Contemporary Educational Psychology 60 (2020) 101835

Contents lists available at ScienceDirect

Contemporary Educational Psychology

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

"It takes two": The added value of structured peer-assisted writing in explicit writing instruction

Fien De Smedt[a,*], Steve Graham[b], Hilde Van Keer[a]

[a] Department of Educational Studies, Ghent University, H. Dunantlaan 2, B-9000 Ghent, Belgium
[b] Mary Lou Fulton Teachers College, Arizona State University, PO Box 871811, Tempe (AZ), United States

ARTICLE INFO

Keywords:
Peer-assisted writing
Explicit writing instruction
Writing performance
Self-efficacy for writing
Writing motivation

ABSTRACT

High-quality writing instruction is vital to supporting developing writers as they learn to plan, compose, and revise text. It is equally important that such instruction enhances students' self-efficacy for writing as well as their motivation to write. The main aim of the present study was to investigate the incremental effect of peer-assisted writing in an explicit writing instruction program on Flemish upper-elementary students' writing performance, self-efficacy for writing, and writing motivation. A randomized control design, using multilevel analyses, was conducted to determine the differential effectiveness of two experimental writing treatments (EI + PA and EI + IND) compared to a business as usual control condition (BAU). Both experimental writing treatments involved explicit instruction in writing, with students in one condition writing with a peer (EI + PA) and students in the other condition writing individually (EI + IND). Participating classes (N = 431 students, N = 20 teachers) were randomly assigned to the three conditions and students were assessed before and after instruction. EI + PA students outperformed both EI + IND and BAU students on the writing measure in the instructed genre but not in the uninstructed genre. Additionally, although EI + PA students were more confident as to their capability (self-efficacy) to generate ideas when compared to their EI + IND counterparts, EI + PA students' writing motivation, characterized by internal or external motives, was significantly lower than EI + IND students. The findings of the present study corroborate and extend the limited number of prior studies illustrating the surplus value of peer-assisted writing in explicit writing instruction programs.

1. Introduction

Alarming results on students' poor writing have been documented in several large-scale writing assessments, revealing that students throughout the world often and consistently show difficulties with writing (Inspectie van het Onderwijs, 2010; National Center for Education Statistics, 2012; Vlaams Ministerie van Onderwijs en Vorming, 2019). This is especially true for students in upper-elementary grades (i.e., grades 5 and 6, 11–12 year olds), as they are just learning how to carry out the complex processes involved in planning, composing, and revising texts (Bereiter & Scardamalia, 1987; Flower & Hayes, 1981; McCutchen, 2008). Compared to students in the lower grades, upper-elementary students are able to engage in higher-order writing skills and strategies (e.g., planning) because the automatization of lower-order writing skills (e.g., handwriting) is increased. However, applying these higher-order writing strategies remains cognitively challenging for this age group (Cameron & Moshenko, 1996; McCutchen, Covill, Hoyne, & Mildes, 1994; McCutchen, Francis, & Kerr, 1997).

In addition to the concern regarding cognitive writing processes and outcomes, motivational challenges are equally important to consider as these are critical predictors of upper-elementary students' writing performance (De Smedt et al., 2018; Graham, Berninger, & Fan, 2007; Pajares & Valiante, 1997; Troia, Harbaugh, Shankland, Wolbers, & Lawrence, 2013). In this respect, Bruning and Kauffman (2016) argued that the challenges writers face, are at least as much related to motivational factors as they are to cognitive and linguistic factors. Educational writing practices, however, prioritize writing for evaluative purposes over writing to communicate thoughts, knowledge, and feelings. Consequently, students often experience writing as a difficult, effortful, and uninteresting activity (Hidi & Boscolo, 2006) and their motivation for writing starts to decline at the end of elementary education (Cleary, 1991; De Smedt et al., 2019).

Taking into account the abovementioned cognitive and motivational challenges in writing, the present study focusses explicitly on the age group of upper-elementary students. A critical issue in instruction

* Corresponding author.
E-mail addresses: Fien.DeSmedt@UGent.be (F. De Smedt), steve.graham@asu.edu (S. Graham), Hilde.VanKeer@UGent.be (H. Van Keer).

https://doi.org/10.1016/j.cedpsych.2019.101835

Available online 17 December 2019
0361-476X/

for these young students is how to support them instructionally so that they acquire needed writing processes, strategies, and skills as well as confidence as a writer and positive motivations for writing. Numerous meta-analyses on effective instructional writing practices in elementary grades have been conducted over the past years identifying multiple evidence-based writing practices (e.g., Graham, McKeown, Kiuhara, & Harris, 2012; Koster, Tribushinina, de Jong, & van den Bergh, 2015; Rogers & Graham, 2008). In the meta-analysis of Koster et al. (2015), that specifically focused on effective writing practices in upper-elementary grades, explicit writing instruction and peer-assistance were identified as two very promising evidence-based practices to foster upper-elementary students' writing. *Explicit writing instruction* refers to the explicit and systematic teaching of writing knowledge and strategies (Graham & Perin, 2007). In this respect, students acquire genre knowledge (i.e., knowledge on the aim, content, and structure of specific text genres) and learn strategies for planning, revising, and/or editing texts. *Peer-assisted writing* involves students working together to plan, draft, and/or revise texts (Graham & Perin, 2007). The main aim of the present study was to investigate the added value of peer-assisted writing in an explicit writing instruction program in view of fostering upper-elementary students' writing performance, self-efficacy for writing, and writing motivation. To this end, we compared students who received (a) explicit writing instruction with peer practice (EI + PA), (b) explicit instruction with individual practice (EI + IND), and (c) business as usual writing instruction (BAU). As the Flemish attainment targets state that students should be able to copy and write descriptive texts by the end of elementary education (Flemish Ministry of Education and Training, 2005) and considering that at this educational level, the descriptive text genre becomes increasingly important (Alexander & Jetton, 2000), this study specifically focusses on descriptive texts.

1.1. A social cognitive view on writing

The *social cognitive theory (SCT)* of Bandura (1986, 1989), that explains human behavior in terms of reciprocal determinism instead of one-sided determinism, serves as the underlying theoretical framework of the current study. More particularly, SCT posits that a person's behavior is influenced by the interaction between three determinants, namely (a) personal, (b) behavioral, and (c) environmental factors. Concerning the interaction between personal and behavioral factors, a person's thoughts and feelings affect his/her behavior, which in turn can influence his/her thought patterns and emotional reactions. Regarding the interaction between personal and environmental factors, a person's thoughts, feelings, and cognitive capabilities are affected by social influences through modeling, instruction, and social persuasion. Finally, a person's behavior affects his/her environment and is, in turn, influenced by the environmental conditions he/she creates. Based on this triadic reciprocal determinism of human behavior, SCT provides a framework for understanding how people actively shape and are shaped by their social environment. More particularly, Bandura (1986, 1989) claimed that people can acquire, expand, or transfer knowledge and skills by observing models they encounter in their environment.

Next to the prominent role of the social environment and observational learning in the SCT, Bandura (1977, 1997) theorizes that self-efficacy (i.e., beliefs in one's capability to perform a given task) is key in influencing human motivation, affect, and behavior. According to SCT, self-efficacy beliefs are derived from four sources of information: (a) performance accomplishments (i.e., successful accomplishments raise mastery expectations while repeated failures lower these), (b) vicarious experience (i.e., observing identifiable models who succeed or fail), (c) verbal persuasion (i.e., being encouraged to believe in oneself), and (d) emotional arousal (i.e., relating stressful situations to personal performance).

Based on SCT, Zimmerman and Risemberg (1997) developed *a social cognitive model of writing* that acknowledges writing as a self-regulated activity involving cognitive as well as social, motivational, and behavioral processes. Parallel with the three determinants of human behavior outlined in SCT, Zimmerman and Risemberg (1997) theorize that there are three forms of writer self-regulation: (a) personal or covert self-regulation involving the adaptive use of cognitive or affective strategies, (b) behavioral self-regulation pertaining strategies such as self-monitoring, self-consequences, and self-verbalizations, and (c) environmental self-regulation referring to the adaptive use of context-related strategies, such as environmental structuring or self-selected models. These self-regulatory processes interact reciprocally via strategic feedback loops that enable the writer to monitor the effectiveness of his/her self-regulatory strategies and to react accordingly (i.e., continuing successful strategy use or modifying and adjusting less successful strategy use). In the current study, we specifically focus on studying upper-elementary writers' *personal or covert self-regulation*, with particular attention for the application of cognitive writing strategies. Zimmerman and Risemberg (1997, p. 79) define these cognitive writing strategies as "rule governed methods for organizing, producing, and transforming written text". Previous research showed that explicit writing instruction is a promising instructional practice to enhance this kind of strategy use during writing (Koster et al., 2015). Given the prominent role of the writer's social environment in general, and the role of peers as important sources of social learning in particular (Bandura, 1986, 1989; Zimmerman & Risemberg, 1997), the present study focusses specifically on studying the incremental effect of peer-assisted writing in an explicit writing instruction program.

Within this triadic system of writing self-regulation, the strategic feedback loops are considered not only to foster writers' self-regulation but also to affect their *self-efficacy for writing*. Favorable feedback will increase the writer's sense of self-efficacy, whereas unfavorable feedback will result in decreased self-efficacy. Zimmerman and Risemberg (1997), therefore, posit that the triadic system of self-regulatory writing processes is closely related to writers' underlying self-efficacy for writing in two ways. First, novice writers who become more effective in applying self-regulatory methods during writing also feel more capable in writing. Second, increased self-efficacy beliefs will predict self-regulation during writing and play a major role in motivating writers to sustain their efforts and to ultimately write better texts. In this respect, the current study investigates cognitive writing outcomes, namely writing performance, as well as motivational writing outcomes, such as self-efficacy for writing and writing motivation.

1.2. Explicit writing instruction and peer-assisted writing: fostering students' writing performance, self-efficacy for writing, and writing motivation

An extensive body of evidence supports the effectiveness of *explicit instruction of writing knowledge and strategies* to promote upper-elementary students' use of cognitive writing strategies and, in turn, promote their writing performance (see *meta*-analyses: Graham et al., 2012; Koster et al., 2015). More particularly, past research demonstrated that upper-elementary students who are explicitly taught how to write become better writers (Bean & Steenwyk, 1984; De Smedt & Van Keer, 2018b; Fitzgerald & Teasley, 1986). Additionally, prior research also revealed that explicitly teaching students when, how, and why to plan and revise texts had a positive impact on students' overall writing performance (Bouwer, Koster, & van den Bergh, 2018; De Smedt & Van Keer, 2018b; Fidalgo, Torrance, Rijlaarsdam, van den Bergh, & Alvarez, 2015; Limpo & Alves, 2013; McKeown et al., 2016; Rietdijk, Janssen, van Weijen, van den Bergh, & Rijlaarsdam, 2017). Regarding the effect on motivational outcomes, there is only a small number of available studies and results obtained from these studies were inconclusive. Some studies found a positive effect of explicit instruction on elementary students' self-efficacy (Graham & Harris, 1989), while others did not (De Smedt, Graham, & Van Keer, 2018; Graham, Harris, & Mason, 2005). Similarly, some studies did not find effects of explicit instruction on students' motivation (e.g., Harris, Graham, & Mason, 2006), whereas

a recent study showed that students receiving explicit instruction evidenced more controlled motivated beliefs afterwards (e.g., De Smedt, Graham, & Van Keer, 2018). It must be noted, however, that the majority of these studies focused on lower-elementary grades (e.g., Graham et al., 2005; Harris et al., 2006) with only one study focusing on upper-elementary students (e.g., De Smedt, Graham, & Van Keer, 2018). The inconclusive findings might, therefore, be related to age-related differences. Next to possible developmental differences, the different motivational outcomes might also be affected by other internal or external factors (e.g., students' topic knowledge, interest).

Although research on the impact of *peer-assisted writing* in upper-elementary grades is scarce, the existing research in this area supports the suggestion that student collaboration during planning, composing, and/or revising text is effective (see *meta*-analyses: Graham et al., 2012; Koster et al., 2015). Peer-assisted writing is in this respect used as an overarching concept entailing different applications, such as peer tutoring (e.g., Grünke, Wilbert, Tsiriotakis, & Agirregoikoa, 2017; Yarrow & Topping, 2001), peer discussion and support (e.g., Graham et al., 2005; Harris et al., 2006), and peer review and feedback (e.g., Holliway, 2004; Hoogeveen & van Gelderen, 2018; Philippakos, 2017). These different peer-assisted writing applications have been effective in promoting upper-elementary students' writing (Graham et al., 2005; Harris et al., 2006; Yarrow & Topping, 2001). It must be noted, however, that most of the empirical studies conducted to date (see Graham et al., 2012; Koster et al., 2015) focus on peer revising and only a few studies involve students engaging in all writing processes together. As to the effect on motivational outcomes, research showed that peer-assistance can foster upper-elementary students' self-efficacy (e.g., Paquette, 2009) and motivation (e.g., De Smedt, Graham, & Van Keer, 2018).

1.3. Integrating explicit writing instruction and peer-assisted writing

Given the effectiveness of both abovementioned writing practices in upper-elementary grades (Koster et al., 2015) and taking into account the prominent role of the social environment to enhance writers' personal self-regulation, self-efficacy, and motivation (Bandura, 1986, 1989; Zimmerman & Risemberg, 1997), integrating peer-assisted writing in explicit writing instruction programs is considered to be beneficial. This combination may be especially fruitful due to their complementary nature (De Smedt & Van Keer, 2018b; Ferretti and Lewis, 2013), with explicit writing instruction focusing on teaching students writing knowledge and skills (Graham, 2006) and peer-assisted writing providing students with the opportunity to apply the knowledge and strategies taught (Daiute & Dalton, 1993) as well as providing students the opportunity to learn from each other (Graham, 2018). Notwithstanding the promising nature of this combination, prior research merging both evidence-based writing practices in elementary grades is scarce.

