EFFECTS OF WRITING INSTRUCTION ON ADOLESCENTS' KNOWLEDGE OF WRITING AND TEXT QUALITY IN HISTORY

JANNET VAN DRIE, TANJA JANSSEN & TALITA GROENENDIJK

Research Institute of Child Development and Education, University of Amsterdam

Abstract

Developing adolescents' writing ability is an important goal of present-day education, but hard to attain. In order to write well, writers must possess knowledge about what to write (content knowledge) and how to write well in a particular genre (metacognitive knowledge). Becoming a proficient writer across genres and disciplines requires systematic and explicit instruction.

We investigated the effects of domain-specific writing instruction on students' knowledge of writing and text quality in the context of argumentative writing in history. Participants were three history teachers and one Dutch language teacher and their classes (Grade 8, 10 and 11). The teachers each designed a writing intervention based on design principles for effective writing instruction. A pre-posttest quasi-experimental design was used to investigate the effects of each intervention. Results showed a positive effect of instruction on writing knowledge; students in all three interventions produced more writing advice, especially more genre-specific and product-related recommendations. With respect to text quality an improvement on genre-specific aspects was found. Correlational analyses between knowledge of writing and text quality after the intervention only yielded positive effects for the intervention in Grade 8. In sum, this study shows that teacher-designed interventions in secondary education may improve adolescents' knowledge of writing and quality of writing.

Keywords: writing in the disciplines, history, argumentative writing, metacognitive knowledge of writing

van Drie, J., Janssen, T. & Groenendijk, T. (2018). Effects of writing instruction on adolescents' knowledge of writing and text quality in history. Contribution to a special issue in honor of Gert Rijlaarsdam Making Connections: Studies of Language and Literature Education. *L1-Educational Studies in Language and Literature, 18, pp. 1-28. https://doi.org/10.17239/L1ESLL-2018.18.03.08*

Corresponding author: Jannet van Drie, Research Institute of Child Development and Education, University of Amsterdam, Nieuwe Achtergracht 127, 1018 WS Amsterdam, the Netherlands. Email: J.P.vanDrie@uva.nl

1. INTRODUCTION

Preparing adolescents for our ever-changing society requires that they are equipped with high levels of literacy skills (Biancarosa & Snow, 2006; Monte-Sano, 2010; Shanahan & Shanahan, 2008). They should be able to read, understand and write a variety of different text-genres within various disciplines. Or in the words of Shanahan and Shanahan (2008, p. 43): "Students must attend to sophisticated genres, specialized language conventions, disciplinary norms of precision and accuracy, and higher-level interpretive processes." Moreover, from secondary education onwards, content-area knowledge becomes more advanced and the texts and tasks students are confronted with require specialized forms of knowledge, related to the specific disciplines. Becoming a proficient writer across genres and disciplines requires systematic and explicit instruction (Miller, Scott, & McTigue, 2018). In the Netherlands, recently more attention is given to reading and writing in different school subjects, both in policy and teaching practices. For example, in the referential levels for language (Expert Group Learning Trajectories, 2009) it is explicitly stated that attention for language, reading and writing is not only a matter for the language subjects, but for all school subjects.

In the school subject of history writing plays an important role. History is prominently a literate discipline, as it is rooted in the analysis of texts, including historical sources, and the (re-)construction of interpretations in written form (Monte-Sano, 2010; van Drie, van Boxtel, & Braaksma, 2014). The language demands set by these kinds of activities are high and highly discipline-specific. Students for example face difficulties with understanding the interpretative and constructed nature of history and why argumentation is needed (Nokes & De La Paz, 2018), with the role of evidence (McCarthy Young & Leinhardt, 1998; Monte-Sano, 2010; Wineburg, 1991). Explicit domain-specific writing instruction is needed to overcome these challenges (e.g., Nokes & De La Paz, 2018; van Drie, Braaksma, & van Boxtel, 2015).

