Effectiveness of an SRSD writing intervention for low- and high-SES children

Naymé Salas[1] · **Marilisa Birello[1]** · **Teresa Ribas[1]**

Abstract

SRSD interventions on writing have shown to be effective across a myriad of contexts and populations. Less is known, however, about their effectiveness for improving the writing skills of disadvantaged children, relative to their efficacy with other types of students. This is important, because low-SES children have been found to often display poor levels of written composition, which is key to preventing school failure. This study thus aimed to test the efficacy of an SRSD program for opinion essay writing in low- and mid-high SES classrooms. Participants were 645 children, attending 2nd and 4th grade classrooms, who were quasi-randomly assigned to the experimental (SRSD) or to the control condition. Children were assessed at pre- and post-test, where measures of text productivity, structural elements, and text quality were obtained. In addition, children's text spelling accuracy and reading comprehension skills were evaluated to test for developmental cascading effects. Children in the SRSD condition outscored the control group in all outcomes, regardless of their SES background. Moreover, some children in the SRSD group improved reading comprehension, but this added benefit was only observed in the mid-high-SES children. No group made gains on spelling accuracy at posttest. We discuss the theoretical and educational implications of our findings.

Keywords SRSD · Writing intervention · Catalan · SES · Developmental cascades

Writing is an intrinsic aspect of language development and it is crucial for personal and academic achievement (Ravid & Tolchinsky, 2002). It is imperative, then, that writing instruction is as efficient as possible. Any instructional writing practice should deliver observable improvement, attributable to the nature of the instruction. In addition, it should benefit a diversity of children, who show varying levels

✉ Naymé Salas
  nayme.salas@uab.cat

[1] Dept. de Didàctica de la Llengua i la Literatura, i de les Ciències Socials, Facultat de Ciències de l'Educació, Universitat Autònoma de Barcelona, Campus de Bellaterra, Edifici G5, 08193 Bellaterra, Barcelona, Spain

🖄 Springer

of writing skills, exposure to print, and general knowledge. Finally, researchers should be able to explain the mechanisms that make a given methodology effective. In this study, we examined whether a specific type of writing instruction, based on the development of self-regulated writing skills (self-regulated strategy development, SRSD, (Graham & Harris, 2009) was as effective for low socioeconomic status (SES) children as for children from a mid- or mid-high SES background. This is important because low-SES children have been found to lag behind age-matched peers on writing skills, increasing the risk for academic failure (e.g., Coker, 2006; Kim, Puranik, & Otaiba, 2015). We also explored whether the intervention, which targeted planning strategies for opinion-essay writing, showed cascading effects to two related skills: spelling accuracy (a within-domain skill) and reading comprehension (a cross-domain skill). This exploration would allow us to better understand the mechanisms underlying the efficacy of the writing intervention.

Self-regulated strategy development

Self-regulated Strategy Development (SRSD) for writing instruction was developed in the 1980's by Karen Harris (Graham & Harris, 2018) and is deeply rooted in a process view of writing (Graham & Harris, 2009; Hayes & Flower, 1980). For this reason, SRSD interventions usually target a specific writing process (e.g., planning, revising) by providing explicit strategies that are easy to remember and, therefore, to apply. Process-based writing strategies work best when paired with strategies that target a child's self-regulation skills. Furthermore, SRSD interventions often revolve around a concrete discourse genre or communicative situation, such as opinion-essay writing, narrative, or informational text writing (e.g., Graham & Perin, 2007; Graham, McKeown, Kiuhara, & Harris, 2012).

A typical SRSD writing intervention includes six recursive stages, which are meant to be combined in each session, depending on the goals of the teacher and the needs of students. Activities that include the *Develop and activate background knowledge* stage aim to make children think about what they know about a certain genre (e.g., opinion essay writing). The teacher may induce discussions about the structure of texts, their purpose or audience, among other characteristics. Procedural knowledge is focused on at this stage. At the *Discuss it* stage, children reflect on their own writing skills and their progress in using writing and self-regulation strategies. In the *Model it* stage the teacher shows students how she uses the target writing skills, engaging in a sort of "writing-aloud" process, where she verbalizes her thoughts as she composes text. Often the teacher expresses obstacles and doubts that emerge during writing, showing how the strategies are useful to overcome them. Moreover, she overtly displays positive attitudes towards writing. The *Memorize it* stage consists of the mnemonics designed to facilitate the rote memorization of the strategies. They are introduced relatively early in the intervention and they are supported throughout. The *Support it* stage involves the teacher's helping and guiding students during the intervention, often complemented with materials for charting and monitoring progress, sheets for self-evaluation, among other resources. Finally, the *Independent performance* stage is characterized by a progressive fading out of

support, as children become more apt at using the writing and self-regulation strategies. In sum, SRSD interventions for improving writing are characterized by explicitly teaching children knowledge about writing, including genre-specific elements (e.g., text structure), and providing scaffolded know-how about writing processes (e.g., planning, revising). SRSD instruction is discourse rich and it is paired with the explicit teaching of self-regulation strategies that help children better orchestrate the complex nature of writing processes and subprocesses. Finally, the teacher's role is geared towards encouraging independent behavior, promoting generalization and maintenance, while emphasizing positive attitudes towards writing (e.g., Graham & Harris, 2009; Harris & Graham, 1996; Graham & Harris, 2018).

SRSD has proven to be a useful framework to conduct writing instruction. There is substantial evidence that it works similarly well across discourse genres (e.g., stories, argumentative texts, informational texts) and administration conditions (e.g., single, small group, or whole class, de la Paz & Graham, 2002; Graham et al., 2012; Rogers & Graham, 2008). Children with various learning and language disorders greatly benefit from this type of instruction (Graham & Harris, 2003). SRSD has also been reported to improve the writing skills of disadvantaged children (e.g., Graham & Harris, 2005; Mason, Snyder, Sukhram, & Kedem, 2006). However, to the best of our knowledge, no study to date has compared the efficacy of SRSD writing interventions of disadvantaged versus socioeconomically advantaged peers. This is a crucial goal, given that the improvement of low-SES children needs to be greater in order to avoid Matthew Effects (e.g., Juel, 1988).

Impact of SES on writing development and instruction

The relationship between SES and academic achievement is well documented. SES affects academic performance in general (e.g., Sirin, 2005), and literacy in particular (Coker, 2006; Kim et al., 2015; Korat & Schiff, 2005; Perry & McConney, 2010). It has even been found to be a relevant factor for neurocognitive development, including executive functions (e.g., Hackman & Farah, 2009; Sarsour et al., 2011), which are essential to written composition (e.g., Salas & Silvente, 2019). The gap in academic achievement attributable to SES differences tends to be fairly stable throughout most of the elementary school years, although it increases in adolescence (Sirin, 2005). The disadvantage for low-SES students appears to hold across tasks (e.g., math, literacy) and countries (van Ewijks & Sleegers, 2010), and it has been observed from as early as preschool (Smith & Dixon, 1995).

Given the attested role of SES in the development of writing, it seems logical to consider whether it influences the relative efficacy of writing instruction programs. In this sense, several SRSD interventions have been conducted which targeted children from disadvantaged backgrounds with overall good results (e.g., Coker & Erwin, 2011; Harris, Graham, & Mason, 2006; Harris, Ray, Graham, & Houston, 2019; Graham & Harris, 2005; Mason et al., 2006). However, to the best of our knowledge, no study has performed a systematic comparison of the effectiveness of SRSD writing instruction as a function of children's or school SES. In the present study, we conducted identical versions of an intervention on opinion-essay writing

in "vulnerable" schools (VSs) and in "non-vulnerable" schools (NVSs), where the former included public primary education centers in which a majority of students were considered to be at risk for school failure and social exclusion, whereas the latter included public education centers in which most students belonged to mid- or mid-high SES backgrounds.