To our knowledge, only five previous studies investigated the combined effect of explicit writing instruction and peer-assisted writing on elementary students' writing. Below, we briefly discuss the scope and results of these studies. Two of these intervention studies focused on lower-elementary students (Graham et al., 2005; Harris et al., 2006), while the others targeted upper-elementary students (De Smedt, Graham, & Van Keer, 2018; De Smedt & Van Keer, 2018b; Hoogeveen & van Gelderen, 2018). Notwithstanding our focus on upper-elementary students, we opted to include a description of the studies of Graham et al. (2005) and Harris et al. (2006) focusing on lower-elementary grades, because these are exemplary for the current study.

First, Graham et al. (2005) studied the effect of Self-Regulated Strategy Development (SRSD) with and without peer support (N = 73 third-grade struggling writers). SRSD is the most extensively studied explicit writing instruction program focusing on improving students' writing self-regulation and its effectiveness has been validated in numerous studies (Graham, 2006; Graham et al., 2012). Graham et al. (2005) particularly studied whether peer assistance had an incremental effect in terms of (a) maintenance and generalization of planning and composing strategies taught during SRSD instruction and (b) students' writing motivation. Revising strategies as a third important writing process and self-efficacy for writing as a motivational factor were, however, not addressed in this investigation. Peer support was operationalized in terms of two peers working together to identify and discuss other contexts in which they could apply the strategies taught and consequently promote strategy use outside SRSD-lessons. The results confirmed the effectiveness of SRSD to support students' writing knowledge and performance and additionally showed that peer support augmented SRSD instruction by fostering students' knowledge and by facilitating generalization to two uninstructed genres. However, no incremental effect of peer-assistance was found for students' writing motivation

Harris et al. (2006) conducted a follow-up study to Graham et al. (2005) with 63 second-grade struggling writers. In line with the earlier study, this second investigation did not address revising strategies, but focused solely on teaching students planning and composing strategies. In contrast with the earlier study, this investigation focused on students' self-efficacy for writing instead of on writing motivation. The operationalization of peer support was slightly different when compared to the earlier investigation, as the instructor was also involved in promoting strategy use outside SRSD-lessons. The results from this second investigation by Harris and colleagues were consistent with the earlier study by Graham et al. (2005) confirming the effectiveness of SRSD and the incremental effect of peer help to support generalization. Finally, no additional effect of peer-assistance was found for students' self-efficacy for writing.

The third and fourth studies reported on a recent randomized controlled trial (N = 206 fifth and sixth graders), studying the distinct and combined impact of explicit writing instruction and peer-assisted writing on students' writing performance (De Smedt & Van Keer, 2018b) and students' self-efficacy for writing and writing motivation (De Smedt, Graham, & Van Keer, 2018). In contrast to the studies by Graham et al. (2005) and Harris et al. (2006), peer assistance was operationalized as an unprescribed form of collaboration, implying that students were not provided with specific guidelines, routines, and structures to coordinate their activities while planning, composing, and revising together. More particularly, students were grouped into fixed dyads and after the instructional phase (i.e., students were explicitly taught planning, composing, and revising strategies) they were given the assignment to collaborate during practice. Students were free on how to approach the collaboration, there were no fixed formats (e.g., peer feedback) nor guidelines or routines on how to write together. Regarding the effectiveness of explicit instruction of writing knowledge and strategies, the results were in line with both previous studies just described. Contrary to the studies conducted by Graham et al. (2005) and Harris et al. (2006), however, no additional additive effect was found for peer-assisted writing on students' writing performance. Finally, no additional effect of peer-assistance was found for students' self-efficacy for writing nor for their writing motivation.

The fifth study (N = 140 sixth graders) focused on the combined effect of peer assistance and explicit instruction of genre knowledge involving functions of linguistic indicators of time and place (Hoogeveen & van Gelderen, 2018). In contrast to the previous studies, this study exclusively focused on teaching students genre knowledge without providing them explicit strategy instruction. Furthermore, this study did not include any motivational outcome measures. In line with the previous studies, the results showed that peer-assisted writing combined with instruction in genre knowledge resulted in better writing.

1.4. The present study

Based on the literature overview on the effectiveness of blending

peer-assisted writing and explicit writing instruction, we can conclude that the empirical evidence underlines and supports the social cognitive view on writing (Zimmerman & Risemberg, 1997). More particularly, explicitly teaching upper-elementary students writing knowledge and strategies and providing them opportunities to work with peers, helps them in becoming more self-regulated writers. However, several gaps or unresolved issues remain in the research literature, warranting attention in the present study.

First, the number of prior studies focusing on both cognitive (i.e., writing performance) and motivational outcome measures (i.e., self-efficacy for writing and writing motivation) is limited. The studies that did focus on both types of outcome measures either studied writing performance and motivation (Harris et al., 2006) or writing performance and self-efficacy (Graham et al., 2005). Additionally, as the results on motivational outcome measures are inconclusive, no firm conclusions on the effect on students' motivation and self-efficacy for writing can be drawn. The present study, therefore, focusses on studying writing performance, self-efficacy, and motivation as outcome measures.

Second, although prior studies provided useful insights on the incremental effect of peer-assisted writing, these studies involved a relatively small number of students in comparison to the present investigation (De Smedt, Graham, & Van Keer, 2018; De Smedt & Van Keer, 2018b; Hoogeveen & van Gelderen, 2018), and half of the available studies focused just on struggling writers (Graham et al., 2005; Harris et al., 2006).

Third, the operationalization of peer assistance in previous studies was at a rather basic level by implementing peer support to promote transfer of writing strategies (Graham et al., 2005; Harris et al., 2006), by implementing unprescribed peer support (De Smedt, Graham, & Van Keer, 2018; De Smedt & Van Keer, 2018b), or by implementing peer assistance to support the acquisition of genre knowledge (Hoogeveen & van Gelderen, 2018). In this respect, none of the previous studies took into account the integration of peer-assisted writing during each phase of the writing process (e.g., focussing on planning and composing but not on revising). The present investigation is, therefore, the first to report on the effectiveness of integrating a structured system of peer-assisted writing during planning, composing, and revising processes. This application of peer-assisted writing is referred to as 'a structured system' as writer roles were imposed during each step of the writing process and collaboration between writing groups was integrated (De Smedt & Van Keer, 2018a).

Finally, in line with previous studies (Graham et al., 2005; Harris et al., 2006), we also study possible transfer effects to an uninstructed genre. More particularly, this intervention study focused on providing explicit instruction of descriptive writing knowledge and on teaching students general planning, composing, and revising strategies. Because the strategies were general and not genre-specific, we aimed to examine whether students were able to spontaneously transfer these strategies to an uninstructed genre, namely narrative texts. In contrast to previous research in which transfer was stimulated (Graham et al., 2005; Harris et al., 2006), this study is the first to investigate spontaneous transfer of writing strategies.

By taking into account the abovementioned issues, the present study moves the field forward by (a) focusing on both cognitive and motivational outcome measures, (b) conducting a larger-scale study with average achieving writers at the end of elementary education, (c) designing, implementing, and evaluating a structured system of peer-assisted writing systematically integrated in the explicit instruction of the complete writing process (i.e., planning, composing, and revising), and (d) studying spontaneous transfer of writing strategies to an uninstructed genre.

Based on the social cognitive view on writing (Zimmerman & Risemberg, 1997) and prior empirical research (Koster et al., 2015), the following hypotheses are put forward. As a first hypothesis, we predict that *experimental students (EI + PA and EI + IND) as compared to BAU students will perform better, will have higher levels of self-efficacy for writing, and will be more motivated to write after the intervention.* This hypothesis is based on the inclusion of three effective instructional approaches drawing upon the social cognitive model of writing (Zimmerman & Risemberg, 1997). First, EI + PA and EI + IND students are provided opportunities to compare and contrast model texts (Abbuhl, 2011; De Smedt & Van Keer, 2018b). In this way, EI + PA and EI + IND students can discover and acquire writing knowledge that enables them to become better writers. Second, EI + PA and EI + IND students can observe teachers who act as coping models for personal self-regulation during writing. More particularly, EI + PA and EI + IND teachers model cognitive writing strategies by explaining, demonstrating, and verbalizing their thoughts and actions during planning, composing, and revising texts (e.g., Bouwer et al., 2018; Graham et al., 2005). According the social cognitive view on writing (Zimmerman & Risemberg, 1997), observing teachers self-regulate their writing helps novice writers in gaining insights in the usefulness and consequences of the writing behaviour being modelled. Furthermore, EI + PA and EI + IND students' vicarious experience of observing their teachers during the act of writing might enhance their sense of self-efficacy and ultimately their motivation to write. Third, after studying model texts and observing their teachers, EI + PA and EI + IND students are provided with ample opportunities to practice the knowledge and strategies taught, while teachers gradually diminish their guidance (De Smedt & Van Keer, 2018b; Rietdijk, Janssen, van Weijen, van den Bergh, & Rijlaarsdam, 2017). In this way, the instruction moves from teacher modeling, to guided practice, to independent student writing. In this respect, we anticipate that as EI + PA and EI + IND students will become better self-regulated writers, their underlying sense of self-efficacy will increase, which in turn plays a major role in increasing EI + PA and EI + IND students' writing motivation (Zimmerman & Risemberg, 1997).

As a second hypothesis, we anticipate that *EI + PA students will perform significantly better, will have a higher sense of self-efficacy, and will be more motivated to write after the intervention compared to their EI + IND counterparts.* This hypothesis is based on the inclusion of the structured system of peer assistance in the EI + PA program. In contrast to EI + IND students, EI + PA students work together, observe each other, and learn from each other during each step of the complete writing process. As peers are considered important sources of observational learning (i.e., students can identify themselves more easily with peers than with their teachers, Couzijn, 1999), we predict that peer assistance will have an incremental effect on EI + PA students' writing performance, self-efficacy, and writing motivation. As a third hypothesis, we expect *no transfer effect* of the cognitive writing strategies that are explicitly taught in the EI + PA and EI + IND interventions to an uninstructed genre (i.e., narrative text). Since EI + PA and EI + IND students have to spontaneously transfer these strategies without any instructional support nor modeling on how to apply these strategies when writing narrative texts, we do not expect significant differences between EI + PA, EI + IND, and BAU students on the narrative posttest.

2. Method

2.1. Participants

Schools participating in the present study were recruited via an open call for participation in a wide-spread teacher journal. Schools who were interested could sign up for participation. In total, 20 teachers and their 431 fifth and sixth graders from 10 Flemish schools (Belgium) volunteered to participate in the study. This study with 20 participating classes is a more large-scale intervention study compared to previous studies that focused on combining explicit writing instruction and peer-assisted writing (i.e., 10 classes in De Smedt & Van Keer, 2018b, 12 classes in Graham et al., 2005; 11 classes in Harris et al., 2006; 5 classes in Hoogeveen & van Gelderen, 2018). As to the participating teachers, the majority were female (75%) and they taught

writing in either fifth grade (n = 7), sixth grade (n = 9), or in a multigrade fifth and sixth grade class (n = 4). Teachers' average age was 37.79 years (SD = 8.98, Min. = 25.21, Max. = 55.44) and they had on average 16.45 years of teaching experience in elementary education (SD = 8.81, Min. = 5, Max. = 35). Prior to the intervention, teachers' attitudes towards writing and writing instruction, their teacher efficacy for writing, the quality of their teacher preparation concerning writing, and the extent to which they already integrate peer-assisted writing into their everyday classroom practice was investigated by means of 5-point Likert scales (for more information on the scales, see De Smedt, Van Keer, & Merchie, 2016). Generally, the participating teachers indicated that they were positive towards writing (M = 3.11, SD = 0.69) and teaching writing (M = 3.61, SD = 0.83). Although teachers were dissatisfied with their teacher preparation concerning the teaching of writing (M = 2.45, SD = 0.83), they on average felt self-efficacious to teach struggling writers (M = 3.57, SD = 0.49) and attributed students' successful writing to their own instructional efforts (M = 3.35, SD = 0.58). Finally, teachers reported they infrequently applied peer-assisted writing in their everyday classroom practice (M = 2.72, SD = 0.76).

Concerning the students, 198 fifth graders (M age = 10.86, SD = 0.56) and 232 sixth graders (M age = 11.76, SD = 0.59) participated. Boys (50.1%) and girls (49.9%) were equally represented in the sample. The majority of the students were Dutch-speaking, that is the language of instruction in Flanders, Belgium (73.3%), while 13.0% of the students were bilingual (i.e., speaking Dutch and a foreign language at home) and 13.7% of the students spoke solely a foreign language at home (e.g., Turkish).

2.2. Design of the study

2.2.1. Conditions

A randomized control design with two measurement occasions (i.e., pretest and posttest) and three research conditions was applied (i.e., two experimental conditions: EI + PA and EI + IND and one business as usual condition: BAU). EI + PA and EI + IND students received explicit instruction regarding writing knowledge and strategies and either practiced writing with a peer (EI + PA) or individually (EI + IND). BAU students followed regular writing lessons presented by their teachers who applied their traditional writing approach by means of their regular textbooks and manuals to teach writing. Based on previous research in which the division of teachers/classes across conditions was not equally distributed (Palermo & Thomson, 2018), we randomly assigned the participating schools to one of the three conditions resulting in 8 EI + PA classes, 8 EI + IND classes, and 4 BAU classes. Consequently, teachers from the same school were assigned to the same condition to avoid possible design contamination effects (Rhoads, 2011). Concerning the randomization of the schools, we took into account the percentage of students being identified as an 'SES-student'[1] as an indicator to assign schools to one of the three conditions, thereby taking into account that schools with lower-class, middle-class, and upper-class students were equally represented in each condition.

