)	2.89 (.86)	3.14 (.79)
	Group C	1.21 (.48)	1.23 (.38)	1.20 (.38)	1.14 (.33)	1.18 (.41)	1.30 (.38)	1.53 (.44)	2.35 (1.04)	2.87 (.96)	3.11 (.89)	2.98 (1.03)
Structure	Group A	1.69 (.51)	2.63 (.67)	2.71 (.86)	2.93 (.77)	2.50 (.72)		3.00 (.65)	2.31 (.68)	2.93 (.81)	2.71 (.83)	2.98 (.70)
	Group B		1.90 (.62)	2.19 (.91)	2.50 (.81)	2.45 (.97)	2.95 (.93)	2.89 (.60)	2.52 (.79)	3.07 (.84)	2.86 (.76)	3.21 (.68)
	Group C	1.87 (.52)	2.03 (.47)	1.65 (.52)	1.75 (.43)	1.68 (.54)	2.03 (.53)	2.20 (.47)	2.65 (.80)	3.16 (.62)	3.03 (.72)	3.30 (.71)

Note: Group C is control for Phase 1. Groups A and B are controls for Phase 2. Values are from probes immediately following each component (or equivalent in control conditions). M&SR = Modelling and Shared Reflection.

the use of paragraphs, with the exception of a decrease in use by Group C following the Peer Feedback component. For both measures, increases after Modelling-and-Shared-Reflection remained after the subsequent three components, and in Groups A and B, transferred to, and were maintained during, Phase 2.

As can be seen from Table 5 neither word count only increased, relative to controls, in two of the four groups following the Modelling-and-Shared-Reflection component, and showed no further statistically significant change after subsequent components. Lack of effect for word count suggests that improvements in quality were not simply artefacts text length.

3.3. *Writing process*

Total time on task, reported in Table 6, increased across the Modelling-and-Shared-Reflection component, relative to controls, in all three groups. This effect was statistically significant in three out of the four possible contrasts, the exception being for Group C compared to Group B in Phase 2. There were no further clear effects of instructional components. We did not observe consistent effects of intervention on the time taken in the three different writing-process phases (planning, drafting, revising). Time in the planning phase increased relative to controls across the Modelling-and-Shared-Reflection component by a mean of between 2 and 3 minutes in all three groups, but this effect only reached statistical significance for Group A in Phase 1 and for Group C relative to Group A in Phase 2. There were no clear effects on planning time for other training components. We did not find clear effects of training on time spent drafting text or on time spent revising.

Analyses of the notes that students made prior to drafting full text showed clear effects of training (Table 7). Again, these effects were associated almost exclusively with the Modelling-and-Shared-Reflection component. Notes showed substantially more structure after Modelling-and-Shared-Reflection, with statistically significant effects for all comparisons. Prior to training, students did not, with very few exceptions, indicate rhetorical goals in their notes. There was an increasing tendency to do this throughout training, with the largest increase resulting from the Modelling-and-Shared-Reflection component (statistically significant for all four comparisons). Students' notes only very rarely mentioned strategy, and there was no reliable evidence that this was affected by intervention.

4. Discussion

Consistent with previous findings (Brunstein & Glaser, 2011; De La Paz & Graham, 2002; Glaser & Brunstein, 2007; Limpo & Alves, 2013; Torrance et al., 2007) this study confirms the benefits of strategy-focused writing instruction for typically developing late-primary and early-secondary students. All three classes showed gains in the quality of their writing following instruction, relative to controls who received instruction that was not strategy-focused. This was apparent both in reader-based and text-analytic measures of text quality. Following intervention students produced texts that were assessed as being more coherent, better structured, and being of generally better quality. Use of sophisticated devices for maintaining coherence increased, as did use of paragraphing. Increases in text quality were not consistently associated with more time spent planning in advance of drafting. However students' written plans showed consistent increase in tendency to specify text structure, and to state communicational goals. These effects were tested and found in three separate groups, and transferred to a different expository genre. Our study arguably therefore provides particularly robust additional evidence for the benefits of strategy-focused writing instruction.

Table 5
Observed means for text-based measures from final text. Standard deviations in parenthesis.