Impact of SRSD interventions on related skills

In developmental psychology, complex tasks such as writing, which involves numerous skills and knowledge sources (e.g., Alamargot & Chanquoy, 2001), are arguably good candidates for "developmental cascades". A cascading effect occurs when one or more skills underlying a task have an impact on other, related skills, both within- and across-domains, thus affecting their development. Examination of cascading effects is crucial for expanding our knowledge of the very nature of the complex task at hand, as well as to determine causal (direct or indirect) relations (Fry & Hale, 1996; Masten & Cicchetti, 2010). Although it has not always been entirely possible to tease apart which, exactly, are the features of SRSD that make it so successful (Graham, Harris, & McKeown, 2014, p. 429), it could be argued that training of self-regulatory skills, of genre-specific knowledge, the explicit discussion and training of writing processes, as well as increasing awareness about a writer's goals or intentions (and of readers' reactions or expectations) will have an effect on other writing skills, as well as in other domains (e.g., reading), even if the intervention did not explicitly target them.

A cascading effect underlay studies that examined whether training of transcription skills (i.e., spelling or handwriting) has an impact on written composition scores (Berninger et al., 1997, 2002; Graham, Harris, & Chorzempa, 2002; Graham et al., 2012; Graham & Santangelo, 2014; Limpo, Vigário, Rocha, & Graham, 2020; Santangelo & Graham, 2016). Because poor transcription skills severely constrain the composition process (Berninger, 1999; Berninger & Winn, 2006; Bourdin & Fayol, 1994; Juel, Griffith, & Gough, 1986), training transcription is likely to make it more automatic, with the consequent reduction in cognitive load. More available cognitive resources, in turn, facilitate attention to other writing skills (e.g., text generation, planning). Thanks to a snowball effect, writers will eventually be able to make more sophisticated writing representations and, therefore, produce richer texts that intend to accomplish more challenging rhetorical goals (McCutchen, 2000; Salas & Silvente, 2019).

Developmental cascades in writing could also start at higher levels and affect low-level skills (Olive, 2014). Arguably, an SRSD intervention that targets planning strategies and that explicitly teaches self-regulatory skills should contribute to lessen part of the cognitive effort during composing, as they are useful to help children manage attentional resources. Therefore, the gain in cognitive resources could be devoted (at least in part) to spelling accuracy, especially when considering the well-established finding that spelling accuracy is inversely related to task demands (de Manrique & Signorini, 1998).

Several studies have examined the potential cascading effects of writing to reading comprehension. There is little disagreement that reading and writing, though distinct domains, have a common knowledge base and influence each other throughout development (e.g., Abbott, Berninger, & Fayol, 2010; Fitzgerald & Shanahan, 2000; Tierney & Shanahan, 1991). Accordingly, a meta-analysis on the effects of writing instruction on reading comprehension yielded positive results, although it is still an under-researched area (Graham & Hebert, 2011). The emphasis of an SRSD intervention on genre-specific elements and their hierarchical structure, the explicit discussions about writers' goals (e.g., persuading), and the evaluation of various texts (the child's own production, model texts, as well as texts produced by peers, as was the case in the current intervention) could also be mobilized when reading.

In short, we aimed to examine if cascading effects were observable within the context of an SRSD intervention for planning opinion essays. Specifically, we examined a within-domain effect from writing to spelling skill; and a cross-domain effect, from writing to reading comprehension skill. For this reason, we assessed children's in-text spelling accuracy at both pre- and posttest, as well as reading comprehension tests also at pre- and post-test. Of interest to the current study was whether such effects (if attested) could be observed across grades and for both low- and mid-high-SES populations.

This study

The current study is a large-scale SRSD intervention on opinion-essay writing skills for grade 2 and grade 4 writers, and was part of a larger project on literacy development in Catalan. The intervention, which had a pretest-posttest design with two conditions (SRSD and control), taught children planning strategies, because of the impact of preplanning activity on the overall quality of the text (Berninger, Whitaker, Feng, Swanson, & Abbott, 1996; Harris et al., 2006). It was schools, rather than classrooms or individual students, that were assigned to each condition. This procedure allowed us to offer teacher training to all the teaching staff at participating schools.[1] At least two schools per condition (SRSD and control) and per school status (vulnerable and non-vulnerable) were included, in order to avoid potential confounds (see Table 1 for the overall distribution of participants). Assignment to either condition was quasi-random: schools were assigned to either condition, with only one exception: Teachers who were familiar with the "didactic sequence" method (Camps i Mundó, 1994) were not assigned to control groups, as it is a teaching methodology that shares some characteristics with SRSD.[2] In other words, there were no inclusionary criteria for teachers to be

---

[1] For ethical reasons, the same training was offered to all control schools after the intervention had concluded.

[2] A "didactic sequence" for writing instruction typically consists of a series of tasks that are oriented towards a real-life goal, which is thought to help children understand the function of different text genres and to better take into account readers' needs. It may also provide procedural strategies.

Table 1 Distribution and demographic information of participants

	Control				SRSD			
	Grade 2		Grade 4		Grade 2		Grade 4	
	VSs	NVSs	VSs	NVSs	VSs	NVSs	VSs	NVSs
N (boys–girls)	62 (30–32)	50 (27–23)	68 (33–35)	50 (20–30)	99 (43–56)	125 (56–69)	71 (44–27)	120 (65–55)
M age in months (SD)	89.9 (4.41)	90.16 (3.9)	113.45 (4.84)	112.82 (4.48)	88.51 (4.93)	87.35 (4.26)	112.74 (3.96)	112.24 (4.61)
M ISEI score	44.71 (15.22)	54.80 (13.07)	41.53 (10.88)	52.87 (13.72)	35.93 (8.71)	56.00 (12.64)	35.46 (8.62)	54.24 (13.56)

VS vulnerable schools, *NVSs* non-vulnerable schools, *ISEI* International Socio-economic Index (Ganzeboom et al., 1992)

assigned to either condition, but teachers already familiar with didactic sequences were not assigned to the control condition.

In order to determine the effectiveness of the intervention, we obtained measures of *planning quality, productivity* (number of words and clauses)*, text quality*, and number of genre-appropriate elements or *structural elements*. To investigate developmental cascading effects to spelling and reading comprehension, we obtained measures of *spelling accuracy* and administered *reading comprehension* tests, respectively.

Our research questions were the following:

(RQ1) Is SRSD writing instruction as effective for children in vulnerable schools as it is for children in non-vulnerable schools?
(RQ2) Do children in the SRSD condition make fewer spelling mistakes at posttest, in comparison with children in the control condition?
(RQ3) Do children in the SRSD condition improve their reading comprehension skills at posttest in comparison with the control group?

With regards to RQ1, we hypothesized that low-SES children would benefit from the SRSD intervention, given that previous studies with similar populations have shown improvement in writing using this type of instruction (e.g., Graham et al., 2012). However, we were ambivalent regarding the comparison with their non-vulnerable peers. On the one hand, one could predict that low-SES children would benefit from the intervention to a lesser extent than children who are from a higher SES background. This is because children who grow up in a low-SES environment have had, in general, less exposure to print and print-related activities (Korat, Klein, & Segal-Drori, 2007). Therefore, they could be at a disadvantage in terms of profiting from the SRSD instruction. On the other hand, low-SES children could show a greater benefit, precisely because they have been less exposed to literate activities, so the intervention could make a larger impact. Alternatively, SES status may not mediate the efficacy of the intervention. Previous studies comparing the efficacy of SRSD in different population types (at-risk versus general) did not find evidence of a moderating effect of student status (Graham et al., 2014). With regards to RQ2, we hypothesized that children in the SRSD condition would show significantly fewer spelling mistakes at posttest than their control counterparts. This prediction is based on the arguably lessened cognitive demands that the composing task would pose to children in the SRSD condition, which could be redirected towards spelling accuracy. Finally, regarding RQ3, we expected that the knowledge base shared by writing and reading comprehension, such as the focus of the intervention on genre-specific elements, would lead to a significant improvement in the latter for children in the SRSD condition.

Method

Participants

Six-hundred and forty-five children participated in the study. Children were recruited from 27 s ($n = 336$) and fourth ($n = 309$) grade classrooms in 12 public schools in

prompt was not a factor in the success of the intervention. The same prompts were used in 2nd and 4th grade.