To check the comparability between the conditions regarding teacher and student characteristics, chi-square and ANOVA analyses were performed. Regarding *the teachers*, chi-square analyses showed no significant differences in the distribution of grade ($\chi^2(4) = 8.01, p = .08$) and gender ($\chi^2(2) = 1.33, p = .51$) across conditions. Based on the ANOVA analyses, there was a statistically significant difference between the conditions regarding teachers' attitudes towards writing ($F(2, 18) = 4.54, p < .05, \eta^2 = 0.36$) indicating that BAU teachers (M = 2.38, SD = 0.32) were significantly less positive towards writing compared to EI + PA (M = 3.47, SD = 0.56) and EI + IND teachers (M = 3.13, SD = 0.68). Based on previous research (De Smedt, Van Keer, & Merchie, 2016; Graham, Harris, Fink, & MacArthur, 2001; Rietdijk, Van Weijen, Janssen, van den Bergh, & Rijlaarsdam, 2018; Tschannen-Moran, Hoy, & Hoy, 1998), teachers' attitudes towards writing are less likely to influence their instruction compared to teachers' efficacy for teaching writing. There were no significant differences between conditions in terms of teachers' attitudes towards writing instruction ($F(2, 18) = 0.69, p = .52$), age ($F(2, 18) = 0.24, p = .79$), teaching experience ($F(2, 18) = 0.55, p = .59$), evaluation of their education in writing instruction ($F(2, 18) = 0.01, p = .99$), efficacy for teaching struggling writers ($F(2, 18) = 0.40, p = .68$), efficacy for attributing students' successful writing to their own instruction ($F(2, 18) = 1.00, p = .39$), and the extent to which they already implemented peer-assisted writing into everyday writing activities ($F(2, 18) = 1.53, p = .25$).

As to *the students*, chi-square analyses revealed significant differences in the distribution of grade ($\chi^2(2) = 11.20, p < .01$), gender ($\chi^2(2) = 6.82, p < .05$), and home language ($\chi^2(4) = 41.36, p < .001$). More particularly, the percentage of sixth graders was higher in the EI + PA condition, more female students were included in the EI + IND condition, and a higher percentage of students speaking either a foreign language or Dutch in combination with a foreign language was represented in the EI + IND condition. Table 1 contains a summary of student information per research condition. In our main analyses, we verified whether students' gender, grade, and home language was significantly related to students' pretest scores. In case of significant relationships, gender, grade, and/or home language were included in the models to control for the initial student differences at pretest (see data analysis).

2.2.2. Procedure

To study the effectiveness of both the experimental and BAU writing programs, a stepwise procedure was applied: (a) an information session for all participating teachers, (b) a training session for the experimental teachers, (c) pretest administration, (d) a 10-week intervention period, and (e) posttest administration.

First, EI + PA, EI + IND, and BAU teachers were invited by the lead author at the university for an *information session* on how to administer the pretest and posttest (i.e., three identical information sessions were held for each group of teachers). Immediately after this information session, EI + PA and EI + IND teachers also received a training session. Regarding the information session, experimental and BAU teachers participated in a 30-min informational session in which the lead author explained the protocol for administering the writing tests and questionnaires during pretest and posttest. Additionally, all teachers received a document outlining the administering protocol in detail and all writing tests and questionnaires were delivered to the teachers.

Second, to support the teachers in implementing the experimental writing program, EI + PA and EI + IND teachers followed a *researcher-directed and condition-specific training session*. More particularly, two training sessions were organized: one session was intended for the EI + PA teachers while the other session was organized for the EI + IND teachers. Both sessions contained a 3-h group training in which the teacher manuals (EI + PA: 92 pages and EI + IND: 81 pages), containing a comprehensive description of the background, aims, organization of the intervention, and detailed lesson scenarios, were discussed. Additionally, during hands-on practices all experimental teachers learned how to explicitly teach writing knowledge and strategies using the explicit instruction procedures. EI + PA teachers were also provided with specific guidelines on how to implement and structure peer-assisted writing while the EI + IND teachers were instructed on how to implement individual practice.

Third, *pretest* data were collected by the teachers, following a detailed protocol, within the classroom context and during regularly

[1] In Flemish elementary education this indicates that a student receives a school allowance, or that the student is a non-native speaker, or that the student's mother has not obtained a certificate of secondary education. In light of privacy, these data are only available at school level.

Table 1
Overview of student characteristics per research condition.

	EI + PA		EI + IND		BAU	
	N	%	*N*	%	*N*	%
Grade						
Fifth grade	70	37.2	72	49.7	56	57.1
Sixth grade	117	62.2	73	50.3	42	42.9
Missing	1	0.6%	0	0.0	0	0.0
Total	188	100	145	100	98	100
Gender						
Male	101	53.7	60	41.4	55	56.1
Female	87	46.3	85	58.6	43	43.9
Missing	0	0.0	0	0.0	0	0.0
Total	188	100	145	100	98	100
Home language						
Dutch	140	74.5	83	57.2	87	88.8
Other language	15	8.0	38	26.2	5	5.1
Dutch + other language	30	16.0	20	13.8	5	5.1
Missing	3	1.5	4	2.8	1	1.0
Total	188	100	145	100	98	100

Note. EI + PA = Explicit writing instruction + peer-assisted writing. EI + IND = Explicit writing instruction + individual writing. BAU = Business as usual.

scheduled class hours. To avoid testing overload for students, teachers administered the writing tests on two different days.

Fourth, a 10-week *intervention period* took place during which each teacher was observed once by the main researcher (see Fidelity of implementation). After the observed lesson, the teachers received individual coaching if needed in implementing the intervention. We opted for this time-based approach as the sequence of the writing lessons had to fit within the trimestral system in Flemish schools.

Finally, *posttest* data were again collected by the teachers, following a detailed protocol, within the classroom context and during regularly scheduled class hours. Similarly to prestest data collection, the tests were administered on two different days.

2.2.3. Intervention

Two experimental writing lesson programs were developed: the EI + PA and EI + IND program. Both programs focused on explicitly teaching students writing knowledge (i.e., genre knowledge and text structure knowledge) and how to plan, compose, and revise descriptive texts. The difference between both experimental programs was in the mode of delivery: EI + PA students practiced writing with a peer, while EI + IND students practiced writing individually. For more information on the design principles, instructional teaching activities, learning activities, and concrete writing lesson programs and instructional materials, see the systematic and analytic description of the instructional writing programs by (De Smedt & Van Keer, 2018a).

Overview of the writing lessons. Both EI + PA and EI + IND writing programs consisted of 11 writing lessons, spread over ten consecutive weeks. The experimental teachers taught one lesson of 50 min per week, with the exception of the first week in which they taught two lessons (i.e., EI + PA teachers taught lesson 2A while EI + IND teacher taught lesson 2B (see Table 2 for an overview of the writing lessons).

Explicit writing instruction (EI + PA and EI + IND). Both the EI + PA and EI + IND program were characterized by three instructional writing practices: (a) explicit instruction of writing knowledge, (b) explicit strategy instruction, and (c) providing optimal writing opportunities so students can practice while gradually diminishing guidance.

Concerning the *explicit instruction of writing knowledge*, the intervention consisted of one instruction lesson in which students were taught genre-specific knowledge such as, the content, goal, and text structure of the descriptive genre so students get acquainted with how such texts are composed (cf., lesson 1). The teachers introduced the descriptive writing genre by offering students two varying descriptive model texts. Students had to compare and contrast these models to discover, identify, and label the goal, content, and structure of descriptive texts. Afterwards, students received a memory card summarizing the important characteristics of the genre.

Regarding the *explicit strategy instruction*, the intervention consisted of four instruction lessons in which students were explicitly taught how to plan (cf., lesson 3), write (cf., lesson 4), and revise (cf., lesson 6) descriptive texts and how to apply and integrate all writing processes together (i.e., planning, composing, and revising; cf., lesson 7). More particularly, students were taught strategies for: (a) gathering and organizing main ideas and additional information by means of a planning card and scheme, (b) composing their text based on their planning taking into account the structure and content of descriptive texts, and (c) revising the content and structure of their text. These strategies are explicitly taught using a stepwise instructional procedure. First, the teachers pointed out the value of a specific strategy (i.e., planning, composing, revising). Second, the teachers activated students' background knowledge by discussing whether, when, and how students already applied specific planning, composing, or revising strategies. Third, the teacher modelled each strategy in front of the class by visualizing the strategy on the board and by thinking aloud. While modelling, the teacher encouraged students to actively participate in the planning, composing, or revising process. After teacher modelling, students received strategy cards, summarizing the important steps of the different writing strategies. Immediately after each strategy was modelled and discussed, teachers provided short writing tasks so students could practice the writing strategies separately. During lesson 7, the teacher guided students throughout the entire writing process (i.e., planning, composing, and revising) and offered students an integration card, summarizing the previous cards (i.e., memory card and strategy cards) in a nutshell.

Regarding the final instructional practice, teachers provided *optimal writing opportunities for students to practice while gradually diminishing guidance* during four practice lessons (i.e., lesson 5, 8, 9, 10). More particularly, the teacher offered students challenging writing tasks in view of practicing the writing strategies taught. During practice, the teacher provided feedback concerning students' texts and writing process. If teachers identified (groups of) writers who were struggling with the writing assignment, the teacher supported these (groups of) writers by providing additional instruction and by offering the help they needed (e.g., using the separate strategy cards). Furthermore, the teacher challenged (groups of) writers who became more proficient in approaching the writing assignments to gradually diminish the use of the supporting materials (e.g., diminishing the use of the strategy cards).

Peer-assisted writing (EI + PA). During lesson 2A, the EI + PA teachers first tried to create engagement and mutual trust between writing partners by: (a) grouping their students into heterogeneous dyads taking into account students' writing proficiency level on the one hand and matching students' personalities on the other hand[2], (b) keeping the dyads fixed for the duration of the intervention, and (c) organizing a class discussion in which students could agree on rules that foster peer-assisted writing (e.g., listening to each other). Second, the teacher structured the collaboration by introducing three roles. The role of 'the thinker' applied to both students in the dyads implying that each student always had to think of and invent good ideas to write about, think

[2] Teachers ranked all their students ranging from 'the most skilful writer' to the 'the most struggling writer'. Subsequently, they split the ranking in half, so they were able to pair the most skilful writer in the first half to the most skilful writer in the second half. They followed this procedure until all students had a writing partner. If a dyad consisted of students with clashing personalities, the teacher adjusted the pairing procedure. In case of an uneven number of students in the class, the teacher exceptionally created one group of three students.

Table 2
Overview of the writing lesson programs.

Week	Lesson (50 min/lesson)	Focus of the lesson	EI + PA	EI + IND
1	1	**Instruction lesson**: Explicitly teaching students **writing knowledge** by comparing and contrasting two model texts	✓	✓
	2A	**Agree on rules on writing collaboratively**	✓	
	2B	**Agree on rules on writing individually**		✓
2	3	**Instruction lesson**: Explicitly teaching students the **planning strategy**	✓	✓
3	4	**Instruction lesson**: Explicitly teaching students the **composing strategy**	✓	✓
4	5	**Practice lesson:** Planning and writing a text	✓	✓
5	6	**Instruction lesson**: Explicitly teaching students the **revising strategy**	✓	✓
6	7	**Practice lesson:** Revising a text	✓	✓
7	8	**Instruction lesson**: **Integrating the strategies**	✓	✓
8	9	**Practice lesson:** Planning, composing, and revising a text	✓	✓
9	10	**Practice lesson:** Planning, composing, and revising a text	✓	✓
10	11	**Practice lesson:** Planning, composing, and revising a text	✓	✓

Note. EI + PA = Explicit writing instruction + peer-assisted writing. EI + IND = Explicit writing instruction + individual writing.

about the content and structure of the text, think about ways to improve the text, … The second and third role were exchangeable implying that the dyads switched these roles each lesson. 'The strategy card reader' was responsible for reading and following the strategy cards and monitoring the writing process, while 'the reporter' was responsible for writing notes, writing down the text they constructed in pairs, and revising the text if they jointly decided to make adjustments. Next to these roles, the teacher also structured the collaboration by offering a shared writing portfolio to each dyad. Third, the teachers also modelled how students could write in pairs by demonstrating the role of the strategy card reader and by modelling appropriate collaboration and interactions skills (e.g., negotiating or compromising). Finally, the teacher also created collaboration opportunities across the different dyads. More particularly, each dyad had to exchange their written work with another dyad so they were able to provide peer feedback on each other's work.

Individual writing (EI + IND). During lesson 2B, the EI + IND teacher created an individual writing environment by: (a) discussing the importance and added value of independent and individual work during writing, (b) organizing a class discussion in which students could agree on rules that foster individual and independent writing (e.g., work quietly), and (c) offering students individual writing portfolios.

2.3. Fidelity of implementation

2.3.1. Students' writing portfolios

Based on the number of completed writing tasks in the students' writing portfolio, we verified that on average, 99.20% of the experimental students completed the eleven writing lessons.

2.3.2. Teacher logbooks

The experimental teachers provided information on the date, hour, and total time spent on each lesson via logbooks with structured protocols (De Smedt & Van Keer, 2018b). Teachers reported spending on average 47.67 min on each writing lesson (SD = 6.90) that approximates the prescribed time of 50 min per lesson. There were no significant differences between both experimental conditions in this respect ($t(12) = 0.71$, $p = .49$).

2.3.3. Observations in experimental classes

The lead author observed one lesson of each experimental teacher, resulting in 16 observations in total. First, *teachers' time on/off task* was measured (Bouwer, Koster, & van den Bergh, 2018; De Smedt & Van Keer, 2018b). Observational results showed that teachers spent on average 48.12 min on the observed lessons (SD = 4.75), that approximates the prescribed time of 50 min per lesson. There were no significant differences between EI + PA and EI + IND teachers ($t(14) = 1.17$, $p = 0.26$). Teachers were on task on average 97.94% of the total observed lesson time. Almost half of their time was devoted to classroom interaction (43.08%) while 31.91% was spent on plenary instruction and 25.01% on monitoring students' writing progress during student practice.