Component		Phase 1							Phase 2				
	Group A	Baseline	M&SR	Declarative	Peer feedback	Solo			Baseline	M&SR	Declarative	Peer feedback	Solo
	Group B		Baseline	M&SR	Declarative	Peer feedback	Solo	Group C					
Simple coherence ties	Group A	14.9 (4.1)	18.3 (2.9)	16.7 (3.5)	17.4 (4.5)	17.5 (2.3)			11.4 (2.8)	10.1 (3.4)	12.0 (3.2)	10.7 (4.5)	13.3 (4.0)
	Group B		17.7 (5.7)	17.1 (3.1)	17.1 (4.3)	16.9 (4.0)	15.5 (3.5)		11.5 (4.2)	11.0 (4.3)	11.9 (3.5)	10.4 (3.6)	12.0 (4.6)
	Group C	18.6 (7.6)	14.7 (4.1)	16.1 (4.6)	18.4 (4.2)	17.6 (4.4)	15.6 (2.6)		11.6 (3.6)	9.5 (3.8)	10.7 (3.3)	9.1 (3.1)	12.1 (3.2)
Advanced coherence ties	Group A	.24 (.64)	1.45 (1.41)	2.51 (1.44)	2.17 (.96)	2.72 (1.83)			1.79 (1.52)	1.64 (1.69)	1.47 (1.60)	1.51 (1.25)	1.74 (1.51)
	Group B		.27 (.59)	1.98 (2.13)	2.69 (1.70)	3.22 (1.93)	3.89 (2.81)		4.09 (3.45)	2.79 (1.69)	2.80 (1.78)	2.66 (1.49)	2.87 (1.61)
	Group C	.58 (.90)	.46 (.82)	.56 (.88)	.69 (.86)	.30 (.64)	.19 (.48)		.70 (1.40)	2.56 (1.98)	4.04 (2.30)	4.39 (2.85)	3.46 (2.33)
Paragraph count	Group A	1.57 (.81)	2.45 (1.19)	3.14 (1.20)	3.45 (.76)	3.05 (1.15)			2.81 (1.12)	2.19 (1.33)	2.57 (1.17)	2.95 (1.24)	2.48 (1.25)
	Group B		1.76 (1.00)	2.76 (1.09)	3.29 (.85)	3.43 (.98)	2.95 (1.15)		2.95 (.95)	2.73 (1.16)	2.55 (1.06)	3.14 (.94)	3.00 (1.18)
	Group C	1.74 (.73)	1.90 (1.17)	2.05 (1.00)	1.94 (1.06)	2.10 (1.62)	1.55 (.69)		1.80 (1.06)	2.75 (1.71)	3.53 (1.74)	3.00 (1.56)	3.35 (1.63)
Word count	Group A	88.3 (24.3)	92.4 (30.6)	102.2 (29.1)	104.5 (28.4)	95.5 (29.4)			92.9 (28.5)	77.4 (24.9)	80.4 (21.5)	82.1 (16.5)	92.1 (28.2)
	Group B		82.0 (21.1)	83.8 (21.5)	91.4 (20.2)	87.3 (22.2)	96.3 (20.5)		69.9 (13.6)	72.7 (17.6)	73.0 (13.2)	75.1 (13.8)	80.5 (18.5)
	Group C	99.1 (25.3)	83.1 (28.0)	80.4 (27.5)	90.9 (32.2)	83.6 (26.5)	81.1 (23.6)		73.4 (34.9)	82.8 (26.2)	78.1 (25.1)	70.6 (25.9)	87.2 (26.3)

Note: Group C is control for Phase 1. Groups A and B are controls for Phase 2. Values are from probes immediately following each component (or equivalent in control conditions). M&SR = Modelling and Shared Reflection.

Table 6
Observed mean time taken in planning, drafting and revising texts (minutes). Standard deviations in parenthesis.