Children were given up to 10 min to plan their text and up to 20 min to compose it. Because several children were likely to be unfamiliar with the term "plan", we explained to them that they were given time to make notes that could help them write the text and compared this action to warming up before running. Testing took place in children's regular classrooms and was conducted by experienced research assistants. No feedback was provided to children on what to say, spelling, or other aspects. Rather, they were encouraged to write as best as they could.

Measures obtained

Planning quality

Children's plans were assessed on a scale from 1 to 5, based on Whitaker, Berninger, Johnston, and Swanson (1994). Plans that were left blank, contained only the prompt, or consisted of a single sentence, were afforded a score of 1. The same score was assigned to plans that were virtually identical to the final text, suggesting that the child had composed the text during planning time and then copied from it. Samples with "minimal planning" received a score of 2. Minimal planning involved some kind of rearrangement of (almost) identical sentences, as well as cases in which the child copied up to a point from the plan, but then went on and added at least a few more sentences. Plans that included instructions on how to build the text (e.g., "say I agree" or "say why I think so") received the same score of 2. Plans that included lists of keywords or topics that were to be developed in the text received a score of 3. Plans that showed some emergent structure or hierarchy, such as an abbreviated version of a topic indicating the part of the text to which they belong (e.g., Introduction) received a score of 4 or 5. A score of 4 was assigned to plans in which only one of the structural elements was developed, while a score of 5 was assigned to plans that developed two or more of the topics and subtopics. A trained research assistant scored all plans and the second author scored a random selection of 20%. The Kappa coefficient was 0.71, indicating that there was substantial reliability between scorers (Landis & Koch, 1977).

Number of words

All texts were transcribed in a text editor and then transformed into CLAN-compatible files using the TEXTIN command (MacWhinney, 2017). Two research assistants transcribed all texts, correcting spelling and punctuation mistakes to avoid bias from evaluators of text quality and other measures. Transcription reliability was determined on the basis of 100 transcriptions, ICC = .996 (95% C.I.: .994–.997). Words were counted automatically in CLAN using the STATFREQ command.

Fig. 1 Screenshots of the *TextHandler* app, used to analyze the quality (top) and the structure (bottom) of the opinion essays randomly. The text is shown on the left hand side, with red letters displaying the participant's code and blue letters displaying the prompt. Raters completed the form on the right hand side

Number of clauses

Texts were segmented into clauses during transcription (after Berman and Slobin, 1994). Reliability across transcribers was ascertained on the basis of 100 texts, ICC = .979 (95% C.I.: .969–.985). Clauses were counted automatically using the STAT-FREQ command in CLAN (MacWhinney, 2017).

Text quality

Texts were uploaded to *TextHandler* (https://text-handler-production.surge.sh/), a bespoke app to facilitate the evaluation of texts. Raters log in the app and choose the genre (argumentative) and the educational level (2nd or 4th) of the writers and the app displays a text that complies with these conditions (Fig. 1). The app is built so that the search of texts to be evaluated is otherwise completely random and unknown to the rater (i.e., no classroom, condition, or time-point information is displayed).

Texts were placed on a scale from 1 (poor quality; well below grade level) through to 6 (excellent quality; well above grade level).

Two undergraduate students scored all 2nd and 4th grade text samples. Raters were unfamiliar with the purpose of the study and were told to factor in the content and organization of the texts, their communicative effectiveness, as well as vocabulary richness, and grammatical accuracy. Raters were initially trained with a sample of 40 texts from their corresponding educational level, which had been scored by the research team, in consultation with highly experienced teachers. These texts were thus used as a practice-with-feedback training phase and anchor texts for quality levels 2, 4, and 6 were identified and used during subsequent scoring (see, for example, Verheyden, Van den Branden, Rijlaarsdam, Van den Bergh, & De Maeyer, 2010, for a similar procedure). The second author rescored all texts, and large discrepancies were rescored by a third rater. Inter-rater reliability was good, ICC = .84 and .80, for grades 2 and 4, respectively. The final score was the average between the two (or three) scores.

Structural elements

The TextHandler app was also used for assessing the presence of key genre-appropriate elements in the opinion essays. Raters were required to determine whether the text included the following elements: thesis (1 point), reasons (1 point per reason), explanations (1 point per explanation), and conclusion (1 point). The app displayed a random text, without providing information of the time point, condition, or the writer's grade (Fig. 1). Raters were two undergraduate students of Primary Education and Speech Therapy. They received extensive training (approximately, 6 h) on detecting the various structural elements before they rated texts in the app and were given a manual detailing the criteria. A random sample of 20% of the texts were scored by both. The final score was the sum of all the structural elements identified in the text. Inter-rater reliability was excellent, ICC = .926.

Spelling accuracy

Spelling mistakes were identified in the texts and then counted automatically using the STATFREQ function in CLAN (MacWhinney, 2017). If the same word contained more than one spelling mistake, it was nevertheless counted as one error, while errors that were committed more than once were counted every time. The score was the proportion of spelling mistakes over the total number of words. Inter-rater reliability was ascertained on the basis of 10 texts, ICC = .922 (95% C.I. = .883–.947).

Reading Comprehension

At pretest children were administered an adaptation of the *Avaluació de la Comprensió Lectora* (ACL) test, which includes standardized tests of reading comprehension for 1st through to 6th grade (Catalá, Catala, Molina, & Monclús, 1998). This means

that there are different instruments depending on grade level. The test is a collection of short texts with questions that target comprehension of literal information or inference making across various genres (e.g., narrative, expository, argumentative, informative, etc.). As per the test instructions, we administered the grade 1 instrument to our 2nd graders, and the grade 3 instrument to our 4th graders. We included two additional texts from the next grade level, in order to avoid a potential negative skew. Cronbach's alpha for the pretest instrument was .939, for the 2nd graders, and .972, for the 4th graders. The score was the percentage of correct responses.

At posttest we administered the reading comprehension subtest in PROLEC-R, a test battery to assess reading skills in Catalan (Cuetos, Rodríguez, Ruano, & Arribas, 2007). The test is designed for individual administration, so we adapted it to a multiple-choice format. The 2nd grade instrument included two texts, one expository and one narrative, and the 4th grade instrument included two texts of each genre. In addition, we added an opinion essay, also with multiple-choice questions, to the 2nd and the 4th grade version. Therefore, both the format of the posttest instruments (short texts of different genres, with multiple choice questions) and the skills they evaluated (e.g., literal and non-literal information) were the same as in the pretest instruments. To avoid bias, we created two different orders of presentation and counterbalanced them. The test showed moderate reliability, with Cronbach's alpha = .71, for the grade 2 instrument (12 items), and .66, for the 4th grade instrument (20 items). The score was the percentage of correct responses.

The intervention

We developed two adaptations of Limpo and Alves' (2013) writing intervention for 5th and 6th grade students, for our grade 2 and grade 4 participants. The overall goal of the interventions was to teach children to plan opinion essays, while developing self-regulation strategies. Both interventions consisted of 11 sessions of roughly 60 min each, that were to be held twice a week. The interventions started in March–April and finished in May. Pretest took place one to 2 weeks before the start of the intervention, and posttest, 1 week after the intervention had concluded. All intervention sessions (but not the testing sessions) were conducted by the children's regular teachers.