Second, the observer assessed the *global quality* (i.e., the quality of instruction, class management, and student engagement) of each observed lesson on a 5-point Likert scale (based on Vaughn et al., 2011). For the *quality of instruction* and *class management*, a set of quality indicators (e.g., 'uses intervention time efficiently' or 'redirects off-task behavior quickly and efficiently') were outlined. Taking into account the number of quality indicators that were discerned by the observer, the quality of instruction and class management were scored as follows: score 1 = no quality indicators were determined, score 2 = a minority of the quality indicators were determined, score 3 = half of the quality indicators were determined, score 4 = a majority of the quality indicators were determined, and score 5 = all quality indicators were determined). *Student engagement* was scored based on the number of students involved during the learning activities (i.e., score 1 = almost no one, score 2 = the minority of the students, score 3 = half of the students, score 4 = the majority of the students, and score 5 = almost all students). Observational data showed that the quality of instruction was high (EI + PA: M = 4.88, SD = 0.35; EI + IND: M = 3.75, SD = 0.50), that experimental teachers managed their class effectively (EI + PA: M = 4.50, SD = 0.53; EI + IND: M = 4.25, SD = 0.71), and that students were actively engaged during the observed lessons (EI + PA: M = 4.75, SD = 0.46; EI + IND: M = 4.38, SD = 0.52) in both experimental conditions.

Finally, the observer evaluated the *quality of implementation* by assessing whether the critical elements of the interventions concerning explicit strategy instruction and mode of delivery were implemented as intended. More particularly, the quality of implementation was assessed by means of a 5-point Likert scale ranging from 'not observed' to 'observed with high alignment with the teacher manual' (De Smedt & Van Keer, 2018b; Vaughn et al., 2011). Table 3 shows that both EI + PA and EI + IND teachers implemented the intervention in high alignment with the teacher manual.

2.3.4. Observations in BAU classes

The lead author also observed one lesson of each BAU teacher, with the exception of one teacher who did not give his permission. This resulted in a total of 3 observations in the BAU classes. BAU teachers spent on average 50.00 min (SD = 7.55) on the observed lessons and were 94.39% of the total observed lesson time on task. About half of their time was devoted to monitoring students' progress during practice (47.31%) while 34.41% was spent on interacting with students. BAU teachers spent only 18.28% of their time on plenary instruction. As to the overall quality of instruction, BAU teachers' instruction was of high quality (M = 4.67, SD = 0.58), they managed their class effectively (M = 4.67, SD = 0.58), and BAU students were engaged (M = 4.67,

Table 3
The quality of implementation[a]: observational data assessing critical elements of the intervention.

	M (SD)	
	EI + PA	EI + IND
Explicit strategy instruction		
Pointing out the value of the strategy	5.00 (0.00)	4.00 (2.00)
Discussing students' strategy use	4.00 (1.85)	4.00 (2.00)
Modeling	4.75 (0.46)	4.38 (0.43)
Introducing strategy cards	2.06 (1.82)	3.75 (1.04)
Individual writing		4.63 (0.58)
Peer-assisted writing	5.00 (0.00)	

Note. EI + PA = Explicit writing instruction + peer-assisted writing. EI + IND = Explicit writing instruction + individual writing.

[a] To assess the quality of implementation, the critical elements of the intervention concerning the explicit strategy instruction and the mode of delivery were measured using a 5-point Likert scale ranging from 'not observed' to 'observed with high alignment with the teacher manual'.

SD = 0.58). The observer also assessed the instructional writing practices and mode of delivery to determine whether there was any contamination of instruction in the BAU classes. More particularly, the observer checked whether critical ingredients of the interventions were also implemented in the BAU conditions by assessing each critical aspect of the EI + PA or EI + IND interventions as follows: 0 = 'not observed', 1 = 'observed but not in alignment with the EI + PA or EI + IND intervention', or 2 = 'observed with high alignment with the EI + PA or EI + IND intervention'. Table 4 shows that none of the critical aspects of the EI + PA or EI + IND were identically implemented in the BAU classes. Observational data did indicate that 66.7% of the BAU teachers implemented individual writing in their writing lessons. This is, however, not surprising as individual practice is more common during writing instruction compared to peer-assisted writing (De Smedt, Van Keer, & Merchie, 2016). Although the majority of BAU students practiced individually, the text genre and writing assignments differed from the EI + IND students. More particularly, 2 BAU teachers focused on writing stories (i.e., inventing a story based on pictures and writing a fairytale) while one BAU teacher focused on how to write an e-mail to express your interest in participating in a competition.

2.4. Measures

2.4.1. Writing performance

Writing tests. In total, students had to complete four writing tests focusing on two writing genres. More particularly, students wrote one descriptive and one narrative text at pretest and one descriptive and one narrative text at posttest. The writing assignments were similar to ensure a stable level of complexity across measurement occasions. As for the descriptive texts, students were asked to present themselves to a new classmate (i.e., pretest) and to present one of their family members (i.e., posttest). As for the narrative texts, students were asked to write a story on how and why a bottle washed ashore (i.e., pretest) and on how and why a treasure ended up in a forest (i.e., posttest). To avoid additional variance in the study design and as the aim of the study was to compare conditions instead of measurement occasions, the writing topics were not counterbalanced.

Assessing text quality. To minimize presentation effects, students' handwritten texts were typed and spelling, punctuation, and capitalization errors were corrected (Graham, Harris, & Hebert, 2011). The texts were assessed by combining two comparative rating procedures, namely (a) comparative judgement (Lesterhuis, Verhavert, Coertjens, Donche, & De Maeyer, 2017; Pollitt, 2012) and (b) a benchmark rating procedure (De Smedt & Van Keer, 2018b; Tillema, van den Bergh, Rijlaarsdam, & Sanders, 2012; Bouwer et al., 2016). Following the *comparative judgement procedure*, pairs of texts are compared and the best text of each pair is chosen by multiple raters. In this way, the quality of the texts can be ranked on a scale ranging from very low to very high text quality. Regarding the *benchmark rating procedure*, experts construct a continuous scale with benchmark texts that represent different levels of writing quality. Following this procedure, texts are not compared pairwise, but texts are compared to the set of benchmark texts. Previous research has shown that both rating procedures are, in terms of reliability and validity, promising for the assessment of writing (Bouwer, Lesterhuis, De Smedt, Van Keer, & De Maeyer, 2019).

Recently, the value of integrating both procedures has been pointed out in writing assessment research (Lesterhuis, Verhavert, Coertjens, Donche, & De Maeyer, 2017; Bouwer, Lesterhuis, De Smedt, Van Keer, & De Maeyer, 2019). More particularly, experts recommend that a (sub) set of texts should first be assessed by means of comparative judgement to calibrate and select the benchmark texts for the rating scale. Second, this benchmark rating scale can then be used for the assessment of the full set of texts. Following this two-stage process, we first randomly selected a subset of texts written during pretest (i.e., 150 descriptive texts and 150 narrative texts). These subsets were assessed by 64 undergraduate students who were trained at rating texts written by late-elementary school students. More particularly, 32 raters made the pairwise comparisons for the descriptive texts, while the other 32 raters made the pairwise comparisons for the narrative texts. The Digital Platform for the Assessment of Competences (www.d-pac.be) was used for the comparative judgement. Concerning the descriptive texts, an average of 22.1 comparisons per text were made, resulting in a reliable rank order (SSR = 0.84). As to the narrative texts, an average of 18.9 comparisons per text were made, also resulting in a reliable rank order (SSR = 0.80). Based on these rank orders, we selected the benchmark texts by transforming the logitscores on the rank orders to standardized scores and by selecting texts with a standardized *z*-score of −2, −1, 0, 1, 2 that represented the baseline range in text quality (Schoonen, 2005). The five benchmark texts were placed on a continuous scale in which the score for the benchmark text with an average text quality was 100 and the interval between the benchmark texts was 15. Finally, all pretests and posttests were randomly assigned to two independent

Table 4
The quality of implementation: observational data assessing contamination of instruction in the business as usual classes.

	Not observed	Observed but not in alignment with EI + PA or EI + IND	Observed and in high alignment with EI + PA or EI + IND
Explicit instruction of writing knowledge	33.3%	66.7%	0.0%
Explicit strategy instruction			
Pointing out the value of a strategy	100.0%	0.0%	0.0%
Discussing students' strategy use	100.0%	0.0%	0.0%
Modeling	100.0%	0.0%	0.0%
Other operationalizations of explicit strategy instruction	100.0%	0.0%	0.0%
Peer-assisted writing	33.3%	66.7%	0.0%
Individual writing	33.3%	0.0%	66.7%

Note. EI + PA = Explicit writing instruction + peer-assisted writing. EI + IND = Explicit writing instruction + individual writing.

trained raters who assessed text quality using the benchmark rating scale. In total, 17.95% of the texts were double-scored and interrater reliability was high (descriptive text quality: Pearson $r = 0.84$, $p < .001$ and Krippendorff's $\alpha = 0.83$; narrative text quality: Pearson $r = 0.79$, $p < .001$ and Krippendorff's $\alpha = 0.80$).

2.4.2. Self-efficacy for writing

Based on the writing self-efficacy framework of Bruning, Dempsey, Kauffman, McKim, and Zumbrunn (2013), students' self-efficacy for ideation (i.e., believing in one's capabilities to generate ideas), conventions (i.e., believing in one's capabilities to apply language and writing conventions), and regulation (i.e., believing in one's self-regulation capabilities during writing) was measured by means of the Self-Efficacy for Writing Scale (SEWS; Bruning et al., 2013).

The structure and the fit of the scales has been tested in prior research with late-elementary students in Flanders (De Smedt et al., 2018) and the scales have also been used in previous intervention research studying self-efficacy for writing as an outcome measure (De Smedt, Graham, & Van Keer, 2018). The SEWS consists of twelve items on a 100-point scale. More particularly, four items measure self-efficacy for ideation (e.g., "I can put my ideas into writing"), four items measure self-efficacy for conventions (e.g., "I can punctuate my sentences correctly"), and four items measure self-efficacy for regulation (e.g., "I can punctuate my sentences correctly"). The CFA results showed that the stability of the three-factor model provided a good fit to the data in the current sample at pretest (SB $\chi^2(51) = 88.33$, $p < .001$, CFI = 0.96, RMSEA = 0.04, SRMR = 0.05) and posttest (SB $\chi^2(51) = 105.29$, $p < .001$, CFI = 0.94, RMSEA = 0.05, SRMR = 0.04). In the current sample, internal consistency for both subscales was high at both pretest (i.e., ideation: Bentler's ρ = 0.89; conventions: Bentler's ρ = 80, and regulation: Bentler's ρ = 78) and posttest (i.e., ideation: Bentler's ρ = 0.87; conventions: Bentler's ρ = 84, and regulation: Bentler's ρ = 83; Bentler, 2009).

2.4.3. Writing motivation

Students' writing motivation was measured by means of the SRQ-Writing motivation scale (De Smedt et al., 2018), that is based on the SRQ-Reading motivation scale (De Naeghel, Van Keer, Vansteenkiste, & Rosseel, 2012) and is rooted in the self-determination theory (SDT; Ryan & Deci, 2000). SDT distinguishes autonomous from controlled writing motivation. The former more particularly refers to motives such as writing because of intrinsic pleasure or because of the identified value of writing, while the latter refers to motives such as writing because of external or internal pressure.

The structure and the fit of the scales has been tested in prior research with late-elementary students in Flanders (De Smedt et al., 2018) and the scales have also been used in previous intervention research studying writing motivation as outcome measure (De Smedt, Graham, & Van Keer, 2018). The SRQ-Writing motivation contains seventeen items on a five-point Likert scale, ranging from 1 (disagree a lot) to 5 (agree a lot). Eight items measure students' autonomous writing motivation (e.g., "I write a text because it's fun to write" or "I write a text because it is important to me to write") and nine items measure students' controlled writing motivation (e.g., "I write a text because I will feel ashamed of myself if I don't write" or "I write a text because others will punish me if I don't write"). We conducted CFA to examine the structure of the SRQ-Writing motivation in the current sample. The results revealed that the fit of the two-factor model was acceptable at pretest (SB $\chi^2(1\ 1\ 6) = 259.94$, $p < .001$, CFI = 0.93, RMSEA = 0.06, SRMR = 0.06) and posttest (SB $\chi^2(1\ 1\ 6) = 291.38$, $p < .001$, CFI = 0.91, RMSEA = 0.06, SRMR = 0.06). In the current sample, internal consistency for both subscales was high at both pretest (i.e., autonomous writing motivation: Bentler's ρ = 0.92 and controlled writing motivation: Bentler's ρ = 0.80) and posttest (i.e., autonomous writing motivation: Bentler's ρ = 0.92 and controlled writing motivation: Bentler's ρ = 0.81; Bentler, 2009).

2.5. Data analysis

Due to the nested data structure (i.e., students at level 1, classes at level 2, schools at level 3), multilevel analyses were conducted using MLwiN 2.29 (Rasbash, Charlton, Browne, Healy, & Cameron, 2009). By taking into account the hierarchical nesting of the data in multilevel models, efficient estimates of regression coefficients (using the iterative generalized least squares estimation procedure - IGLS), correct standard errors, and significance tests are obtained (Goldstein, 1995; Rasbash, Steele, Browne, & Goldstein, 2012). Not all students were present at both measurement occasions, with respectively 2.32% and 2.55% of the students missing at pretest and posttest. However, the occurrence of missing data, assumed they are missing at random, does not pose a problem for the application of multilevel models because of their flexibility to deal with unbalanced data structures (Jones & Duncan, 1998; Snijders & Bosker, 1999).