From writing research, it is known that in order to become a proficient writer, students do not only need content knowledge, but also knowledge of writing products and processes (McCutchen, 1986; 2011). This meta-cognitive knowledge of writing refers to knowing "[….] what constitutes a good text and which writing strategies are likely to be successful in dealing simultaneously with all the constraints writing a text poses" (Schoonen, van Gelderen, de Glopper et al., 2003, p. 168). Research indicated that metacognitive knowledge is positively related to writing performance, and that successful writers have more declarative, procedural and conditional knowledge about writing than less successful writers (Bouwer & Koster, 2016; Klein & Kirkpatrick, 2010; Schoonen & de Glopper, 1996). This suggests that knowledge of writing products and processes in history might be related to the quality of students' writing in history. However, this has not yet been investigated.

Whereas several studies investigated the effects of domain-specific writing instruction and revealed positive effects on the quality of students' writing in history, as well as their content knowledge (see for a review: Klein & Boscolo, 2016), little is

thus far known about the effects on students' metacognitive knowledge of writing. Bangert-Drowns, Hurley and Wilkinson (2004) highlighted the role of metacognition in writing-to-learn approaches. They found in their meta-analysis of writing-to-learn interventions a positive effect of metacognitive prompts, suggesting that a metacognitive position of the learner contributes to learning in the disciplines. This view is supported by a recent review of Miller et al. (2018). However, their focus is mainly on metacognitive strategies for content learning, for instance through learning to write journals. Klein and Kirkpatrick (2010), did highlight the role of knowledge of writing in their theory that instruction affects genre knowledge, which affects text quality and that this in turn predicts learning.

The present study aims to advance our insight in effects of domain-specific writing instruction on adolescents' metacognitive knowledge about writing and the quality of their writing in history. We investigate this in the context of three teacher-designed domain-specific writing interventions in Dutch secondary history education. This study is thus conducted in a more ecological setting as most often it are researchers who design the interventions (Koster, Tribushinina, de Jong, & van den Bergh, 2015). We are especially interested in what kind of metacognitive knowledge might be fostered by writing instruction, i.e., whether this knowledge is more product or process related and whether it is related to the quality of students' writing in history.

2. THEORETICAL FRAMEWORK

2.1 Knowledge of writing

Writing can be a tool for communicating and learning in content area subjects. However, writing is a complex and demanding task, as writing researchers have repeatedly pointed out (Kellogg, 2008; Rijlaarsdam, Braaksma, Couzijn et al., 2005). In the writing process model of Flower and Hayes (1980), updated by Hayes in 1996, writing is depicted as a complex problem-solving process. The model contains three main components: (a) the task environment (e.g., the task at hand and the text produced so far); (b) the writer's long-term memory, with knowledge about the topic, audience, genres, and task approaches; and (c) a set of cognitive activities (planning, formulating, transcribing, revising, editing). These activities are overseen by a metacognitive monitoring function, which enables the writer to monitor and evaluate how well thinking and writing is going.

This model and other models of the writing process indicate that, in order to write well, writers must possess knowledge about what to write and how to write well in a particular genre. Research showed that more knowledge about the topic leads to a better text (McCutchen, 1986). However, content knowledge is not enough. Writers also need knowledge of the characteristics of a good text and how to achieve that. According to the writing model of Hayes (1996), writers should have knowledge of schemata of different text genres; their structure, their components

and knowledge of typical linguistic markers to relate these components. This knowledge of a text genre and having organized schemata in long-term memory can facilitate planning and revision processes (McCutchen, 2011). Thus, students also need knowledge of written products and of writing processes (McCutchen, 1986; 2011). This kind of knowledge is referred to as metacognitive knowledge (Bouwer & Koster, 2016; Flavell, 1979; Harris, Graham, Brindle & Sandmel, 2009; Schoonen & de Glopper, 1996). Harris et al. (2009) make a distinction, based on general literature on metacognitive knowledge, between declarative, procedural and conditional knowledge of writing. *Declarative knowledge* of writing includes knowledge of genre, structure and goal of the text. This product knowledge is related to knowledge of specific text genres, more general characteristics but also more domain-specific ones. Especially within secondary education, this knowledge of domain-specific aspects of writing becomes more important (Miller et al., 2018). *Procedural knowledge* involves knowledge about the execution of the writing task, how to write, and strategies to attain a specific goal. *Conditional knowledge* is knowledge about when and how to apply strategies in order to attain the intended goal. Throughout schooling years, students' metacognitive knowledge moves from predominantly declarative knowledge in the first grades of primary school to more procedural and conditional knowledge from grade 5 onwards (Bouwer & Koster, 2016). Despite this general trend, Schoonen and de Glopper (1996) concluded that most of the 15 years-old in their study seemed to possess restricted knowledge about writing, mostly focused on formal criteria for the written product and hardly on writing processes.