Component		Phase 1							Phase 2				
	Group A	Baseline	M&SR	Declarative	Peer feedback	Solo			Baseline	M&SR	Declarative	Peer feedback	Solo
	Group B		Baseline	M&SR	Declarative	Peer feedback	Solo	Group C					
Planning	Group A	4.5 (5.4)	7.3 (5.9)	10.7 (8.3)	9.9 (9.3)	7.5 (7.4)			4.5 (5.9)	2.4 (3.7)	3.9 (4.1)	2.6 (3.9)	4.7 (4.8)
	Group B		7.6 (4.2)	9.9 (5.9)	9.1 (4.4)	8.0 (6.2)	6.7 (6.1)		6.0 (4.4)	6.3 (4.3)	5.0 (3.9)	7.6 (6.5)	7.0 (4.2)
	Group C	7.3 (6.1)	7.0 (3.8)	7.6 (4.1)	7.7 (4.1)	7.7 (3.7)	6.4 (3.7)		6.1 (3.5)	8.8 (3.9)	8.4 (3.7)	7.0 (3.8)	9.4 (4.9)
Drafting	Group A	10.0 (5.4)	11.2 (5.7)	12.1 (4.7)	12.1 (7.5)	11.8 (4.9)			10.9 (5.1)	9.8 (5.7)	8.6 (3.5)	10.7 (3.7)	9.1 (3.4)
	Group B		6.8 (3.3)	9.4 (7.9)	6.7 (2.9)	7.6 (3.3)	8.1 (5.3)		6.3 (3.5)	6.6 (3.5)	6.5 (3.4)	9.5 (4.1)	6.0 (2.4)
	Group C	9.7 (6.1)	7.5 (4.6)	7.9 (3.9)	7.4 (5.4)	7.1 (2.8)	7.1 (2.3)		7.0 (3.4)	7.5 (3.8)	6.8 (2.5)	6.5 (3.4)	7.4 (3.2)
Revising	Group A	1.0 (2.0)	.48 (.68)	1.1 (1.2)	1.0 (1.2)	.71 (1.3)			1.0 (1.1)	.70 (.90)	1.3 (1.4)	1.5 (1.7)	1.7 (1.8)
	Group B		.27 (.46)	1.3 (1.6)	1.2 (1.5)	.68 (.78)	.41 (.67)		1.2 (1.6)	1.0 (1.7)	1.0 (.9)	1.1 (1.0)	1.2 (1.3)
	Group C	1.3 (1.8)	1.6 (1.8)	1.4 (1.0)	1.2 (1.4)	1.3 (1.4)	1.0 (1.7)		1.6 (1.1)	1.2 (1.2)	1.8 (1.6)	1.4 (1.3)	1.8 (1.7)
Total time on task	Group A	15.8 (6.2)	19.0 (7.0)	23.9 (9.1)	23.0 (13.4)	20.0 (8.2)			16.3 (6.8)	12.9 (6.2)	13.7 (3.4)	14.9 (3.8)	15.2 (4.9)
	Group B		14.7 (6.4)	20.6 (9.5)	17.0 (7.2)	16.3 (7.3)	15.2 (7.6)		13.5 (4.1)	13.8 (4.2)	12.4 (4.2)	17.8 (4.8)	14.2 (5.5)
	Group C	18.3 (7.9)	16.1 (3.9)	16.9 (2.3)	16.3 (7.1)	16.1 (3.9)	1.0 (1.7)		14.5 (3.9)	14.7 (3.7)	17.9 (5.4)	16.9 (4.6)	14.8 (4.9)

Note: Group C is control for Phase 1. Groups A and B are controls for Phase 2. Values are from probes immediately following each component (or equivalent in control conditions). M&SR = Modelling and Shared Reflection.

Table 7
Observed means from evaluation of notes written prior to drafting. Standard deviations in parenthesis.