A four-step procedure was implemented. First, the fully unconditional three-level null model was computed for each posttest response variable: (a) quality of the descriptive text, (b) quality of the narrative text, (c) self-efficacy for ideation, (d) self-efficacy for conventions, (e) self-efficacy for regulation, (f) autonomous writing motivation, and (g) controlled writing motivation. Because the variability in all response variables could not be attributed to differences between schools, we decided to proceed with multilevel analyses considering two levels (students at level 1 and classes at level 2). Consequently, the fully unconditional two-level null model was computed for each posttest response variable. Second, we verified whether students' gender, grade, and home language was significantly related to students' pretest scores. In case of significant relationships, gender, grade, and/or home language were included in the models to control for the initial student differences at pretest. Additionally, students' pretest scores on each of the response variables (centered around the mean score) was included as a respective covariate to control for baseline performance, self-efficacy, and motivation. The third step of the procedure consisted of adding the categorical variable 'condition' to the model. To represent the three research conditions two dummy-variables were used, contrasting the experimental conditions with the BAU condition and comparing the EI + PA and EI + IND condition to each other. Finally, standardized regression coefficients (*SD*) were calculated to interpret the effect sizes for all significant effects (Cohen, 1977).

3. Results

3.1. Descriptive results

Table 5 displays the descriptive statistics for all study variables at pretest and posttest and the correlations between the variables. Pretest variables showed moderate to high positive correlations with posttest variables (i.e., ranging from $r = 0.27$, $p < .01$ to $r = 0.71$, $p < .01$). Students' descriptive writing performance showed small to moderate correlations with their writing performance on the narrative text (i.e., ranging from $r = 0.11$, $p < .05$ to $r = 0.35$, $p < .01$). The different types of self-efficacy were moderately to strongly correlated (i.e., ranging from $r = 0.32$, $p < .01$ to $r = 0.67$, $p < .01$).

3.2. Multilevel results

Tables 6–12 present the summaries of the model estimates for the two-level analysis of students' descriptive text quality (Table 6), narrative text quality (Table 7), self-efficacy for ideation (Table 8), self-efficacy for conventions (Table 9), self-efficacy for regulation (Table 10), autonomous writing motivation (Table 11), and controlled writing motivation (Table 12). The intercepts (β_{0ij}) in the fixed part of the null models represent the overall mean of each response variable for all students in all classes. The random part in the null models revealed that, the variances at class level (σ^2_{u0}) were either significantly different

Table 5
Correlations and descriptive statistics for all study variables.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
Descriptive text quality (pretest)	1.00													
Descriptive text quality (posttest)	0.29**	1.00												
Narrative text quality (pretest)	0.30**	0.25**	1.00											
Narrative text quality (posttest)	0.11*	0.35**	0.27**	1.00										
Autonomous motivation (pretest)	0.22**	0.24**	0.15**	0.21**	1.00									
Autonomous motivation (posttest)	0.21**	0.21**	0.15**	0.11*	0.71**	1.00								
Controlled motivation (pretest)	0.04	−0.5	−0.3	−0.12*	−0.00	0.01	1.00							
Controlled motivation (posttest)	−0.05	−0.11*	−0.10	−0.12*	−0.14**	−0.16**	0.52**	1.00						
Self-efficacy ideation (prestest)	0.16**	0.17**	0.19**	0.25**	0.49**	0.34**	−0.02	−0.11*	1.00					
Self-efficacy ideation (posttest)	0.14**	0.18**	0.23**	0.19**	0.35**	0.39**	−0.01	−0.06	0.55**	1.00				
Self-efficacy conventions (pretest)	0.15**	0.21**	0.09	0.18**	0.34**	0.27**	−0.05	−0.10	0.41**	0.39**	1.00			
Self-efficacy conventions (posttest)	0.08	0.21**	0.11*	0.13*	0.29**	0.25**	−0.02	−0.05	0.35**	0.67**	0.56**	1.00		
Self-efficacy regulation (pretest)	0.18**	0.16**	0.17**	0.15**	0.51**	0.37**	0.05	−0.02	0.48**	0.36**	0.48**	0.32**	1.00	
Self-efficacy regulation (posttest)	0.14**	0.16**	0.18**	0.18**	0.37**	0.42**	0.05	0.02	0.33**	0.59**	0.41**	0.54**	0.57**	1.00
M	100.26	100.54	96.17	96.25	3.21	3.22	2.77	2.71	68.17	71.62	77.64	76.05	69.11	65.84
SD	13.26	12.49	11.47	9.91	0.98	0.97	0.75	0.80	23.10	20.80	17.34	18.19	20.24	22.45
N	421	420	414	416	404	414	397	389	425	423	425	421	424	420

Note. *Correlations are significant at p < .05; ** correlations are significant at p < .01.

Table 6
Summary of the model estimates for the two-level analysis of students' descriptive text quality at posttest.

	Model 0	Model 1	Model 2[a]
Fixed part			
β_{0ij} Intercept	100.61 (1.47)***	99.27 (1.59)***	96.84 (2.18)***
β_{1ij} Gender (girl)		2.32 (1.07)*	2.37 (1.07)*
β_{2ij} (Pretest score − 100)		0.28 (0.05)***	0.28 (0.05)***
β_{3j} EI + PA			6.43 (2.93)*
β_{4j} EI + IND			1.01 (3.56)
Random part			
σ^2_{u0} class-level variance	37.19 (13.64)**	39.72 (14.24)**	28.65 (10.80)**
Proportion of the variance at class level	23.92%	27.67%	21.63%
σ^2_{e0} student-level variance	118.27 (8.36)***	103.81 (7.42)***	103.79 (7.42)***
Proportion of the variance at student level	76.08%	72.33%	78.37%
Loglikelihood	3236.66	3117.65	3111.92
Reference Model		Model 0	Model 1

Note. EI + PA = Explicit writing instruction + peer-assisted writing. EI + IND = Explicit writing instruction + individual writing.
Standard error estimates are placed between brackets.
[a]Model equation with business as usual condition as reference condition as an example:
$y \sim N(X\beta, \Omega)$.
$y_{ij} = \beta_{0ij} + \beta_1\ \mathrm{Girl}_{ij} + \beta_2\ \mathrm{Pretest}(-100)_{ij} + \beta_3\ \mathrm{EI+PA}_j + \beta_4\ \mathrm{EI+IND}_j$
$\beta_{0ij} = \beta_0 + u_{0j} + e_{0ij}$
$[u_{0j}] \sim N(0, \Omega u)$: $\Omega_u = [\sigma 2_{u0}]$
$[e_{0ij}] \sim N(0, \Omega e)$: $\Omega e = [\sigma 2_{e0}]$
***p < .001. **p < .01. *p < .05.

from zero (i.e., descriptive text quality: $\chi^2(1) = 7.44$, $p < .01$; self-efficacy for ideation: $\chi^2(1) = 5.11$, $p < .05$; self-efficacy for conventions: $\chi^2(1) = 4.77$, $p < .05$; self-efficacy for regulation $\chi^2(1) = 4.27$, $p < .05$; controlled writing motivation: $\chi^2(1) = 4.90$, $p < .05$) or not significantly different from zero (students' narrative text quality: $\chi^2(1) = 3.56$, $p = .06$ and autonomous writing motivation: $\chi^2(1) = 3.47$, $p = .06$). Based on these results, it can be concluded that next to differences between students (see the variances at student level, σ^2_{e0}), the variability in all response variables, with the exception of students' narrative text quality and their autonomous motivation, could be attributed to differences between classes. Consequently, these results justify the use of multilevel analyses in the current study.

3.2.1. Quality of the descriptive text

After controlling for students' gender and pretest score, the results presented in Table 6 showed that EI + PA students statistically outperformed both EI + IND ($\chi^2(1) = 4.80$, $p < .05$, 0.51 *SD*) and BAU students ($\chi^2(1) = 4.44$, $p < .05$, 0.60 *SD*) at posttest. No significant differences were found between EI + IND and BAU students ($\chi^2(1) = 0.08$, $p = .78$).

3.2.2. Quality of the narrative text

After including students' gender, grade, home language, and pretest score as covariates, the results presented in Table 7 revealed no statistical differences between EI + PA and EI + IND students ($\chi^2(1) = 0.00$, $p = 1.00$), nor between the experimental students and the BAU students (EI + PA: $\chi^2(1) = 1.48$, $p = .22$ and EI + IND: $\chi^2(1) = 1.34$, $p = .25$).

3.2.3. Self-efficacy for ideation

After controlling for students' gender and pretest score, the multilevel results presented in Table 8 showed that EI + PA students were statistically more self-efficacious to come up with good ideas to write at posttest compared to their EI + IND counterparts ($\chi^2(1) = 4.20$, $p < .05$, 0.28 *SD*). No significant differences were found between the experimental and the BAU students (EI + PA: $\chi^2(1) = 0.88$, $p = .35$ and EI + IND: $\chi^2(1) = 0.62$, $p = .43$, respectively).

3.2.4. Self-efficacy for conventions

After including students' gender and pretest score as covariates, no statistical differences between EI + PA and EI + IND students ($\chi^2(1) = 2.19$, $p = .14$), nor between the experimental and the BAU students (EI + PA: $\chi^2(1) = 0.00$, $p = .97$ and EI + IND: $\chi^2(1) = 1.43$, $p = .23$) were found (see Table 9).

3.2.5. Self-efficacy for regulation

After controlling for students' gender, home language, and pretest score, no statistical differences between EI + PA and EI + IND students ($\chi^2(1) = 0.08$, $p = .78$), nor between the experimental and the BAU students (EI + PA: $\chi^2(1) = 2.30$, $p = .13$ and EI + IND: $\chi^2(1) = 1.43$, $p = .23$) were found (see Table 10).

3.2.6. Autonomous writing motivation

After controlling for students' gender, home language, and pretest score, no statistically significant differences between EI + PA and EI + IND students ($\chi^2(1) = 0.18$, $p = .67$), nor between the experimental students and the BAU students (EI + PA: $\chi^2(1) = 0.60$, $p = .44$ and EI + IND: $\chi^2(1) = 1.18$, $p = .28$) were found (see Table 11).

Table 7
Summary of the model estimates for the two-level analysis of students' narrative text quality at posttest.

	Model 0	Model 1	Model 2[a]
Fixed part			
β_{0ij} Intercept	96.21 (0.75)***	95.07 (1.20)***	93.44 (1.73)***
β_{1ij} Gender (girl)		1.76 (0.95)	1.73 (0.95)
β_{2ij} Grade (6th grade)		2.37 (1.33)	2.15 (1.30)
β_{3ij} Home language (foreign language)		−4.34 (1.56)**	−4.62 (1.59)**
β_{4ij} Home language (Dutch + foreign language		−1.95 (1.47)	−2.15 (1.48)
β_{5ij} (Pretest score – 96.2)		0.22 (0.04)***	0.22 (0.04)***
β_{6j} EI + PA			2.35 (1.94)
β_{7j} EI + IND			2.35 (2.03)
Random part			
σ^2_{u0} class-level variance	6.70 (3.55)	7.31 (3.66)	6.21 (3.32)
Proportion of the variance at class level	6.80%	8.29%	7.12%
σ^2_{e0} student-level variance	91.76 (6.52)***	80.95 (5.91)***	81.03 (5.92)***
Proportion of the variance at student level	93.20%	91.81%	92.88%
Loglikelihood	3078.79	2876.74	2875.10
Reference Model		Model 0	Model 1

Note. EI + PA = Explicit writing instruction + peer-assisted writing. EI + IND = Explicit writing instruction + individual writing.
Standard error estimates are placed between brackets.
[a]Model equation with business as usual condition as reference condition as an example:
$y \sim N(XB, \Omega)$
$y_{ij} = \beta_{0ij} + \beta_1\ \text{Girl}_{ij} + \beta_2\ \text{6th grade}_{ij} + \beta_3\ \text{Foreign language}_{ij} + \beta_4\ \text{Dutch + Foreign language}_{ij} + \beta_5\ \text{Pretest}(-96.2)_{ij} + \beta_6\ \text{EI + PA}_j + \beta_7\ \text{EI + IND}_j$
$\beta_{0ij} = \beta_0 + u_{0j} + e_{0ij}$
$[u_{0j}] \sim N(0,\Omega u)$: $\Omega_u = [\sigma 2_{u0}]$
$[e_{0ij}] \sim N(0,\Omega e)$: $\Omega e = [\sigma 2_{e0}]$
***p < .001. **p < .01. *p < .05.

Table 8
Summary of the model estimates for the two-level analysis of students' self-efficacy for ideation at posttest.

	Model 0	Model 1	Model 2[a]
Fixed part			
β_{0ij} Intercept	71.31 (1.82)***	71.23 (1.62)***	70.83 (2.79)***
β_{1ij} Gender (girl)		−0.30 (1.68)	−0.04 (1.68)
β_{2ij} (Pretest score – 68.2)		0.49 (0.04)***	0.49 (0.04)***
β_{3j} EI + PA			3.10 (3.31)
β_{4j} EI + IND			−2.68 (3.39)
Random part			
σ^2_{u0} class-level variance	47.19 (20.88)*	24.87 (12.31)*	17.15 (9.79)
Proportion of the variance at class level	10.77%	8.17%	5.77%
σ^2_{e0} student-level variance	390.87 (27.54)***	279.48 (19.80)***	279.93 (19.84)***
Proportion of the variance at student level	89.23%	91.83%	94.23%
Loglikelihood	3750.13	3561.59	3557.78
Reference Model		Model 0	Model 1

Note. EI + PA = Explicit writing instruction + peer-assisted writing. EI + IND = Explicit writing instruction + individual writing.
Standard error estimates are placed between brackets.
[a]Model equation with business as usual condition as reference condition as an example:
$y \sim N(XB, \Omega)$
$y_{ij} = \beta_{0ij} + \beta_1\ \text{Girl}_{ij} + \beta_2\ \text{Pretest(-68.2)}_{ij} + \beta_3\ \text{EI + PA}_j + \beta_4\ \text{EI + IND}_j$
$\beta_{0ij} = \beta_0 + u_{0j} + e_{0ij}$
$[u_{0j}] \sim N(0,\Omega u)$: $\Omega_u = [\sigma 2_{u0}]$
$[e_{0ij}] \sim N(0,\Omega e)$: $\Omega e = [\sigma 2_{e0}]$
***p < .001. **p < .01. *p < .05.