In session 1, teachers told students that they would participate in a project where children share what they think with other children, teachers, and the general public. Their written texts were going to be uploaded to a blog that we created specifically for the project. This procedure aimed to provide a real-life goal to motivate children to write their essays as best as they could (Graham & Perin, 2007). This session was also used to discuss what children knew about expressing opinions and key terms were introduced: planning, strategy, goal. The session concluded with a general understanding that goals are best achieved by planning in advance and making an effort. The planning strategy was introduced in Session 2. The grade 2 strategy was *PER*, the Catalan acronym for *Penso* '(I) think' (an idea), *3 Explicacions* '3 explanations', and *Resum* 'Summarize'. The grade 4 strategy was *CREC*, where C stands for *Crec* '(I) believe', R stands for *3 Raons* '3 reasons', E stands for *3 Explicacions* '3

explanations', and C stands for *Conclusió* 'conclusion'. Goal-setting was also introduced in this session. Session 3 introduced self-evaluation and progress monitoring. Self-evaluation (and a corresponding self-evaluation sheet) was first practiced on a model text about the same topic as the students' pretest. After evaluating the model text, children evaluated their own pretest texts. As most children got very low scores, the class was led to conclude that the PER/CREC strategy was needed to improve their writing. Progress-monitoring was introduced in this session. In session 4 the teacher modeled the use of the strategy for planning. The teacher used a graphic organizer, and afterwards the whole class discussed the self-instructions she used during planning. Children planned texts as a class in session 5, and individually in session 6. Session 7 involved the teacher's modelling of the entire writing process, from planning to completion of the composition of the text; this time without the help of the graphic organizer. In session 8, children worked in pairs to write a text using the strategies, and in session 9 they worked individually. Children wrote another text individually for homework between sessions 8 and 9. Teachers were instructed to evaluate the homework text, prioritizing some difficulties over others: Difficulties with planning (e.g., not using the strategy, leaving plans incomplete, etc.) were to be addressed first; difficulties with the transition from the plan to the text (e.g., disregarding the plan once complete), should be next in importance; finally, difficulties with the text itself (e.g., not developing ideas further in the text and simply copying from the plan) should come third. In session 10, based on the result of the teacher's assessment, children were organized in groups according to the main difficulty they had shown. Teachers discussed with each group what their difficulty was and helped develop an additional self-instruction that targeted it. Session 11 was devoted to individual writing.

Teachers were encouraged to keep graphic organizers and other supporting material after session 7 as and when students needed them, so as to tailor the intervention to children's needs as much as possible. They were also told to follow the same feedback criteria as in session 10 throughout the intervention. Furthermore, some sessions included activities to (1) promote the use of the strategy in other contexts; (2) discuss the quality of reasons and explanations; and (3) create situations for children to reflect on their own progress writing opinion essays.

Control group

Teachers in the control condition were invited to keep a log book of all literacy-related activities that they carried out while the interventions in the SRSD condition were taking place. Log books were then checked by the first author, confirming that the nature of the literacy activities in these groups did not share the key features of the intervention. Children in the control condition produced the same number of texts as the children in the intervention condition, using the same prompts, in order to provide children across conditions the same amount of writing practice.

Table 2 Intra-class correlation coefficients (ICCs) for all outcome variables

Outcome	ICC
Plans	0.607
Words	0.301
Clauses	0.211
Structure	0.367
TQ	0.132
Spelling	0.090
ReadC	0.090

Plans = planning score (range: 1–5); TQ = text quality score (range 1–6); Spelling = text spelling accuracy (proportion over number of words); ReadC = reading comprehension percentage correct

Treatment fidelity

Treatment fidelity was ensured via a series of actions. First, teachers in the SRSD condition attended an 8-hour training course, divided into four sessions imparted at the schools by the first author. Training aimed to (1) present the goal of the intervention and its theoretical and practical bases; (2) introduce the teachers' book, which contained detailed information of each session, model scripts for the modelling sessions, and answer keys to all activities; and (3) discuss teachers' concerns. Second, teachers were provided checklists with 10–12 points per session, which targeted the key aspects of the intervention (e.g., *explain the planning strategy going over each letter in the acronym*; or *discuss the self-instructions used during modeling*). Third, the first author conducted a close follow-up before and after each session, in which she contacted teachers and discussed any issues with the prior session, and then reviewed goals and concerns for the following session. Checklists were used as a point of reference for these follow-up calls; that is, researcher and teacher would review that each point had been accomplished in the previous session, and discussed key aspects and potential issues with the following session. With this procedure, teachers reported having completed 100% of the key points for each session. Fourth, teachers were required to audio-record at least 30% of the sessions, preferably from sessions 1, 2, 3, 4, and 7, in which they played a larger role. All recordings were then listened to by a trained research assistant who also completed the checklists. Coverage of sessions ranged between 96 and 100%, indicating that the teachers had followed the activities in the intervention closely.

Data analysis

We aimed to explain the difference between posttest and pretest scores for each outcome variable; that is, the gain at posttest, relative to pretest scores. Because children were nested into classrooms, we computed intra-class correlation coefficients (ICCs), which indicated that a moderate to large proportion of the variance in each
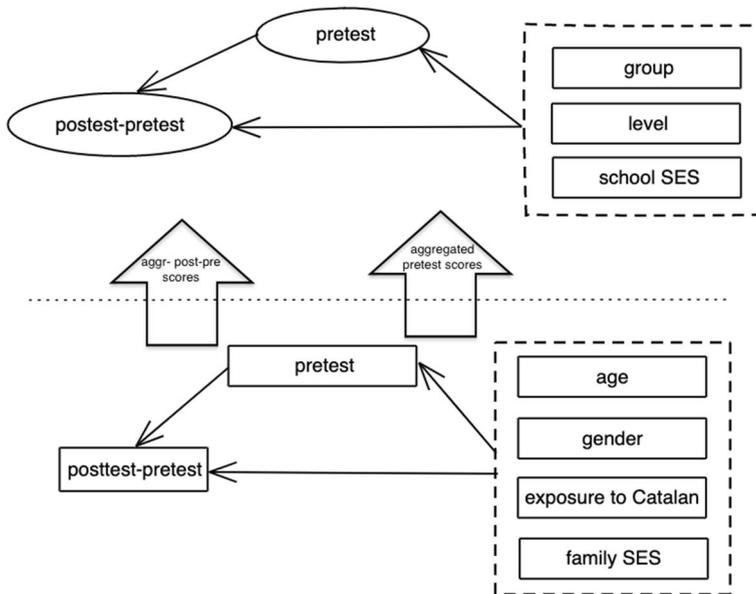
Fig. 2 Multi-level path-analysis model. Pretest and change (posttest-pretest) outcomes were assessed as latent random factors. Level 1 variables are displayed below the dotted line, while Level 2 variables are displayed above. Wide arrows indicate cross-level assessment

outcome variable was due to children being nested in classrooms (Table 2, Heck & Thomas, 2015).

The efficacy of the intervention was then tested via a series of multilevel path analysis models in MPlus 8.2 (Muthén & Muthén, 2017), in which pre-intervention and change in performance were set as simultaneous outcomes, with the pre-intervention regressed on the change, and personal- and class-level independent characteristics were regressed on the two outcomes. Figure 2 shows that the two outcomes were assessed as latent random factors, that is, randomly varying across classes. This process is illustrated in (1) the wide arrows that show the cross-level assessment, and (2) the ellipses for latent variables. All explanatory variables were observed regardless of their level affiliation. Correlations across explanatory variables were further allowed to a minimum extent for meeting model fit requirements.

Results

Descriptive statistics are presented in Table 3. Data were normally distributed, with skewness and kurtosis values within acceptable levels: skewness values ranged between -0.938 and 1.994, while kurtosis values ranged between -1.85 and 7.016 (Kline, 2011). Some of the children were absent during the pre- or the posttest sessions. Missing values amounted to 8.5% ($n = 55$) and 7.9% ($n = 51$) at pretest and posttest, respectively. The proportion of missing data did not differ between the

Table 3 Descriptive statistics (means and SDs) by condition, grade level, and school status for pretest and posttest