		Phase 1						Phase 2				
	Group A	Baseline	M&SR	Declarative	Peer feedback	Solo		Group C				
	Group B		Baseline	M&SR	Declarative	Peer feedback	Solo	Baseline	M&SR	Declarative	Peer feedback	Solo
Structure	Group A	.52 (.68)	1.60 (1.19)	1.81 (1.40)	1.95 (1.28)	1.81 (1.33)		1.38 (1.40)	.81 (1.08)	1.00 (1.18)	.80 (1.24)	1.19 (1.21)
	Group B		1.05 (.38)	1.86 (1.15)	2.19 (.98)	1.90 (1.14)	1.89 (1.20)	2.05 (1.17)	2.05 (1.17)	1.59 (1.18)	2.05 (1.09)	2.38 (.97)
	Group C	1.05 (.71)	1.15 (.49)	1.25 (.64)	1.28 (.46)	1.05 (.69)	1.00 (.73)	1.35 (.67)	2.10 (.91)	2.53 (.61)	2.42 (.77)	2.50 (.69)
Strategy	Group A	.00 (.00)	.30 (.73)	.57 (.87)	.65 (.75)	.38 (.67)		.29 (.64)	.10 (.44)	.19 (.60)	.30 (.73)	.19 (.60)
	Group B		.05 (.22)	.19 (.51)	.19 (.51)	.05 (.22)	.00 (.00)	.00 (.00)	.00 (.00)	.00 (.00)	.00 (.00)	.00 (.00)
	Group C	.11 (.32)	.10 (.45)	.05 (.22)	.11 (.32)	.00 (.00)	.10 (.45)	.05 (.22)	.00 (.00)	.11 (.46)	.11 (.46)	.15 (.49)
Objective	Group A	.00 (.00)	.50 (.89)	.76 (1.00)	1.00 (1.03)	1.05 (1.02)		.76 (1.00)	.38 (.81)	.52 (.87)	.20 (.62)	.52 (.87)
	Group B		.00 (.00)	.67 (.97)	.76 (1.00)	1.19 (.98)	1.37 (.96)	1.09 (1.02)	1.23 (.97)	1.00 (1.02)	1.18 (1.01)	1.33 (.97)
	Group C	.11 (.46)	.00 (.00)	.00 (.00)	.00 (.00)	.10 (.45)	.00 (.00)	.10 (.45)	.90 (1.02)	.95 (1.03)	1.37 (.96)	1.20 (1.01)

Note: Group C is control for Phase 1. Groups A and B are controls for Phase 2. Values are from probes immediately following each component (or equivalent in control conditions). M&SR = Modelling and Shared Reflection.

Groups A and B returned to non-strategy focused instruction in Phase 2. Despite this, and the fact that assessments in Phase 2 involved a different, though related, genre, quality improvements were maintained. Again consistent with findings from previous research, benefits of strategy-focused instruction were persistent.

Our main focus was on the individual effects of each component. Improvements were reliably associated with just the first, Modelling-and-Shared-Reflection component, in which students observed a mastery model of the target planning and drafting strategies, followed by structured reflection on what they had observed. The ground covered by the instruction in the Modelling-and-Shared-Reflection component and in the Declarative component was the same. In both students were exposed to strategies associated with identifying audience needs, setting rhetorical and communicational goals, generating content, and so forth. During modelling these remained fairly implicit in the instructor's script. The instructor modelled the use of all of the target strategies, but did not at any point provide names or labels for these or their sub-components. In her think aloud the instructor identified strategic steps in her writing procedure using meta-comments ('Now I'd better switch to…; Do I know anything more about…?). Strategies were, however, bound to the production of a particular text, rather than being abstracted as general principles. Students then reflected, as a group, on what they had observed, with the instructor guiding but not contributing substantively to the discussion. This pooling of observations may have served to make explicit and possibly to label some of the strategies used by the model. It will also mean that learning for each individual student will have been based not just on their own observation, but partly on the observations of other students in the class. Crucially, however, reference to strategy in this component was bound to the process of writing a particular text and not explicitly presented as general principles.

It was not until the Declarative component that strategies were abstracted and bound to mnemonics. The learning associated with this component served to maintain gains achieved as a result of Modelling-and-Shared-Reflection. In two of the three groups declarative instruction resulted in improved reader-based quality ratings, relative to control, although these were only statistically reliable in one of these groups. In the Group C (Phase 2, opinion essay) there was no improvement in reader-based scores after declarative instruction (following very large improvements as a result of Modelling-and-Shared-Reflection) but substantial and reliable gains in use of the advanced coherence-tie devices associated with this instruction. One possible interpretation of this is that modelling resulted in a balanced use of these devices but that students then overused them, given the opinion-essay genre, when these were taught again in the Declarative component, with resulting detrimental effect to reader-based quality ratings. Our research design does not permit the strong conclusion that, in the present context, modelling and shared reflection is more effective than declarative instruction. It is possible that if the order of these components were reversed then we might have found similar gains for direct instruction. Our results do, however, show that even without first providing students with meta-knowledge for framing the strategies that they observe, modelling and sharing observations can be effective in teaching writing strategies. This is a new finding. Previous research has found that declarative instruction, in the absence of modelling, can be effective in developing writing skills (Sawyer et al., 1992; but see Danoff et al., 1993) albeit in struggling writers rather than the fully-range classes that were the focus of the present study. Our findings suggest that, for typically developing writers at least, the reverse may also be true.