Table 9
Summary of the model estimates for the two-level analysis of students' self-efficacy for conventions at posttest.

	Model 0	Model 1	Model 2[a]
Fixed part			
β_{0ij} Intercept	75.72 (1.54)***	75.53 (1.55)***	76.99 (2.88)***
β_{1ij} Gender (girl)		−0.61 (1.44)	−0.44 (1.44)
β_{2ij} (Pretest score – 77.6)		0.58 (0.04)***	0.58 (0.04)***
β_{3j} EI + PA			0.10 (3.44)
β_{4j} EI + IND			−4.19 (3.51)
Random part			
σ^2_{u0} class-level variance	32.62 (14.94)*	27.09 (11.75)*	22.77 (10.48)*
Proportion of the variance at class level	9.75%	11.70%	10.02%
σ^2_{e0} student-level variance	301.95 (21.33)***	204.47 (14.54)***	204.48 (14.53)***
Proportion of the variance at student level	90.25%	88.30%	89.80%
Loglikelihood	3622.22	3420.00	3417.57
Reference Model		Model 0	Model 1

Note. EI + PA = Explicit writing instruction + peer-assisted writing. EI + IND = Explicit writing instruction + individual writing.
Standard error estimates are placed between brackets.
[a]Model equation with business as usual condition as reference condition as an example:
$y \sim N(XB, \Omega)$
$y_{ij} = \beta_{0ij} + \beta_1\ \text{Girl}_{ij} + \beta_2\ \text{Pretest(-77.6)}_{ij} + \beta_3\ \text{EI + PA}_j + \beta_4\ \text{EI + IND}_j$
$\beta_{0ij} = \beta_0 + u_{0j} + e_{0ij}$
$[u_{0j}] \sim N(0,\Omega u)$: $\Omega_u = [\sigma 2_{u0}]$
$[e_{0ij}] \sim N(0,\Omega e)$: $\Omega e = [\sigma 2_{e0}]$
***p < .001. **p < .01. *p < .05.

3.2.7. Controlled writing motivation

After including students' pretest score as covariates, the results presented in Table 12 indicated that EI + PA students had statistically lower scores for controlled motivation at posttest compared to EI + IND students ($\chi^2(1) = 7.50, p < .01, -0.30\ SD$). No statistical differences were found between experimental students and BAU students (EI + PA: $\chi^2(1) = 1.44$, $p = .23$ and EI + IND: $\chi^2(1) = 1.39$, $p = .24$, respectively).

3.3. Social validity of the interventions

Based on the teachers' logbooks we were able to gain insight into the experimental teachers' satisfaction with the EI + PA and EI + IND writing program. Table 13 shows that teachers were generally satisfied

Table 10
Summary of the model estimates for the two-level analysis of students' self-efficacy for regulation at posttest.

	Model 0	Model 1	Model 2[a]
Fixed part			
β_{0ij} Intercept	65.79 (1.81)***	66.11 (1.67)***	69.37 (2.67)***
β_{1ij} Gender (girl)		0.85 (1.86)	0.95 (1.87)
β_{2ij} Home language (foreign language)		−4.61 (2.86)	−4.19 (2.92)
β_{3ij} Home language (Dutch + foreign language		−0.80 (2.83)	−0.34 (2.83)
β_{4ij} (Pretest score – 69.1)		0.61 (0.05)***	0.61 (0.05)***
β_{5j} EI + PA			−4.77 (3.15)
β_{6j} EI + IND			−3.98 (3.32)
Random part			
σ^2_{u0} class-level variance	42.74 (20.70)*	15.50 (10.04)	11.88 (8.88)
Proportion of the variance at class level	8.44%	4.54%	3.52%
σ^2_{e0} student-level variance	463.73 (32.79)***	325.77 (23.26)***	325.91 (23.27)***
Proportion of the variance at student level	91.56%	95.46%	96.48%
Loglikelihood	3791.73	3566.66	3564.40
Reference Model		Model 0	Model 1

Note. EI + PA = Explicit writing instruction + peer-assisted writing. EI + IND = Explicit writing instruction + individual writing.
Standard error estimates are placed between brackets.
[a]Model equation with business as usual condition as reference condition as an example:
$y \sim N(XB, \Omega)$
$y_{ij} = \beta_{0ij} + \beta_1\ \text{Girl}_{ij} + \beta_2\ \text{Foreign language}_{ij} + \beta_3\ \text{Dutch + Foreign language}_{ij} + \beta_4\ \text{Pretest(-69.1)}_{ij} + \beta_5\ \text{EI + PA}_j + \beta_6\ \text{EI + IND}_j$
$\beta_{0ij} = \beta_0 + u_{0j} + e_{0ij}$
$[u_{0j}] \sim N(0,\Omega u)$: $\Omega_u = [\sigma 2_{\ u0}]$
$[e_{0ij}] \sim N(0,\Omega e)$: $\Omega e = [\sigma 2_{e0}]$
***p < .001. **p < .01. *p < .05

with the provided instructional materials and lessons to foster students' writing. There were no significant differences between the conditions (attainment lesson objectives: $t(12) = 0.54$, $p = .60$; evaluation instructional materials: $t(12) = 0.71$, $p = .49$; value of the lessons: $t(12) = 1.28$, $p = .23$; ease of implementation: $t(12) = 0.90$, $p = .39$; level of difficulty for students: $t(12) = 0.99$, $p = .34$).

4. Discussion

4.1. The impact of the EI + PA and EI + IND program on students' writing performance

4.1.1. The impact on descriptive text quality

As predicted, the results showed that students writing with a peer outperformed both students practicing individually and the business as usual students in terms of the quality of their descriptive text. These findings confirm the basic tenets of the social cognitive view on writing by highlighting the pivotal role of observational learning during planning, composing, and writing with a peer (Zimmerman & Risemberg,

Table 11
Summary of the model estimates for the two-level analysis of students' autonomous writing motivation at posttest.

	Model 0	Model 1	Model 2[a]
Fixed part			
β_{0ij} Intercept	3.12 (0.07)***	3.04 (0.07)***	3.14 (0.12)***
β_{1ij} Gender (girl)		0.03 (0.07)	0.04 (0.07)
β_{2ij} Home language (foreign language)		−0.10 (0.11)	−0.08 (0.12)
β_{3ij} Home language (Dutch + foreign language		0.04 (0.11)	0.05 (0.11)
β_{4ij} (Pretest score – 3.21)		0.70 (0.04)***	0.69 (0.04)***
β_{5j} EI + PA			−0.11 (0.14)
β_{6j} EI + IND			−0.16 (0.15)
Random part			
σ^2_{u0} class-level variance	0.06 (0.03)	0.03 (0.02)	0.03 (0.02)
Proportion of the variance at class level	6.32%	6.52%	6.52%
σ^2_{e0} student-level variance	0.89 (0.06)***	0.43 (0.03)***	0.43 (0.03)***
Proportion of the variance at student level	93.68%	93.48%	93.48%
Loglikelihood	1142.61	791.25	790.11
Reference Model		Model 0	Model 1

Note. EI + PA = Explicit writing instruction + peer-assisted writing. EI + IND = Explicit writing instruction + individual writing.
Standard error estimates are placed between brackets.
[a]Model equation with business as usual condition as reference condition as an example:
$y \sim N(XB, \Omega)$
$y_{ij} = \beta_{0ij} + \beta_1\ \text{Girl}_{ij} + \beta_2\ \text{Foreign language}_{ij} + \beta_3\ \text{Dutch + Foreign language}_{ij} + \beta_4\ \text{Pretest(-3.21)}_{ij} + \beta_5\ \text{EI + PA}_j + \beta_6\ \text{EI + IND}_j$
$\beta_{0ij} = \beta_0 + u_{0j} + e_{0ij}$
$[u_{0j}] \sim N(0,\Omega u)$: $\Omega_u = [\sigma 2_{\ u0}]$
$[e_{0ij}] \sim N(0,\Omega e)$: $\Omega e = [\sigma 2_{e0}]$
***p < .001. **p < .01. *p < .05.

Table 12
Summary of the model estimates for the two-level analysis of students' controlled writing motivation at posttest.

	Model 0	Model 1	Model 2[a]
Fixed part			
β_{0ij} Intercept	2.73 (0.07)***	2.31 (0.06)***	2.14 (0.09)***
β_{1ij} (Pretest score – 2.77)		0.54 (0.05)***	0.53 (0.05)***
β_{2j} EI + PA			−0.12 (0.10)
β_{3j} EI + IND			0.12 (0.10)
Random part			
σ^2_{u0} class-level variance	0.07 (0.03)*	0.01 (0.01)	0.00 (0.00)
Proportion of the variance at class level	11.11%	2.17%	0.00%
σ^2_{e0} student-level variance	0.56 (0.04)***	0.45 (0.03)***	0.45 (0.03)***
Proportion of the variance at student level	88.89%	97.83%	100.00%
Loglikelihood	901.54	746.17	739.50
Reference Model		Model 0	Model 1

Note. EI + PA = Explicit writing instruction + peer-assisted writing. EI + IND = Explicit writing instruction + individual writing.
Standard error estimates are placed between brackets. ***p < .001. **p < .01. *p < .05.
[a]Model equation with business as usual condition as reference condition as an example:
$y \sim N(XB, \Omega)$
$y_{ij} = \beta_{0ij} + \beta_1 \text{ Pretest}(-2.77)_{ij} + \beta_2 \text{ EI + PA}_j + \beta_3 \text{ EI + IND}_j$
$\beta_{0ij} = \beta_0 + u_{0j} + e_{0ij}$
$[u_{0j}] \sim N(0, \Omega u)$: $\Omega_u = [\sigma 2_{u0}]$
$[e_{0ij}] \sim N(0, \Omega e)$: $\Omega e = [\sigma 2_{e0}]$
***p < .001. **p < .01. *p < .05.

1997). Furthermore, the results corroborate previous empirical studies that revealed the incremental effect of peer-assistance in explicit writing instruction programs (Graham et al., 2005; Harris et al., 2006). The present study, however, goes beyond these prior studies by investigating the effect of 'a structured system of peer-assisted writing'. Whereas the operationalization of peer assistance in previous studies was at a rather basic level, such as peer support to promote transfer of writing strategies (Graham et al., 2005; Harris et al., 2006), peer assistance to support the acquisition of genre knowledge (Hoogeveen & van Gelderen, 2018), or unprescribed peer collaboration (De Smedt & Van Keer, 2018b), the present study provides more in-depth insight into possible essential prerequisites for 'a structured system of peer assistance' to be powerful and effective (Daiute & Dalton, 1993).

Our findings and approach provide support for the importance of *promoting mutual engagement* between writing partners (Daiute & Dalton, 1993). In the present study, this was done by taking into account students' matching personalities when grouping them into dyads and by including a discussion on collaboration rules that the dyad had to agree to. Second, in line with research results on more unprescribed applications of peer collaboration (De Smedt & Van Keer, 2018b), the present study highlights the need to *structure the collaboration* to support students in how to approach this collaboration (Yarrow & Topping, 2001). In the current EI + PA program, peer assistance was more directly structured by means of role assignment and the use of shared writing portfolios. The inclusion of the roles (i.e., the thinker, the strategy card reader, and the reporter) supported students in their strategic writing behavior, as they learned what kind of behaviors were expected during each step of the writing process (Yarrow & Topping, 2001). Moreover, the use of shared writing portfolios clearly directed students to work on shared documents instead of merging individual texts. Third, in line with SCT (Bandura, 1986, 1989) the need to *model the collaboration* was emphasized. More particularly, EI + PA teachers and students modelled how to approach the collaboration by demonstrating the use of the roles and by modeling appropriate collaboration and interaction skills, such as negotiating, compromising, or active listening. In this way, EI + PA students did not only acquire cognitive writing strategies by observing teacher and peer models, but also learned how effective collaboration and interaction with a peer during writing helped them in becoming better writers. Finally, the present study emphasizes the need to *create a 'share-your-writing-culture'* in the writing classroom. This view on writing as a social and shared activity, rather than as an individual and solitary activity, empirically confirms the emphasis on social learning processes in writing development according to Zimmerman and Risemberg (1997). In view of this final prerequisite for effective collaboration, the EI + PA teachers provided collaboration opportunities between the dyads by sharing their texts with other dyads, who provided feedback, which in turn helped students to improve their work. In this way, EI + PA students could observe how their peers read, perceived, and experienced their texts and this reader/peer feedback could help them to enhance their texts (Couzijn, 1999).

Although this study provided in-depth insights into possible crucial components of a powerful and structured application of peer-assisted writing, the present research design did not enable us to identify which component(s) account for the overall effectiveness of the EI + PA program. In this respect, future studies should conduct component analyses to determine the differential effectiveness of these and other aspects of the intervention to uncover the most powerful ingredients of peer-assisted writing (e.g., Fidalgo et al., 2015; López, Torrance, Rijlaarsdam, & Fidalgo, 2017).

Contrary to our hypotheses, EI + IND students did not outperform BAU students at posttest. Previous research, however, repeatedly revealed the effectiveness of explicitly teaching students writing knowledge and strategies (Graham et al., 2012; Koster et al., 2015). These rather unexpected results, however, confirm the added value of peer-assistance in time-based explicit writing instruction programs. More particularly, in time-based approaches students are required to develop

Table 13
Social validity: experimental teachers' satisfaction with the EI + PA and EI + IND writing program.