	SRSD				Control			
	Grade 2		Grade 4		Grade 2		Grade 4	
	VS	NVS	VS	NVS	VS	NVS	VS	NVS
	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)
Plans								
Pretest	1.32 (0.49)	1.33 (0.47)	1.34 (0.54)	1.87 (1.00)	1.24 (0.43)	1.33 (0.52)	1.44 (0.72)	1.39 (0.68)
Posttest	3.45 (1.88)	4.47 (1.21)	4.65 (0.99)	4.09 (1.49)	1.16 (0.46)	1.39 (0.77)	1.23 (0.66)	1.98 (1.08)
Words								
Pretest	22.55 (17.46)	26.05 (20.31)	34.81 (15.69)	50.50 (26.43)	21.46 (11.42)	21.81 (13.81)	50.15 (33.29)	47.98 (25.88)
Posttest	35.69 (17.80)	44.65 (18.02)	78.97 (24.90)	80.95 (23.80)	19.41 (10.06)	22.89 (14.05)	45.74 (27.40)	55.25 (24.31)
Clauses								
Pretest	5.21 (3.45)	6.74 (4.28)	7.94 (3.91)	11.68 (5.77)	4.35 (2.24)	5.69 (2.96)	9.93 (5.65)	11.22 (6.40)
Posttest	9.09 (4.52)	10.66 (3.96)	17.88 (5.79)	17.58 (5.18)	5.25 (2.80)	5.83 (3.27)	11.10 (6.77)	12.48 (5.77)
Structure								
Pretest	1.58 (1.11)	1.72 (0.96)	1.89 (0.99)	2.49 (1.26)	1.35 (0.99)	1.54 (1.03)	2.15 (1.18)	2.49 (1.06)
Posttest	2.53 (1.46)	3.39 (1.37)	4.04 (1.18)	4.53 (1.00)	1.54 (0.81)	1.40 (0.74)	2.11 (0.95)	2.52 (1.15)
TQ								
Pretest	1.47 (0.68)	1.69 (0.83)	1.38 (0.62)	1.77 (0.66)	1.28 (0.77)	1.74 (0.67)	1.52 (0.85)	1.73 (0.80)
Posttest	2.04 (1.46)	2.90 (1.69)	2.63 (1.25)	2.47 (1.69)	1.49 (0.71)	1.74 (0.72)	1.66 (0.74)	1.86 (0.62)
Spelling								
Pretest	0.40 (0.16)	0.35 (0.17)	0.25 (0.11)	0.18 (0.09)	0.26 (0.19)	0.39 (0.19)	0.20 (0.10)	0.25 (0.13)
Posttest	0.40 (0.17)	0.32 (0.15)	0.24 (0.08)	0.19 (0.09)	0.33 (0.16)	0.38 (0.16)	0.23 (0.08)	0.24 (0.13)
ReadC								
Pretest	0.49 (0.19)	0.60 (0.23)	0.48 (0.22)	0.68 (0.19)	0.55 (0.20)	0.65 (0.22)	0.54 (0.23)	0.68 (0.21)

Table 3 (continued)

	SRSD				Control			
	Grade 2		Grade 4		Grade 2		Grade 4	
	VS	NVS	VS	NVS	VS	NVS	VS	NVS
	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)
Posttest	0.62 (0.21)	0.73 (0.19)	0.70 (0.16)	0.77 (0.12)	0.67 (0.22)	0.81 (0.20)	0.70 (0.15)	0.72 (0.12)

Plans = planning score (range: 1–5); TQ = text quality score (range 1–6); Spelling = text spelling accuracy (proportion over number of words); ReadC = reading comprehension percentage correct; *VS* vulnerable schools, *NVS* non-vulnerable schools

control and SRSD conditions, either at pretest, $Chi^2(1) = 1.67$, $p = .197$, or at post-test, $Chi^2(1) = 1.71$, $p = .192$. Missing data were replaced with correlated data using the Full Information Maximum Likelihood (FIML) framework.

Effectiveness of the SRSD interventions in vulnerable and non-vulnerable classrooms

The models were set up so that, at level 1, the level of individual children ($N = 645$), we controlled for children's age, gender, SES, and exposure to Catalan outside school. At level 2, the level of the classrooms ($N = 27$), we tested the effect of children's Condition (control v. SRSD), Grade (2nd v. 4th), classroom SES (vulnerable v. non-vulnerable), which was inherited from the school's SES status. At both levels we controlled for pretest scores. The models computed for each outcome variable showed beyond acceptable fit indices, $Chi^2(7) > 4.43$, $p > .05$, RMSEA = 0.000, CFI = 1.000, TLI > 1.003, $SRMR_{within} < 0.026$, $SRMR_{between} < 0.057$. Tables 4 and 5 display the SEM results at Level 1 and 2, respectively. All models resulted in substantial explained variance, especially at Level 2 (see R^2 values in Table 5). The models also tested for the effect of pretest scores (e.g., number of words at pretest) at both levels (Fig. 2). Finally, Table 6 displays effect sizes (standardized mean difference, using Hedge's g correction; Borenstein, Cooper, Hedges, & Valentine, 2009) for the main effect of Condition across grades and population types.

Planning quality

At Level 1 the gain at posttest in planning quality was explained by children's gender, where girls significantly outperformed boys, b[5] $= 0.31$, $p = 002$. In addition, higher pretest scores were associated with smaller gains at posttest, $b = -0.77$, $p < .001$. At Level 2 there was a significant Condition effect, $b = 2.72$, $p < .001$, whereby children assigned to the SRSD condition wrote more sophisticated plans at posttest than children in the control condition. There was also a significant Grade × School SES interaction, $b = 1.01$, $p = 046$; however, the decomposition of this interaction did not show any significant simple slope.

Number of words

At Level 1, children's gender influenced the number of words that children wrote at posttest, relative to pretest, with girls writing significantly more words, on average, than boys, $b = 6.17$, $p < .001$. Higher numbers of words at pretest was associated with a smaller gain at posttest; that is, a regression to the mean effect, $b = -0.72$, $p < .001$. At level 2 children in the SRSD condition made larger gains in words in comparison to controls, $b = 24.08$, $p < .001$, with 4th graders writing significantly more words than second graders, $b = 18.17$, $p = .035$. There was a

---

[5] All estimates are unstandardized.

Table 4 Level 1 SEM results for all outcome variables (unstandardized coefficients)

	Plans Est. (SE)	Words Est. (SE)	Clauses Est. (SE)	Structure Est. (SE)	TQ Est. (SE)	Spelling Est. (SE)	ReadC Est. (SE)
Age	0.01 (0.01)	0.13 (0.16)	n/a[a]	n/a[a]	< 0.01 (0.01)	<0.01 (<0.01)	n/a[a]
Gender	0.31* (0.01)	6.17* (1.56)	1.15* (0.38)	0.29* (0.09)	−0.01 (0.10)	−0.01 (0.01)	0.02 (0.01)
SES	< 0.01 (< 0.01)	−0.01 (0.07)	0.01 (0.02)	0.01 (<0.01)	<0.01 (0.01)	<0.01* (<0.01)	<0.01 (< 0.01)
CatExp	−0.15 (0.12)	0.09 (1.97)	0.04 (0.50)	−0.26* (0.13)	0.06 (0.14)	0.01 (0.01)	−0.02 (0.02)
Pretest	−0.77* (0.08)	−0.72* (0.04)	−0.71* (0.04)	−0.73* (0.04)	−0.59* (0.07)	−0.70* (0.04)	−0.60* (0.03)
R^2	0.19* (0.04)	0.44* (0.04)	0.34* (0.03)	0.37* (0.09)	0.11* (0.03)	0.41* (0.03)	0.44* (0.03)

Plans = planning score (range: 1–5); TQ = text quality score (range 1–6); Spelling = text spelling accuracy (proportion over number of words); ReadC = reading comprehension percentage correct; SES = ISEI score, n/a = not applicable

[a] Age did not contribute to explaining any significant variance in the model and considerably worsened fit. For this reason, it was dropped from the analysis

*$p < .05$

Table 6 Effect sizes (and SE) for measures with a significant main effect of condition

Measure	Overall	Grade 2	Grade 4	VSs	NVSs
Planning	2.62 (0.11)	2.75 (0.16)	2.49 (0.15)	2.88 (0.17)	2.32 (0.15)
No. words	1.18 (0.09)	1.00 (0.12)	1.37 (0.13)	1.40 (0.13)	0.93 (0.12)
No. clauses	0.98 (0.09)	1.12 (0.12)	1.10 (0.13)	1.16 (0.13)	0.87 (0.12)
Text structure	1.49 (1.00)	1.15 (0.12)	1.89 (0.14)	1.23 (0.13)	1.70 (0.14)
Text quality	0.77 (0.21)	0.69 (0.33)	0.86 (0.37)	0.79 (0.44)	0.76 (0.26)

VSs vulnerable schools, *NVSs* non-vulnerable schools

significant Condition x Grade interaction, $b = 13.24$, $p = .039$, due to a varying Grade effect: 4th graders made larger gains at posttest than 2nd graders only in the SRSD condition.