There are at least three, mutually-compatible reasons why modelling, combined with shared reflection, in the absence of formalized strategy instruction might be a particularly effective way of improving writing performance. First, observation might have

motivational benefits. Observing someone modelling a particular strategy provides information not just about that strategy, but about the fact that it can be performed by another person, with resulting benefits for self-efficacy. Self-regulation through strategy use requires not just knowledge of the strategy, but also the belief that its use is possible, that it will work and that it gives benefit. These beliefs are likely to be better fostered by observing a model than by simply being told that the strategy is beneficial. Second, observation may bind strategy knowledge to performance. Modelling exposes students to the strategy-in-action in a way that does not necessarily occur when strategies are directly taught. It necessarily shows in a direct way how strategies are applied to a real life writing task. Third, it may be that, for developing writers, having to maintain in mind explicit, abstracted strategy knowledge during writing may have detrimental effects on other cognitive processes associated with text production. Rehearsing a strategy mnemonic may provide internal control over the writing process, but this may be at the cost of diverting attention from other, important language processes. Developing writers may lack automaticity in the lower-level mechanics of word retrieval and sentence production. If this is the case, then retrieving explicit strategy meta-knowledge—recalling a mnemonic representing planning steps, for example—may place excessive load on cognitive resources (the "double challenge" discussed by Rijlaarsdam and co-workers; Rijlaarsdam et al., 2011; Rijlaarsdam & Couzijn, 2000). Learning by observation, if it bypasses the need for explicit strategy representations, will reduce this challenge.

There are both theoretical and practical implications of these findings. Historically researchers have tended to assume that, contrary to expertise development in other domains, writing experts make use of strategic (explicit, metacognitive) representations (rhetorical goals, planning procedures, audience models, and so forth; Flower & Hayes, 1980; Hayes & Nash, 1996; Kellogg, 2008; Scardamalia & Bereiter, 1991). Our findings confirm the value of giving students an understanding of, and focusing their attention on, higher-level textual features, and of showing them possible planning behaviours that make space for applying this understanding to their own text. However, allowing this understanding to develop through observation of a model and sharing these observations, appears to be sufficient to give substantial gains in writing performance.

From a practical perspective, our findings suggest that teachers of upper-primary classes will find observational learning, in the absence of more formal instruction, effective in developing students' writing performance. Note, however, that in the present study observation was combined with class discussion ("shared reflection") of what had been observed. Shared reflection did not add new content, but was included to encourage students, and particularly weaker students, to engage with and retain what they had observed. Whether it was in fact necessary to students' learning could usefully be the focus of future studies.

These conclusions need qualifying in two ways. First, and most obviously, the design of the present experiment does not permit the conclusion that modelling (combined with shared reflection) is more effective than declarative instruction. It is possible that if the order in which components were taught were reversed then declarative instruction would have shown similar benefits. A direct comparison of the benefits of these two forms of instruction may be of practical importance, and therefore might usefully be the focus of future research. Our design does, however, provide a particularly robust test of the hypothesis that modelling combined with shared reflection, *in the absence of declarative instruction*, can be effective. Second, our design demonstrates that, consistent with previous findings, the benefits of packages of strategy-focused instruction are sustained. However, it is possible that this sustained effect, demonstrated by the continuing improved performance of Groups A and B in Phase 2, results from some or all of declarative instruction, peer feedback, and solo practice. Our findings do not rule out the possibility that the effects of modelling and shared reflection, on their own, are temporary.

The other social-learning component in our intervention involved peer feedback. As with declarative instruction, this component also maintained gain in performance but did not give reliable additional benefit. This contrasts with findings of previous research (Graham et al., 2005; Harris et al., 2006; Yarrow & Topping, 2001). Failure to find effects of peer feedback may result from the lack of value in feedback, *per se*, although this would seem unlikely. Alternatively in the present context, where tasks were relatively unconstrained, and strategies relatively complex, the quality of the feedback that students received from their peers might have been relatively poor. Note also that for peer feedback to show effects in the present study it would have to give benefit over and above the substantial gains resulting from the first (and in some cases the second) components. Failure to find effects of peer feedback does not, therefore, indicate that there is never value in students scaffolding each-other's writing strategies. Had our sample been at a different developmental stage, been performing a more constrained task and/or been given more extensive training then we may have found effects.