	M (SD)		
	EI + PA	EI + IND	All experimental conditions
Attainment of the lesson objectives[a]	4.10 (0.29)	4.21 (0.47)	4.15 (0.37)
Evaluation of the instructional materials[b]	4.66 (0.51)	4.81 (0.21)	4.72 (0.40)
Value of the lessons to enhance students' writing[c]	8.76 (0.47)	9.12 (0.59)	8.91 (0.54)
Level of ease to implement the lessons[d]	8.15 (0.50)	8.48 (0.84)	8.29 (0.66)
Level of difficulty of the lessons for the students[e]	7.49 (1.05)	8.51 (2.67)	7.93 (1.90)

Note. EI + PA = Explicit writing instruction + peer-assisted writing. EI + IND = Explicit writing instruction + individual writing.
[a] Measured on a 5-point Likert scale ranging from 'not attained' to 'fully attained'.
[b] Measured on a 5-point Likert scale ranging from 'very unclear' to 'very clear'.
[c] Measured on a 10-point scale ranging from 'very invaluable to teach stimulate students' writing' to 'very valuable to stimulate students' writing'.
[d] Measured on a 10-point scale ranging from 'too difficult to implement in the class' to 'very easy to implement in class'.
[e] Measured on a 10-point scale ranging from 'students experienced a lot of difficulties' to 'student experienced no difficulties'.

knowledge and skills within a fixed timeframe (e.g., Koster & Bouwer, 2018). On the contrary, in mastery-based approaches teachers provide instruction until students master the required knowledge and skills (e.g., Graham & Harris, 2018). In the current time-based approach of the EI + PA and EI + IND interventions (i.e., 11 lessons within 10 weeks), EI + IND students perhaps struggled to internalize all the knowledge and strategies taught during individual practice. Consequently, as EI + IND students potentially could not fully benefit from the knowledge and strategies taught within these ten weeks, they did not outperform BAU students. EI + PA students, however, might had the advantage of observing a peer and collaborating with each other during this complex writing process. In line with the prominent role of peer learning in SCT (Bandura, 1986, 1989), EI + PA students were probably better supported in internalizing the knowledge and strategies within this limited timeframe, resulting in an overall better performance at posttest.

4.1.2. The impact on narrative text quality

In line with our hypothesis regarding the absence of a transfer effect of the experimental interventions to an uninstructed writing genre, there were no significant differences between EI + PA, EI + IND, and BAU students on the quality of students' narrative writing. In contrast to the prior studies of Graham et al. (2005) and Harris et al. (2006) in which peer support was specifically designed to promote and facilitate the transfer of strategies to other situations and contexts, the present study examined whether spontaneous transfer of the strategies occurred. In this respect, the EI + PA and EI + IND teachers focused solely on the descriptive text genre without making any references to other writing genres or without discussing how the general writing strategies could be applied in other contexts. Based on the basic premises of the social cognitive view on writing (Zimmerman & Risemberg, 1997) and the results of the previous studies of Graham et al. (2005) and Harris et al. (2006) on the one hand and the present study on the other hand, it is reasonable to argue that to successfully transfer writing knowledge and strategies to uninstructed genres, students need explicit support on how to do so. Teachers can, for instance, discuss and model how to apply or adapt the writing strategies in uninstructed genres and provide students opportunities to practice these (adjusted) strategies. Additional research is needed to test our assertion.

4.2. The impact of the EI + PA and EI + IND program on motivational aspects

4.2.1. The impact on self-efficacy for writing

As expected the results revealed that EI + PA students felt more self-efficacious to invent ideas compared to their EI + IND counterparts, whereas no significant differences were found between EI + PA and BAU students. These findings support our hypothesis that peer assistance augments learning in an explicit writing program in terms of fostering students' perceived competence to come up with ideas. A possible explanation for our finding might be related to the inclusion of the 'thinker-role' that was a fixed role for every writing member during each lesson. According to this role, each student received responsibility to generate ideas and thereby provide input for the planning, composition, and revision of the text. During modeling, the teacher also reminded the students of their responsibility as 'a thinker' by encouraging all students to actively think about ideas and thereby contributing to the shared writing product. Next to teacher modeling, the 'thinker role' was also modeled by the group members during writing practice. In this respect, peer modeling might also play an important role in fostering students' self-efficacy for ideation (Zimmerman & Risemberg, 1997).

Although previous research on peer-assisted writing in elementary grades focused on structuring peer collaboration by means of tutor and tutee roles (e.g., Sutherland & Topping, 1999; Yarrow & Topping, 2001), the present study is the first to investigate how these roles can augment explicit writing instruction. Future research should study more in-depth to what extent the inclusion and explicit modeling of roles in peer-assisted writing can possibly affect students' interactions within writing groups. More particularly, how do these roles manifest during students' interactions and do these roles possibly create more in-depth interactions on the writing process and product between writing partners? By studying students' interactions, we can gain a deeper understanding of the effectiveness of peer-assisted writing by uncovering (in) effective interaction or learning processes that cannot be revealed by solely studying students' writing products (e.g., Allal, 2018; Herder, Berenst, de Glopper, & Koole, 2018).

Contrary to our hypotheses, we did not find any significant differences between EI + PA, EI + IND, and BAU students on their self-efficacy for conventions and regulation. These results suggest that there was no distinct nor combined effect of explicit writing instruction and peer-assisted writing on students' perceived capabilities to adhere to language and writing rules nor to regulate their behavior during writing practice. Both EI + PA and EI + IND programs focused on explicitly teaching students writing knowledge (i.e., text structure and genre knowledge) and writing strategies (i.e., planning, composing, and revising strategies). In this respect, the non-significant results for students' self-efficacy for conventions can be explained by the fact that teaching students basic writing skills, such as spelling or handwriting, was not the focus of both programs. During instruction and modeling, EI + PA and EI + IND teachers primarily focused on correct strategy use and writing knowledge instead of stressing the need to apply writing rules accurately. This might have affected students' impression that during EI + PA or EI + IND lessons their knowledge and capabilities on applying correct language and writing rules did not increase.

With regard to students' self-efficacy for regulation, in contrast to for instance SRSD (Graham et al., 2005; Harris et al., 2006), the EI + PA and EI + IND programs did not focus on explicitly teaching and modeling self-regulated strategies so students could learn how to regulate their behavior during writing. EI + PA and EI + IND students were not explicitly taught strategies, for example, to assess their own writing process or to motivate themselves to persist during writing (Zimmerman & Risemberg, 1997). The lack hereof might have influenced EI + PA and EI + IND students' impression that they did not move forward regarding these self-regulation skills.

Overall, the present findings are inconsistent with the results of earlier studies showing no augmented effect of peer-assistance on students' self-efficacy for writing (De Smedt, Graham, & Van Keer, 2018; Graham, Harris, & Mason, 2005; Harris, Graham, & Mason, 2006). Within this limited number of previous studies, however, only the study of De Smedt, Graham, and Van Keer (2018) also examined self-efficacy for writing as a multidimensional construct (Bruning et al., 2013). Again, due to the limited research base, no valid conclusions can be drawn regarding the added value of peer-assistance in explicit writing instruction programs on the different dimensions of students' self-efficacy (i.e., ideation, conventions, regulation). Future studies should replicate the current study by taking into account the multidimensional character of self-efficacy for writing and extend the current study by: (a) investigating the effect of other writing programs, such as for instance SRSD, on students self-efficacy for ideation, conventions, and regulation, and (b) including genre-specific measures on self-efficacy for writing instead of a general measure across writing genres.

4.2.2. The impact on writing motivation

Contrary to our hypotheses, the results showed no significant differences between EI + PA, EI + IND, and BAU students' autonomous writing motivation nor an augmented effect of peer-assistance in the experimental conditions. Moreover, EI + PA students evidenced significantly less controlled motivation at posttest compared to EI + IND students, whereas no significant differences were found between EI + PA and BAU students. These results indicated that EI + IND students experienced higher levels of controlled motivation, that is, the type of motivation that is associated with external (e.g., grades) or internal

concerns or pressure (e.g., guilt) compared to their EI + PA counterparts. These results are in line with a previous study of De Smedt, Graham, and Van Keer (2018) showing that EI + IND students reported higher controlled writing motivation as compared to students in an individual matched practice condition who did not receive any type of explicit or implicit instruction.

According to SCT, students function as active agents in their own motivation. More particularly, students' ability to evaluate their own attainments and their perceived self-efficacy to fulfill one's standards provide a major cognitive mechanism of motivation (Bandura, 1986, 1989). In this way, accurate self-assessment can lead to either satisfaction (i.e., positive evaluation of a good performance) or action to intensify efforts (i.e., negative evaluation of a substandard performance). Also, students with low self-efficacy beliefs are easily dissuaded by failure, while students with high self-efficacy beliefs are prompted to persist until they succeed. Taking SCT into account, we offer two possible explanations for the lack of motivational growth in the EI + PA condition. First, as mentioned above, EI + PA students' overall self-efficacy for writing, which plays an influential role in students' motivation, did not increase after the intervention. Notwithstanding the fact that EI + PA students felt more self-efficacious to invent ideas, their self-efficacy for conventions and regulations was not enhanced. These findings indicate that students might experience writing as a complex activity involving cognitive, linguistic, and self-regulatory processes. Enhanced self-efficacy in one of these processes (i.e., ideation), and not in the other processes (i.e., conventions and regulation), did not result in increased writing motivation. These findings might indicate that, for students to become more motivated writers, their overall sense of self-efficacy for all processes underlying writing needs to be enhanced. Second, according to Bandura (1989), students show long-lasting motivation in activities at which they feel self-efficacious and from which they derive satisfaction. The process of promoting students' self-efficacy and building enduring motivation by means of performance accomplishments, vicarious experience, verbal persuasion, and emotional arousal takes time (Bandura, 1986, 1989). In this respect, the current time-based approach of the EI + PA intervention (i.e., 11 lessons within 10 weeks), was perhaps too limited to foster students' long-lasting writing motivation.

4.3. Limitations and suggestions for future research

In addition to the limitations and suggestions for future research already discussed above, we conclude with some final thoughts on limitations related to measurement issues and how these can be addressed in future research. First, students' writing performance in both writing genres was measured using only one test per genre. Such a test indicates students' writing performance at a given moment in time. To assess students' writing proficiency in a more valid and reliable way, multiple writing tests per genre should be administered (Bouwer, Béguin, Sanders, & van den Bergh, 2015). Such large data collections are, however, very labor-intensive as all texts need to be corrected and typed to avoid presentation effects (Graham et al., 2011), genre-specific benchmark scales need to be developed (Bouwer et al., 2015), and different panels of raters are required to assess the large number of texts (Bouwer et al., 2018). Within the scope of the present study, such large data collections were unfortunately not feasible.

In addition to the writing measurement issues noted above, we also want to draw attention to the limitations of using questionnaires to measure students' self-efficacy for writing and writing motivation. A well-known drawback related to self-report is the elicitation of socially desirable responses from students (Schellings & Van Hout-Wolters, 2011). Additionally, young students also might overestimate their capabilities when completing the self-efficacy questionnaire (Klassen, 2002). As a result, students' self-reported mean scores on self-efficacy for writing and writing motivation tend to be higher than expected at pretest. In this respect, it is particularly challenging to uncover a significant growth from pretest to posttest regarding students' self-reported self-efficacy and writing motivation. This can be especially difficult when taking into account that experimental students, in contrast to BAU students, might assess their writing capabilities more realistically at posttest due to the EI + PA or EI + IND program possibly affecting students' self-reflection and awareness regarding their writing capacities. Based on these limitations, we recommend that self-report data be complimented with other data, such as observational data or data retrieved from conversational interviews in which the researcher tries to reveal students' motives when writing particular tasks (e.g., Dowson & McInerney, 2003).

References

Abbuhl, R. (2011). Using models in writing instruction: A comparison with native and nonnative speakers of English. *SAGE Open, 1*(3), 1–12. https://doi.org/10.1177/2158244011426295.

Alexander, P., & Jetton, T. (2000). Learning from text: A multidimensional and developmental perspective. In M. Kamil, P. Mosenthal, P. Pearson, & R. Barr (Eds.). *Handbook of reading research* (pp. 285–310). Mahwah, New Jersey: Lawrence Erlbaum Associates Inc.

Allal, L. (2018). The co-regulation of writing activities in the classroom. *Journal of Writing Research, 10*(1), 25–60. https://doi.org/10.17239/jowr-2018.10.01.02.

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. 191-125 *Psychological Review, 84*(2), https://doi.org/10.1037/0033-295X.84.2.191.

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory.* Englewood Cliffs, NJ: Prentice-Hall.

Bandura, A. (1989). Social cognitive theory. In R. Vasta (Ed.). *Annals of child development* (pp. 1–60). Greenwich, CT: JAI Press.

Bandura, A. (1997). *Self-efficacy: The exercise of control.* New York, NY: Freeman.

Bean, T., & Steenwyk, F. (1984). The effect of three forms of summarization instruction on sixth graders' summary writing and comprehension. *Journal of Reading Behavior, 16*(4), 297–306. https://doi.org/10.1080/10862968409547523.

Bentler, P. (2009). Alpha, dimension-free, and model-based internal consistency reliability. *Psychometrika, 74*(1), 137–143. https://doi.org/10.1007/S11336-008-9100-1.

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

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

Bouwer, R., Lesterhuis, M., De Smedt, F., Van Keer, H., & De Maeyer, S. (2019). A comparative approach to the assessment of writing: Benefits of benchmark ratings and comparative judgement. *Manuscript submitted for publication*.

Bouwer, R., Koster, M., & van den Bergh, H. (2016). Benchmark rating procedure, best of both worlds? Comparing procedures to rate text quality in a reliable and valid manner. Bringing writing research into the classroom. The effectiveness of Tekster, a newly developed writing program for elementary students (Doctoral Dissertation). Utrecht University. Utrecht, the Netherlands.

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

Bruning, R., Dempsey, M., Kauffman, D., McKim, C., & Zumbrunn, S. (2013). Examining dimensions of self-efficacy for writing. *Journal of Educational Psychology, 105*(1), 25–38. https://doi.org/10.1037/a0029692.

Bruning, R., & Kauffman, D. (2016). Self-efficacy beliefs and motivation in writing development. In C. MacArthur, S. Graham, & J. Fitzgerald (Eds.). *Handbook of Writing Research* (pp. 160–173). New York: The Guilford Press.