Number of clauses

At level 1 children's gains at posttest were larger for girls than for boys, $b = 1.15$, $p = .002$, while higher pretest scores were associated with smaller gains, $b = -0.71$, $p < .001$. At level 2 the gain at posttest in number of clauses was significantly higher for children in the SRSD condition, than for controls, $b = 4.78$, $p < .001$.

Structural elements

The gain in the number of genre-appropriate structural elements that children included at posttest was explained by children's gender, where girls outperformed boys, $b = 0.29$, $p = .002$. Moreover, having no exposure to Catalan outside school was associated with a larger gain, $b = -0.26$, $p = .038$, as well as a lower pretest score, $b = -0.73$, $p < .001$. At level 2, gains at posttest were significantly larger for children in the SRSD condition than for controls, $b = 1.90$, $p < .001$. Also, 4th graders made larger gains than 2nd graders, $b = 2.35$, $p < .001$, and children in NVSs showed a larger improvement than those in VSs, $b = -0.79$, $p = .041$.

Text quality

At level 1, gains in text quality were only explained by pretest scores, where higher scores were associated with smaller gains, $b = -0.59$, $p < .001$. At level 2, there was a significant main effect of Condition, $b = 0.84$, $p < .001$, indicating that children in the SRSD classrooms made larger gains than controls. Interestingly, there was a significant 3-way interaction, $b = 0.70$, $p = 0.20$, which revealed that the effect of Condition was only significant for children in VSs (both grades), but not for children in either grade in NVs.

Cascading effects to spelling accuracy

There was no indication that the intervention impacted on children's spelling scores at posttest, relative to pretest. At level 1, gains in spelling were associated with children's SES, $b = -0.002$, $p < .001$, and pretest scores, $b = -0.70$, $p < .001$, where higher SES levels were associated with fewer spelling mistakes, and larger pretest scores (i.e., a higher proportion of mistakes) were associated with a higher decrease of mistakes at posttest. At level 2, pretest scores were significantly, and indirectly associated with gains, $b = -0.35$, $p = .018$. We also attested a significant 3-way interaction, $b = -0.05$, $p = 0.42$, indicating that 4th graders decreased the average proportion of spelling mistakes at posttest in comparison with 2nd graders, but only in the SRSD condition in VSs.

Cascading effect to reading comprehension

At Level 1 there was a regression to the mean effect, $b = -0.60$, $p < .001$, meaning that larger pretest scores were associated with smaller gains at posttest. At level 2 there was a significant 3-way interaction, $b = 0.36$, $p = .016$, that indicated that (1) the effect of Grade (4th > 2nd) was only attested for the SRSD condition in VSs; (2) the effect of School SES was significant only for controls in 4th grade, by which NVSs outperformed VSs; and (3) that the effect of Condition was only significant for children in NVSs, in 4th grade.

Discussion

We addressed the effectiveness of SRSD interventions to improve opinion-essay writing skills in low-SES children, relative to peers in non-vulnerable schools. We also examined the potential of SRSD interventions to trigger cascading effects to two related skills: spelling and reading comprehension.

Effectiveness of SRSD for vulnerable and non-vulnerable populations

SRSD was an effective method to improve opinion-essay writing skills for children in both vulnerable and non-vulnerable schools. Children in the SRSD condition made better pre-writing plans and then wrote texts that were longer and included more genre-appropriate structural elements at posttest, than controls. This finding provides further support to reports that SRSD is generally effective in low-SES settings (e.g., Harris et al., 2019; Mason et al., 2006), by offering a systematic comparison of identical implementations in two different contexts.

A few nuances need to be noted, however. For most outcomes, there were no differences as a function of children's socioeconomic context at pretest, either. This might point to children's poor writing skills in non-narrative genres, regardless of SES (Berman, 2008). Indeed, mastery of argumentative writing is a challenging task

that requires intensive, continued instruction (Olive, Favart, Beauvais, & Beauvais, 2009). In contrast, for some writing outcomes, children benefited from the intervention to different degrees as a function of SES. This was the case of structural elements, where children in mid-high SES contexts made larger gains than their peers at posttest, a difference that was also attested at pretest. More privileged children, who are typically raised by adults with higher levels of education, are likely to have more exposure to uses of language (written and spoken) that resemble academic or school-like usage, such as the language required to articulate arguments (e.g., Kesane & Ruiz, 2010). In contrast, it was children in vulnerable schools who made larger gains in text quality. That more privileged children did not improve their text quality was surprising, especially in view that every other measure showed a clear improvement. A possible explanation is that children in vulnerable schools showed a lower starting point, and the intervention allowed them to catch up with their peers, in line with the regression to the mean effects throughout the study. Moreover, while our assessment of text quality included attention to argument elaboration, it may not be sensible enough for differences in the quality of the arguments developed by the children. Future studies should include more fine-grained measures of argument structure and quality to better understand what changes and how after an SRSD intervention.

The interventions worked similarly well for beginner- and intermediate-level writers; that is, for children in 2nd and 4th grade, in line with previous meta-analytic reviews that found that the efficacy of SRSD was not mediated by age (e.g., Graham et al., 2012, p. 886). This means that teachers should provide writing instruction in genres other than narrative to young children, given that it is an essential skill for academic attainments, and that earlier intervention may be essential to avoiding Matthew effects in writing (Juel, 1988).

Developmental cascades and SRSD

There was no evidence that our SRSD intervention targeting planning and self-regulation activated developmental cascades that had a positive impact on children's spelling skills. While SRSD arguably helps reduce cognitive demands by providing writing schemas and a set of self-regulation skills to better orchestrate the writing process, children might devote the newly available cognitive resources to other goals. In addition, some children might simply not count with the necessary orthographic representations to improve the spelling accuracy of their texts. Such a result strengthens the view that spelling requires explicit teaching both in isolation and in combination with the teaching of higher-level skills (Alves, Limpo, Salas, & Joshi, 2018; Berninger et al., 2002).

In contrast, there was some evidence that reading comprehension skills may have been affected by positive cascades stemming from the skills targeted by the intervention. However, only a subgroup of children (mid-high-SES 4th graders) improved their reading comprehension skills as a function of the intervention. This should not be taken to mean that explicit instruction of reading comprehension is unnecessary; rather, it shows that the intervention may have been effective in activating children's

knowledge and familiarity with rhetorical and structural aspects of written language, facilitating reading comprehension. The fact that only older children in more advantaged contexts showed this added benefit may be explained by their longer and richer exposure to written language, both inside and outside school. Although it is hard to pinpoint which aspects of the intervention might have triggered this effect, we would argue that the discussions around a writer's goals and the hierarchical organization of information in the text are likely contenders. In this sense, children from disadvantaged contexts may require more extended contact with and instruction on written language for the same mechanisms to be activated (Korat et al., 2007).

Limitations

A number of limitations needs to be noted. Firstly, conducting research in Catalan, a minority language, involved choosing from the only two reading comprehension standardized tests available and making adaptations. This is far from ideal, which we tried to compensate by rigorously ensuring the comparability and reliability of all instruments. However, it was not a perfect comparison and it is plausible that the cascading effect from SRSD to reading comprehension would have been larger had we administered tests that consisted only of opinion-essay items. Secondly, children were all speakers of both Catalan and Spanish and, in a number of cases, they also spoke additional languages (very often in the case of the children in vulnerable schools, who are usually 1st or 2nd generation immigrants). Therefore, the present findings may be harder to generalize to other contexts. Finally, although teachers were instructed to provide support to students according to their needs (i.e., a criterion-based approach), due to implementation limitations all children received the same number of sessions and were evaluated simultaneously, regardless of their mastery level (a time-based approach, e.g., Harris & Graham, 2017, p. 125). It is likely that this design feature affected at least some of the findings in this study and should be borne in mind during interpretation.