Finally, we would like to comment on the design of the present study. Evaluating whether or not particular forms of instruction result in self-regulated performance in school-age students necessarily requires sustained training over multiple sessions. For both practical reasons and for findings to have practical application, this training will necessarily take place in a classroom context. However, this typically means that intact classes, rather than students, are allocated to conditions (and, arguably, there are methodological reasons why creating new classes through random allocation is undesirable in this context). Problems associated with whole-class allocation are not solved simply by each condition including more than one class, unless the number of classes is large. The lagged-group, cross-panel design of the present research helps to overcome this problem. If hypothesized effects are only claimed when they are seen in all groups of participants, as was the case in the present research, this design provides a particularly robust evaluation of the effects of instruction.

In summary, therefore, our findings suggest that for typically-developing, upper-primary students, observation of a mastery model followed by whole-class reflection is sufficient to promote the development of self-regulated writing strategies. Consistent with previous findings, we found that the development of these strategies resulted in an improvement in the quality of students' texts.

Appendix. Abridged version of the script that formed the basis for teacher modelling in Phase 1 Session 1 (planning a compare–contrast essay). Translated from Spanish

General instructions

Your aim is to make explicit your thoughts as you plan the text.

You should demonstrate effective planning strategies (present a mastery model) but omit explicit reference to these strategies. For instance, avoid expressions such as: *with the vowel A… Audience, now I have to remember the I vowel…what kind of Ideas can I write, the first thing to do in the planning and I have to remember the strategy POD + the vowels*, and so forth.

During the modelling, you should use self-questions and explicit self-instructions answering those self-questions in relation to:

- Regulation of what you are doing. For example: self-questions: *what do I have to do? What is the first thing that I have to do? How will I do it?* Self-instructions: *the first thing that I have to do is*

decide the purpose of my text, and the kind of text that I am going to write…, Also, I have to think…

- Regulation during the writing process of what you have done so far. For example: self-questions: *Have I thought of enough ideas? Is that the correct goal for my text?* 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? Can I start to write my text now?* Self-instructions: *Now, I am ready to write my text. Yes, I have enough differences and similarities between the themes, 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; this is a good idea; I am a really good writer; I am so imaginative!*

Detailed example

Taking into account these general guidelines, the following is a detailed example of the kinds of statements that you should include when you think aloud during the planning of the compare–contrast text (without any reference to the POD + the vowels strategy or the steps of the planning process).

I have to write a compare–contrast essay that explains the similarities and the differences between two aquatic sports, such as, water polo and swimming. I must concentrate on this task and on the text that I have to write. And I am confident that I can write this well. That way I can be sure to get good and original ideas for my text. Well, I think that I am ready for starting planning my text. I'll write some notes for my text where I will write my ideas for the introduction, development and conclusion. But, before doing anything, I should think about who will read my text. I should try to make it understandable and interesting for them. They'll like what I write. So, in this case, my text is for students of 6th grade in my class. So, I have to take this into account when I write my text. I'm going to need to make it suitable for that age. I also need to remember the purpose of my text. That is explaining the differences and similarities that exist between two aquatic sports, like: water polo and swimming. I'll write about those in my introduction [write notes]. *Well, now that I have these two things, I am going to continue, what ideas can I write in my text? First, I am going to think about general ideas that I know. So, first, I am going to think…in what ways are water polo and swimming alike? Or what do they have in common? What are the main ideas here? Well, between similarities, the main ideas are* [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 aloud. Ask whether ideas are main or secondary. Possibly generate more ideas then….] *I think that now there are enough ideas, I will not need more ideas in this part. Now what about the other part of the text? What are the differences between water polo and swimming?* [repeat the same process then…] *I'll read these all again* [read list]. *I am doing this very well! I'm really proud of myself! So have I finished? Ah, no. There's that last part; the conclusion. I need to sum up my ideas and write my thoughts about them. Something like…*[write a concluding sentence] *Ok, I think I'm almost finished. I'll just read it again* [read through notes, with self-praise and other comments, maintaining students' interest].