Cameron, C., & Moshenko, B. (1996). Elicitation of knowledge transformational reports while children write narratives. *Canadian Journal of Behavioural Science, 28*(4), 271–280. https://doi.org/10.1037/0008-400X.28.4.271.

Cleary, M. (1991). Affect and cognition in the writing processes of eleventh graders: A study of concentration and motivation. *Written Communication, 8*(4), 473–508. https://doi.org/10.1177/0741088391008004003.

Cohen, J. (1977). *Statistical power analysis for the behavioral sciences.* New York: Academic Press.

Couzijn, M. (1999). Learning to write by observation of writing and reading processes: Effects on learning and transfer. *Learning and Instruction, 9*(2), 109–142. https://doi.org/10.1016/S0959-4752(98)00040-1.

Daiute, C., & Dalton, B. (1993). Collaboration between children learning to write. Can novices be masters? *Cognition and Instruction, 10*(4), 281–333. https://doi.org/10.1207/s1532690xci1004_1.

De Naeghel, J., Van Keer, H., Vansteenkiste, M., & Rosseel, Y. (2012). The relation between elementary students' recreational and academic reading motivation, reading frequency, engagement, and comprehension: A self-determination theory perspective. *Journal of Educational Psychology, 104*(4), 1006–1021. https://doi.org/10.1037/a0027800.

De Smedt, F., Graham, S., & Van Keer, H. (2018). The bright and dark side of writing motivation: Effects of explicit instruction and peer assistance. *The Journal of Educational Research, 112*(2), 152–167. https://doi.org/10.1080/00220671.2018.1461598.

De Smedt, F., Merchie, E., Barendse, M., Rosseel, Y., De Naeghel, J., & Van Keer, H. (2018). Cognitive and motivational challenges in writing: Studying the relationship with writing performance across students' gender and achievement level. *Reading Research Quarterly, 53*(2), 249–272. https://doi.org/10.1002/rrq.193.

De Smedt, F., Rogiers, A., Heirweg, S., Vandamme, S., Merchie, E., & Van Keer, H. (2019). Reading and writing motivation across grades in primary and secondary education: A self-determination theory perspective. *Submitted for publication.*

De Smedt, F., & Van Keer, H. (2018a). An analytic description of an instructional writing program combining explicit writing instruction and peer-assisted writing. *Journal of Writing Research, 10*(2), 227–279. https://doi.org/10.17239/jowr-2018.10.01.04.

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

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

Dowson, M., & McInerney, D. (2003). What do students say about their motivational goals?: Towards a more complex and dynamic perspective on student motivation. *Contemporary Educational Psychology, 28*(1), 91–113. https://doi.org/10.1016/S0361-476X(02)00010-3.

Ferretti, R., & Lewis, W. (2013). Best practices in teaching argumentative writing. In S. Graham, C. A. MacArthur, & J. Fitzgerald (Eds.), Best practices in writing instruction (Second Edition ed., pp. 113-140). New York: The Guilford Press.

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

Fitzgerald, J., & Teasley, A. (1986). Effects of instruction in narrative structure on children's writing. *Journal of Educational Psychology, 78*(6), 424–432. https://doi.org/10.1037/0022-0663.78.6.424.

Flemish Ministry of Education and Training (2005). Education in Flanders. The Flemish educational landscape in a nutshell. Retrieved from https://planipolis.iiep.unesco.org/sites/planipolis/files/ressources/belgium_flanders_education_nutshell.pdf.

Flower, L., & Hayes, J. (1981). A cognitive process theory of writing. *College Composition and Communication, 32*(4), 365–387. https://doi.org/10.2307/356600.

Goldstein, H. (1995). *Multilevel statistical models.* London: Edward Arnold.

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

Graham, S. (2018). A writer(s) within community model of writing. In C. Bazerman, V. Berninger, S. Graham, J. Langer, S. Murphy, P. Matsuda, D. Rowe, & M. Schleppegrell (Eds.). *The lifespan development of writing* (pp. 271–325). Urbana, IL: National Council of English.

Graham, S., Berninger, V., & Fan, W. (2007). The structural relationship between writing attitude and writing achievement in first and third grade students. *Contemporary Educational Psychology, 32*(3), 516–536. https://doi.org/10.1016/j.cedpsych.2007.01.002.

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

Graham, S., & Harris, K. (2018). An examination of the design principles underlying a Self-Regulated Strategy Development study. *Journal of Writing Research, 10*(2), https://doi.org/10.17239/jowr-2018.10.02.02.

Graham, S., Harris, K., Fink, B., & MacArthur, C. (2001). Teacher efficacy in writing: A construct validation with primary grade teachers. *Scientific Studies of Reading, 5*(2), 177–202. https://doi.org/10.1207/S1532799Xssr0502_3.

Graham, S., Harris, K., & Hebert, M. (2011). It is more than just the message: Analysis of presentation effects in scoring writing. *Focus on Exceptional Children, 44*(4), 1–12.

Graham, S., Harris, K., & 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. https://doi.org/10.1016/j.cedpsych.2004.08.001.

Graham, S., McKeown, D., Kiuhara, S., & Harris, K. (2012). A meta-analysis of writing instruction for students in the elementary grades. *Journal of Educational Psychology, 104*(3), 879–896. https://doi.org/10.1037/A0029185.

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

Grünke, M., Wilbert, J., Tsiriotakis, I., & Agirregoikoa, A. (2017). Improving the length and quality of texts written by fourth graders with learning dissabilities through a peer-tutoring graphic organizing strategy. *Insights into Learning Disabilities, 14*(2), 167–188.

Harris, K., Graham, S., & Mason, L. (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.

Herder, A., Berenst, J., de Glopper, K., & Koole, T. (2018). Nature and function of proposals in collaborative writing of primary school students. *Linguistics and Education, 46*, 1–11. https://doi.org/10.1016/j.linged.2018.04.005.

Hidi, S., & Boscolo, P. (2006). Motivation and writing. In C. A. MacArthur, S. Graham, & J. Fitzgerald (Eds.). *Handbook of writing research* (pp. 144–157). New York: The Guilford Press.

Holliway, D. (2004). Through the eyes of my reader: A strategy for improving audience perspective in children's descriptive writing. *Journal of Research in Childhood Education, 18*(4), 334–349. https://doi.org/10.1080/02568540409595045.

Hoogeveen, M., & van Gelderen, A. (2018). Writing with peer response using different types of genre knowledge: Effects on linguistic features and revisions of sixth-grade writers. *The Journal of Educational Research, 111*(1), 66–80. https://doi.org/10.1080/00220671.2016.1190913.

Jones, K., & Duncan, C. (Eds.). (1998). *Modeling context and heterogeneity: Applying multilevel models*. Oxford, United Kingdom: Oxford University Press.

Klassen, R. (2002). A question of calibration: A review of the self-efficacy beliefs of students with learning disabilities. *Learning Disability Quarterly, 25*(2), 88–102. https://doi.org/10.2307/1511276.

Koster, M., & Bouwer, R. (2018). Describing multifaceted writing interventions: From design principles for the focus and mode of instruction to student and teacher activities. *Journal of Writing Research, 10*(2), https://doi.org/10.17239/jowr-2018.10.02.03.

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

Lesterhuis, M., Verhavert, S., Coertjens, L., Donche, V., & De Maeyer, S. (2017). Comparative judgement as a promising alternative to score competences. In E. Cano, & G. Ion (Eds.). *Innovative Practices for Higher Education Assessment and Measurement.* Hershey: PA: IGI Global.

Limpo, T., & Alves, R. (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.

López, P., Torrance, M., Rijlaarsdam, G., & Fidalgo, R. (2017). Effects of direct instruction and strategy modeling on upper-primary students' writing development. *Frontiers in Pyschology, 8*(1054), https://doi.org/10.3389/fpsyg.2017.01054.

McCutchen, D. (2008). Cognitive factors in the development of children's writing. In C. MacArthur, S. Graham, & J. Fitzgerald (Eds.). *Handbook of writing research* (pp. 115–130). New York: The Guilford Press.

McCutchen, D., Covill, A., Hoyne, S., & Mildes, K. (1994). Individual differences in writing: Implications of translating fluency. *Journal of Educational Psychology, 86*(2), 256–266. https://doi.org/10.1037/0022-0663.86.2.256.

McCutchen, D., Francis, M., & Kerr, S. (1997). Revising for meaning, effects of knowledge and strategy. *Journal of Educational Psychology, 89*(4), 667–676 10.1037/0022-0663.89.4.667.

McKeown, D., Brindle, M., Harris, K., Graham, S., Collins, A., & Brown, M. (2016). Illuminating growth and struggles using mixed methods: Practice-based professional development and coaching for differentiating SRSD instruction in writing. *Reading and Writing, 29*(6), 1105–1140. https://doi.org/10.1007/s11145-016-9627-y.

National Center for Education Statistics (2012). *The nation's report card: Writing 2011.* Washington, D.C.: Institute of Education Sciences, U.S. Department of Education.

Pajares, F., & Valiante, G. (1997). Influence of self-efficacy on elementary students' writing. *The Journal of Educational Research, 90*(6), 353–360. https://doi.org/10.1080/00220671.1997.10544593.

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

Paquette, K. (2009). Integrating the 6 + 1 writing traits model with cross-age tutoring: An investigation of elementary students' writing development. *Literacy Research and Instruction, 48*(1), 28–38. https://doi.org/10.1080/19388070802226261.

Philippakos, Z. (2017). Giving feedback: Preparing students for peer review and self-evaluation. *The Reading Teacher, 71*(1), 13–22. https://doi.org/10.1002/trtr.1568.

Pollitt, A. (2012). The method of adaptive comparative judgement. *Assessment in Education: Principles, Policy & Practice, 19*(3), 281–300. https://doi.org/10.1080/0969594X.2012.665354.

Rasbash, J., Charlton, C., Browne, W., Healy, M., & Cameron, B. (2009). MLwiN Version 2.1. Centre for Multilevel Modelling: University of Bristol.

Rasbash, J., Steele, F., Browne, W., & Goldstein, H. (2012). A User's Guide to MLwiN, v2. 26. Retrieved from Centre for Multilevel Modelling.

Rhoads, C. (2011). The implications of "contamination" for experimental design in education. *Journal of Educational and Behavioral Studies, 36*(1), 76–104. https://doi.org/10.3102/1076998610379133.

Rietdijk, S., Janssen, T., van Weijen, D., van den Bergh, H., & Rijlaarsdam, G. (2017). Improving writing in primary schools through a comprehensive writing program. *Journal of Writing Research, 9*(5), 173–225. https://doi.org/10.17239/jowr-2017.09.02.04.

Rietdijk, S., Van Weijen, D., Janssen, T., van den Bergh, H., & Rijlaarsdam, G. (2018). Teaching writing in primary education: Classroom practice, time, teachers' beliefs and skills. *Journal of Educational Psychology, 110*(2), 640–663. https://doi.org/10.1037/edu0000237.

Rogers, L., & Graham, S. (2008). A meta-analysis of single subject design writing intervention research. *Journal of Educational Psychology, 100*(4), 879–906. https://doi.org/10.1037/0022-0663.100.4.879.

Ryan, R., & Deci, E. (2000). Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary Educational Psychology, 25*(1), 54–67. https://doi.org/10.1006/ceps.1999.1020.

Schellings, G., & Van Hout-Wolters, B. (2011). Measuring strategy use with self-report instruments: Theoretical and empirical considerations. *Metacognition Learning, 6*(2), 83–90. https://doi.org/10.1007/s11409-011-9081-9.

Schoonen, R. (2005). Generalizability of writing scores: An application of structural equation modeling. *Language Testing, 22*(1), 1–30. https://doi.org/10.1191/0265532205lt295oa.

Snijders, T., & Bosker, R. (1999). *Multilevel analysis. An introduction to basic and advanced modeling.* London: Sage.

Sutherland, J., & Topping, K. (1999). Collaborative creative writing in eight-year-olds: Comparing cross-ability fixed role and same-ability reciprocal role pairing. *Journal of Research in Reading, 22*(2), 154–179. https://doi.org/10.1111/1467-9817.00080.

Tillema, M., van den Bergh, H., Rijlaarsdam, G., & Sanders, T. (2012). Quantifying the quality difference between L1 and L2 essays: A rating procedure with bilingual raters and L1 and L2 benchmark essays. *Language Testing, 30*(1), 71–97. https://doi.org/10.1177/0265532212442647.

Troia, G., Harbaugh, A., Shankland, R., Wolbers, K., & Lawrence, A. (2013). Relationships between writing motivation, writing activity, and writing performance: Effects of grade, sex, and ability. *Reading and Writing, 26*(1), 17–44. https://doi.org/10.1007/s11145-012-9379-2.

Tschannen-Moran, M., Hoy, A., & Hoy, W. (1998). Teacher efficacy: Its meaning and measure. *Review of Educational Research, 68*(2), 202–248. https://doi.org/10.3102/00346543068002202.

Inspectie van het Onderwijs. (2010). Het onderwijs in het schrijven van teksten. De kwaliteit van het schrijfonderwijs in het basisonderwijs. Utrecht: Ministerie van Onderwijs, Cultuur en Wetenschap.

Vaughn, S., Klingner, J., Swanson, E., Boardman, A., Roberts, R., Mohammed, S., & Stillman-Spisak, S. (2011). Efficacy of collaborative strategic reading with middel school students. *American Educational Research Journal, 48*(4), 938–964. https://doi.org/10.3102/0002831211410305.

Vlaams Ministerie van Onderwijs en Vorming. (2019). Peiling Nederlands. Lezen, luisteren en schrijven in het basisonderwijs 2018. Retrieved from Leuven: https://peilingsonderzoek.be/wp-content/uploads/2018/11/Brochure_Nederlands_BaO_2018_DEF.pdf.

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

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