Educational implications and concluding remarks

This study provided evidence that SRSD is a valid, effective approach to writing instruction for disadvantaged children. Teaching children about genre elements and writing processes, while providing them with self-regulation strategies, is an efficient way to improve writing skills in beginner and intermediate-level children, regardless of SES background. We have also shown that SRSD may trigger positive developmental cascades to other skills, although more research is needed to understand which specific features of SRSD are responsible for such cascading effects, as well as their overall scope and the conditions under which they emerge. Finally, although the SRSD intervention was generally effective, some of its benefits were not within reach of all students. Most importantly, both vulnerable and non-vulnerable children appeared to benefit from the instruction to a very similar extent for most outcome measures, which is insufficient to bridge the gap between these populations. Educators and policy makers should substantially increase underprivileged

children's opportunities to become in contact with written language, especially non-narrative genres, which are particularly instrumental to academic success. We would argue that these efforts may need to be conducted as early as possible (i.e., preschool), given that SES-related effects appear to be operative already in the earlier elementary school years.

References

Abbott, R. D., Berninger, V. W., & Fayol, M. (2010). Longitudinal relationships of levels of language in writing and between writing and reading in grades 1 to 7. *Journal of Educational Psychology, 102*(2), 281–298. https://doi.org/10.1037/a0019318.

Alamargot, D., & Chanquoy, L. (2001). *Through the models of writing*. Berlin: Springer.

Alves, R. A., Limpo, T., Salas, N., & Joshi, R. M. (2018). Handwriting and spelling. In S. Graham, C. A. MacArthur, & M. Hebert (Eds.), *Best practices in writing instruction* (3rd ed.). New York: Guildford Press.

Berman, R. A. (2008). The psycholinguistics of developing text construction. *Journal of Child Language, 35*(4), 735–771. https://doi.org/10.1017/S0305000908008787.

Berman, R. A., & Slobin, D. I. (1994). *Relating events in narrative: A cross linguistic developmental study*. Hillsdale: Laurence Erlbaum.

Berninger, V. W. (1999). Coordinating transcription and text generation in working memory during composing: Automatic and constructive processes. *Learning Disability Quarterly, 22*(2), 99–112. https://doi.org/10.2307/1511269.

Berninger, V. W., Vaughan, K. B., Abbott, R. D., Abbott, S. P., Rogan, L. W., Brooks, A., et al. (1997). Treatment of handwriting problems in beginning writers: Transfer from handwriting to composition. *Journal of Educational Psychology, 89*(4), 652. https://doi.org/10.1037/0022-0663.89.4.652.

Berninger, V. W., Vaughan, K., Abbott, R. D., Begay, K., Coleman, K. B., Curtin, G., et al. (2002). Teaching spelling and composition alone and together: Implications for the simple view of writing. *Journal of Educational Psychology, 94*(2), 291. https://doi.org/10.1037/0022-0663.94.2.291.

Berninger, V., Whitaker, D., Feng, Y., Swanson, H. L., & Abbott, R. D. (1996). Assessment of planning, translating, and revising in junior high writers. *Journal of School Psychology, 34*(1), 23–52. https://doi.org/10.1016/0022-4405(95)00024-0.

Berninger, V. W., & Winn, W. (2006). Implications of advancements in brain research and technology for writing development, writing instruction, and educational evolution. In C. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (pp. 96–114). New York: Guilford Press.

Borenstein, M., Cooper, H., Hedges, L., & Valentine, J. (2009). Effect sizes for continuous data. In H. Cooper, L. V. Hedges, & J. C. Valentine (Eds.), *The handbook of research synthesis and meta-analysis* (2n ed., pp. 221–235). New York City: Russell Sage Foundation.

Bourdin, B., & Fayol, M. (1994). Is written language production more difficult than oral language production? A working memory approach. *International Journal of Psychology, 29*(5), 591–620. https://doi.org/10.1080/00207599408248175.

Camps i Mundó, A. C. (1994). *L'ensenyament de la composició escrita. [The teaching of written composition]*. Barcelona: Barcanova.

Catalá, G., Catala, M., Molina, E., & Monclús, R. (1998). *Avaluació de la comprensió lectora. Proves ACL [Assessment of Reading comprehension]*. Barcelona: Graó.

Coker, D. (2006). Impact of first-grade factors on the growth and outcomes of urban schoolchildren's primary-grade writing. *Journal of Educational Psychology, 98*(3), 471. https://doi.org/10.1177/00420 85910377426.

Coker, D. L., Jr., & Erwin, E. (2011). Teaching academic argument in an urban middle school: A case study of two approaches. *Urban Education, 46*(2), 120–140. https://doi.org/10.1177/0042085910 377426.

Cuetos, F., Rodríguez, B., Ruano, E., & Arribas, D. (2007). Prolec-r. *Evaluación de los procesos lectores–revisado [Assessment of reading processes-revised].* TEA.

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. https://doi.org/10.1037/0022-0663.94.4.687.

de Manrique, A. M. B., & Signorini, A. (1998). Emergent writing forms in Spanish. *Reading and Writing, 10*(6), 499–517. https://doi.org/10.1023/A:1008019206946.

Fitzgerald, J., & Shanahan, T. (2000). Reading and writing relations and their development. *Educational Psychologist, 35*(1), 39–50. https://doi.org/10.1207/S15326985EP3501_5.

Fry, A. F., & Hale, S. (1996). Processing speed, working memory, and fluid intelligence: Evidence for a developmental cascade. *Psychological Science, 7*(4), 237–241. https://doi.org/10.1111/j.1467-9280.1996.tb00366.x.

Ganzeboom, H. B., De Graaf, P. M., & Treiman, D. J. (1992). A standard international socio-economic index of occupational status. *Social Science Research, 21*(1), 1–56. https://doi.org/10.1016/0049-089X(92)90017-B.

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

Graham, S., & Harris, K. R. (2005). Improving the writing performance of young struggling writers: Theoretical and programmatic research from the center on accelerating student learning. *The journal of Special Education*, *39*(1), 19–33.https://doi.org/10.1177/00224669050390010301.

Graham, S., & Harris, K. R. (2009). Almost 30 years of writing research: Making sense of it all with The Wrath of Khan. *Learning Disabilities Research and Practice, 24*(2), 58–68. https://doi.org/10.1111/j.1540-5826.2009.01277.x.

Graham, S., & Harris, K. (2018). An examination of the design principles underlying a self-regulated strategy development study. *Journal of Writing Research, 10*(2), 139–187. https://doi.org/10.17239/jowr-2018.10.02.02.

Graham, S., Harris, K. R., & Chorzempa, B. F. (2002). Contribution of spelling instruction to the spelling, writing, and reading of poor spellers. *Journal of Educational Psychology, 94*(4), 669. https://doi.org/10.1037/0022-0663.94.4.669.

Graham, S., Harris, K. R., & McKeown, D. (2014). The writing of students with learning disabilities, meta-analysis of self-regulated strategy development writing intervention studies, and future directions: Redux. In H. L. Swanson, K. R. Harris, & S. Graham (Eds.), *Handbook of learning disabilities* (pp. 405–438). New York: The Guilford Press.

Graham, S., & Hebert, M. (2011). Writing to read: A meta-analysis of the impact of writing and writing instruction on reading. *Harvard Educational Review, 81*(4), 710–744. https://doi.org/10.17763/haer.81.4.t2k0m13756113566.

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. https://doi.org/10.1037/a0029185.

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

Graham, S., & Santangelo, T. (2014). Does spelling instruction make students better spellers, readers, and writers? A meta-analytic review. *Reading and Writing, 27*(9), 1703–1743. https://doi.org/10.1007/s11145-014-9517-0.

Hackman, D. A., & Farah, M. J. (2009). Socioeconomic status and the developing brain. *Trends in Cognitive Sciences*, *13*(2), 65–73. https://doi.org/10.1016/j.tics.2008.11.003.