Acknowledgements

During this research funds from Ministerio de Economía y Competitividad, Reference EDU2010-18219 were awarded to the first author. We would like to thank the school Nuestra Madre del Buen Consejo, Padres Agustinos de León and to Mª Ángeles Domínguez for their assistance. Correspondence should be directed to Raquel Fidalgo (rfidr@unileon.es) and to Mark Torrance (mark.torrance@ntu.ac.uk).

References

Alexander, P. A., Graham, S., & Harris, K. R. (1998). A perspective on strategy research: Progress and prospects. *Educational Psychology Review*, *10*(2), 129–154.

Bereiter, C., Burtis, P. J., & Scardamalia, M. (1988). Cognitive operations in constructing main points in written composition. *Journal of Memory and Language*, *27*, 261–278. doi:10.1016/0749-596X(88)90054-X.

Bosque, I., & Demonte, V. (1999). *Gramática descriptiva de la lengua española. Vol 3: Entre la oración y el discurso [Descriptive grammar of the Spanish language: Vol 3: Between the sentence and the discourse]*. Madrid: Espasa Calpe, S.A.

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*, 405–415. doi:10.1037//0022-0663.94.2.405.

Braaksma, M. A. H., Rijlaarsdam, G., Van den Bergh, H., & Van Hout-Wolters, B. H. A. M. (2004). Observational learning and its effects on the orchestration of writing processes. *Cognition & Instruction*, *22*, 1–36. doi: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. doi:10.1037/a0024622.

Couzijn, M. (1999). Learning to write by observation of writing and reading processes: Effects on learning and transfer. *Learning & Instruction*, *9*, 109–142. doi:10.1016/S0959-4752(98)00040-1.

Danoff, B., Harris, K. R., & Graham, S. (1993). Incorporating strategy instruction within the writing process in the regular classroom. Effects on the writing of students with and without learning disabilities. *Journal of Reading Behavior*, *25*, 295–319. doi:10.1080/10862969009547819.

De La Paz, S., & Graham, S. (2002). Explicitly teaching strategies, skills, and knowledge: Writing instruction in middle school classrooms. *Journal of Educational Psychology*, *94*(4), 687–698. doi:10.1037//0022-0663.94.4.687.

Fidalgo, R., Torrance, M., & Garcia, J. N. (2008). The long-term effects of strategy-focussed writing instruction for grade six students. *Contemporary Educational Psychology*, *33*(4), 672–693. doi:10.1016/j.cedpsych.2007.09.001.

Fidalgo, R., Torrance, M., & Robledo, P. (2011). Comparación de dos programas de instrucción estratégica y autorregulada para la mejora de la competencia escrita. [Comparison of two self-regulated and strategic instructional programs for improving writing competence]. *Psicothema*, *23*(4), 672–680.

Flower, L. S., & Hayes, J. R. (1980). The dynamics of composing: Making plans and juggling constraints. In L. W. Gregg & E. R. Steinberg (Eds.), *Cognitive processes in writing* (pp. 31–50). Hillsdale, NJ: Lawrence Erlbaum Associates.

García, J. N., de Caso-Fuertes, A.-M., Fidalgo-Redondo, R., Arias-Gundín, O., & Torrance, M. (2010). Spanish research on writing instruction for students with and without learning disabilities. In C. Bazerman, R. Krut, K. Lunsford, S. McLeod, S. Null, P. Rogers, et al. (Eds.), *Traditions of writing research: Traditions, trends, and trajectories* (pp. 71–82). New York: Routledge.

Glaser, C., & Brunstein, J. C. (2007). Improving fourth-grade students' composition skills: Effects of strategy instruction and self-regulation procedures. *Journal of Educational Psychology*, *99*(2), 297–310. doi:10.1037/0022-0663.99.2.297.

Graham, S., & Harris, K. R. (1989). Components-analysis of cognitive strategy instruction—Effects on learning-disabled students compositions and self-efficacy. *Journal of Educational Psychology*, *81*(3), 353–361. doi:10.1037//0022-0663.81.3.353.

Graham, S., Harris, K. R., & Mason, L. (2005). Improving the writing performance, knowledge, and self-efficacy of struggling young writers: The effects of self-regulated strategy development. *Contemporary Educational Psychology*, *30*(2), 207–241. doi:10.1016/j.cedpsych.2004.08.001.