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

Harris, K. R., & Graham, S. (2017). Self-regulated strategy development: Theoretical bases, critical instructional elements, and future research. In R. Fidalgo Redondo, K. Harris, & M. Braaksma (Eds.), *Design principles for teaching effective writing*. Netherlands: Brill.

Harris, K. R., Graham, S., & Mason, L. H. (2006). Improving the writing, knowledge, and motivation of struggling young writers: Effects of self-regulated strategy development with and without peer support. *American Educational Research Journal, 43*(2), 295–340. https://doi.org/10.3102/00028312043002295.

Harris, K. R., Ray, A., Graham, S., & Houston, J. (2019). Answering the challenge: SRSD instruction for close reading of text to write to persuade with 4th and 5th Grade students experiencing writing difficulties. *Reading and Writing, 32*(6), 1459–1482. https://doi.org/10.1007/s11145-018-9910-1.

Hayes, J. R., & Flower, L. S. (1980). Identifying the organization of writing processes. In I. W. Gregg & E. R. Steinberg (Eds.), *Cognitive processes in writing*. Hillsdale: Erlbaum.

Heck, R. H., & Thomas, S. L. (2015). *An introduction to multilevel modeling techniques: MLM and SEM approaches using Mplus*. Abingdon: Routledge.

Juel, C. (1988). Learning to read and write: A longitudinal study of 54 children from first through fourth grades. *Journal of Educational Psychology, 80*(4), 437–447. https://doi.org/10.1037/0022-0663.80.4.437.

Juel, C., Griffith, P. L., & Gough, P. B. (1986). Acquisition of literacy: A longitudinal study of children in first and second grade. *Journal of Educational Psychology, 78,* 243–255. https://doi.org/10.1037/0022-0663.78.4.243.

Kesane, I., & Ruiz, L. (2010). Contribucions de la comunitat científica internacional sobre aprenentatge de la lectura i superació del fracàs escolar [Contributions of the international scientific community on the learning of reading and overcoming school failure]. *Temps d'Educació, 38,* 115–134.

Kim, Y. S., Puranik, C., & Otaiba, S. A. (2015). Developmental trajectories of writing skills in first grade: Examining the effects of SES and language and/or speech impairments. *The Elementary School Journal, 115*(4), 593–613. https://doi.org/10.1086/681971.

Kline, R. (2011). *Principles and practice of structural equation modeling* (3rd ed.). New York: Guilford Press.

Korat, O., & Schiff, R. (2005). Do children who read more books know "what is good writing" better than children who read less? A comparison between grade levels and SES groups. *Journal of Literacy Research, 37*(3), 289–324. https://doi.org/10.1207/s15548430jlr3703_2.

Korat, O., Klein, P., & Segal-Drori, O. (2007). Maternal mediation in book reading, home literacy environment, and children's emergent literacy: A comparison between two social groups. *Reading and Writing*, 20(4), 361–398. https://doi.org/10.1007/s11145-006-9034-x.

Landis, J. R., & Koch, G. G. (1977). An application of hierarchical kappa-type statistics in the assessment of majority agreement among multiple observers. *Biometrics*. https://doi.org/10.2307/2529786.

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. https://doi.org/10.1016/j.cedpsych.2013.07.004.

Limpo, T., Vigário, V., Rocha, R., & Graham, S. (2020). Promoting transcription in third-grade classrooms: Effects on handwriting and spelling skills, composing, and motivation. *Contemporary Educational Psychology*. https://doi.org/10.1016/j.cedpsych.2020.101856.

MacWhinney, B. (2017). *Tools for analyzing talk. Part 2: The CLAN program*. Pittsburgh: Carnegie Mellon University. Retrieved from http://talkbank.org/manuals/CLAN.pdf.

Mason, L. H., Snyder, K. H., Sukhram, D. P., & Kedem, Y. (2006). TWA+ PLANS strategies for expository reading and writing: Effects for nine fourth-grade students. *Exceptional Children, 73*(1), 69–89. https://doi.org/10.1177/001440290607300104.

Masten, A. S., & Cicchetti, D. (2010). Developmental cascades. *Development and Psychopathology, 22*(3), 491–495. https://doi.org/10.1017/S0954579410000222.

McCutchen, D. (2000). Knowledge, processing, and working memory: Implications for a theory of writing. *Educational Psychologist, 35*(1), 13–23. https://doi.org/10.1207/S15326985EP3501_3.

Olive, T. (2014). Toward a parallel and cascading model of the writing system: A review of research on writing processes coordination. *Journal of Writing Research*. https://doi.org/10.17239/jowr-2014.06.02.4.

Olive, T., Favart, M., Beauvais, C., & Beauvais, L. (2009). Children's cognitive effort and fluency in writing: Effects of genre and of handwriting automatisation. *Learning and Instruction, 19*(4), 299–308. https://doi.org/10.1016/j.learninstruc.2008.05.005.

Perry, L. B., & McConney, A. (2010). Does the SES of the school matter? An examination of socioeconomic status and student achievement using PISA 2003. *Teachers College Record, 112*(4), 1137–1162.

Ravid, D., & Tolchinsky, L. (2002). Developing linguistic literacy: A comprehensive model. *Journal of Child Language, 29*(2), 417–447. https://doi.org/10.1017/S0305000902005111.

Rogers, L. A., & Graham, S. (2008). A meta-analysis of single subject design writing intervention research. *Journal of Educational Psychology, 100*(4), 879. https://doi.org/10.1037/0022-0663.100.4.879.

Salas, N. (2020). Non-phonological strategies in spelling development. *Frontiers in Psychology,*. https://doi.org/10.3389/fpsyg.2020.01071.

Salas, N., & Silvente, S. (2019). The role of executive functions and transcription skills in writing. A cross-sectional study across 7 years of schooling. *Reading and Writing, 33,* 877–905. https://doi.org/10.1007/s11145-019-09979-y.

Santangelo, T., & Graham, S. (2016). A comprehensive meta-analysis of handwriting instruction. *Educational Psychology Review, 28*(2), 225–265. https://doi.org/10.1007/s10648-015-9335-1.

Sarsour, K., Sheridan, M., Jutte, D., Nuru-Jeter, A., Hinshaw, S., & Boyce, W. T. (2011). Family socioeconomic status and child executive functions: The roles of language, home environment, and single parenthood. *Journal of the International Neuropsychological Society, 17*(1), 120–132. https://doi.org/10.1017/S1355617710001335.

Sirin, S. R. (2005). Socioeconomic status and academic achievement: A meta-analytic review of research. *Review of Educational Research, 75*(3), 417–453. https://doi.org/10.3102/00346543075003417.

Smith, S. S., & Dixon, R. G. (1995). Literacy concepts of low-and middle-class four-year-olds entering preschool. *The Journal of Educational Research, 88*(4), 243–253. https://doi.org/10.1080/00220671.1995.9941305.

Tierney, R., & Shanahan, T. (1991). Research on the reading-writing relationship: Interac- tions, transactions, and outcomes. In R. Barr, M. Kamil, P. Mosenthal, & D. Pearson (Eds.), *The handbook of reading research*. Harlow: Longman.

Van Ewijk, R., & Sleegers, P. (2010). The effect of peer socioeconomic status on student achievement: A meta-analysis. *Educational Research Review, 5*(2), 134–150. https://doi.org/10.1016/j.edurev.2010.02.001.

Verheyden, L., Van den Branden, K., Rijlaarsdam, G., Van den Bergh, H., & De Maeyer, S. (2010). Written narrations by 8-to 10-year-old Turkish pupils in Flemish primary education: A follow-up of seven text features. *Journal of Research in Reading, 33*(1), 20–38. https://doi.org/10.1111/j.1467-9817.2009.01430.x.

Whitaker, D., Berninger, V., Johnston, J., & Swanson, H. L. (1994). Intraindividual differences in levels of language in intermediate grade writers: Implications for the translating process. *Learning and Individual Differences, 6*(1), 107–130. https://doi.org/10.1016/1041-6080(94)90016-7.