Graham, S., McKeown, D., Kiuhara, S., & Harris, K. R. (2012). A meta-analysis of writing instruction for students in the elementary grades. *Journal of Educational Psychology*, *104*(4), 879–896. doi:10.1037/a0029185.

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

Halliday, M., & Hassan, R. (1976). *Cohesion in English*. London: Longman.

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

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

Hayes, J. R. (1996). A new framework for understanding cognition and affect in writing. In C. M. Levy & S. Ransdell (Eds.), *The science of writing: Theories, methods, individual differences, and applications* (pp. 1–26). Mahwah, NJ: Erlbaum.

Hayes, J. R., & Flower, L. S. (1980). Identifying the organisation of writing processes. In L. Gregg & E. R. Steinberg (Eds.), *Cognitive processes in writing* (pp. 3–30). Hillsdale, NJ: Erlbaum.

Hayes, J. R., & Nash, J. G. (1996). On the nature of planning in writing. In C. M. Levy & S. Ransdell (Eds.), *The science of writing: Theories, methods, individual differences and applications* (pp. 29–55). Mahwah, NJ: Erlbaum.

Kellogg, R. (2008). Training writing skills: A cognitive development perspective. *Journal of Writing Research*, *1*(1), 1–26.

Kitsantas, A., Zimmerman, B. J., & Cleary, T. (2000). The role of observation and emulation in the development of athletic self-regulation. *Journal of Educational Psychology*, *92*(4), 811–817. doi:10.1037/0022-0663.92.4.811.

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

Quené, H., & Van den Bergh, H. (2004). On multi-level modelling of data from repeated measures designs: A tutorial. *Speech Communication*, *43*, 103–121. doi:10.1016/j.specom.2004.02.004.

Raedts, M., Rijlaarsdam, G., Van Waes, L., & Daems, F. (2007). Observational learning through video-based models: Impact on students' accuracy of self-efficacy beliefs, task knowledge and writing performance. In S. Hidi & P. Boscolo (Eds.), *Studies of writing: Writing and motivation* (Vol. 19, pp. 219–238). Oxford: Elsevier.

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.), (Vol. 3, pp. 189–228). Handbook of educational psychology: Application of educational psychology to learning and teaching. Washington: American Psychological Society.

Rijlaarsdam, G., Braaksma, M., Couzijn, M., Janssen, T., Raedts, M., Steendam, E., et al. (2008). Observation of peers in learning to write. Practice and research. *Journal of Writing Research*, *1*(1), 53–83.

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

Rogers, L., & Graham, S. (2008). A meta-analysis of single subject design writing intervention research. *Journal of Educational Psychology*, *100*, 879–906. doi:10.1037/0022-0663.100.4.879.

Sanders, T. J. M., Spooren, W. P. M., & Noordman, L. G. M. (1992). Toward a taxonomy of coherence relations. *Discourse Processes*, *15*(1), 1–35. doi:10.1080/01638539209544800.

Sawyer, R. J., Graham, S., & Harris, K. R. (1992). Direct teaching, strategy instruction, and strategy instruction with explicit self-regulation—Effects on the composition skills and self-efficacy of students with learning-disabilities. *Journal of Educational Psychology*, *84*(3), 340–352. doi:10.1037//0022-0663.84.3.340.

Scardamalia, M., & Bereiter, C. (1991). Literate expertise. In K. A. Ericsson & J. Smith (Eds.), *Toward a general theory of expertise: Prospects and limits* (pp. 172–194). Cambridge: Cambridge University Press.

Schunk, D., & Zimmerman, B. (1997). Social origins of self-regulatory experience. *Educational Psychologist*, *32*(4), 195–208.

Spencer, S. L., & Fitzgerald, J. (1993). Validity and structure, coherence, and quality measures in writing. *Journal of Reading Behavior*, *25*(2), 209–231. doi:10.1080/10862969309547811.

Thurstone, L., & Thurstone, T. (2004). *Test de aptitudes escolares [Test of scholar aptitudes]* (11th ed.). Madrid: TEA Ediciones.

Torrance, M., Fidalgo, R., & Garcia, J. N. (2007). The teachability and effectiveness of cognitive self-regulation in sixth grade writers. *Learning and Instruction*, *17*(3), 265–285. doi:10.1016/j.learninstruc.2007.02.003.

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

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. doi:10.1037//0022-0663.94.4.660.