Research indicated that metacognitive knowledge of writing and text quality are related (e.g., Klein & Krickpatrick, 2010; McCutchen, 1986; Schoonen & de Glopper, 1996). According to Torrance (1996) familiarity with the specific genre influences writing processes. Schoonen and de Glopper (1996), for instance, found in a study with more than thousand 15 years-old students (9th grade, various educational levels) that better writers had more knowledge about writing (indicated by the number of recommendations they gave in a letter of advice to an imaginary peer). Furthermore, they found that this knowledge was less directed to superficial aspects such as presentation and grammar, but more to higher order aspects, for example the organization of the text.

Writing interventions can affect students' metacognitive knowledge of writing in a positive way. For instance, Klein and Kirkpatrick (2010) found that instruction in content area writing in Grades 5 and 6 affected students' genre knowledge, which in turn affected text quality, which predicted learning during writing. Bouwer and Koster (2016) investigated in a quasi-experimental study the effects of a writing intervention in Grades 4 to 6 and reported an increase in knowledge of writing, especially with regards to higher order aspects of writing (i.e., style, content, and text organization) and writing processes.

Thus far, research on metacognitive knowledge of writing predominantly focused on primary education and not on secondary education and in most studies researchers were responsible for the learning materials used, not teachers themselves

(Koster et al., 2015). Furthermore, none of the previous studies focused on the domain of history, as far as we know.

2.2 Developing students writing in history

An important goal for today's history education is that students acquire the skill of historical reasoning. Historical reasoning aims at reaching 'justifiable conclusions about processes of continuity and change, causes and consequences, and/or differences and similarities between historical phenomena or periods' (van Boxtel & van Drie, 2018, p. 151). It comprises of activities such as asking historical questions, contextualizing, using historical concepts and meta-concepts (i.e., cause, consequence, change), using historical sources and providing arguments to back assertions (van Boxtel & van Drie, 2018; van Drie & van Boxtel, 2008). Writing is one of the means in which this reasoning can be expressed and developed. Monte-Sano (2010) argues that writing and reasoning are intrinsically linked and that writing is essential to learn the substantive and procedural forms of knowledge of the specific discipline.

Within secondary history education students learn to write different types of texts. Coffin (2006) distinguishes three genre families in historical writing in secondary education: recording, explaining and arguing. Each of these genres requires different kinds of historical reasoning (Monte-Sano & De La Paz, 2012; van Drie, van Boxtel, & van der Linden, 2006). But also within a specific text genre different aspects of historical reasoning can be highlighted, depending on the task at hand. For example, the prompt "Were the changes in the behavior of the youth in the Sixties revolutionary?" puts emphasis on describing historical changes and the prompt "Was Germany guilty of the outbreak of the First World War?" focuses on explaining a historical phenomenon. Both prompts belong to the arguing family and require taking a standpoint, providing arguments to support this view, however the first focuses on describing changes, whereas the latter focuses on explaining. So, in order to write well in history students should also have knowledge about specific text-genres in history.

In recent years, several meta-studies have provided us with indications of effective writing instruction (e.g., Graham & Perin, 2007; Graham & Harris, 2018). Among these are for example writing strategy instruction, studying model texts, prewriting activities, and collaboration during writing (i.e., collaborative writing, peer feedback). Some of these approaches put more emphasis on writing processes, others on the writing product. Writing strategy instruction is, for instance, an approach for the systematic teaching of strategies for planning, revising and editing texts, thus focusing on writing processes. Studying text models is more product-oriented. By analyzing examples of a particular type of text and extracting elements of good writing, students become aware of how a good text should look like which they can incorporate in their own writing, which has a positive effect on text quality (Hillocks, 1986; Janssen & Overmaat, 1990).

These insights on effective writing instruction are also important for writing in the disciplines. Recent reviews by Klein and Boscolo (2016) and Miller and colleagues (2018) showed that students' writing in the content areas can be improved by instruction. In addition, van Drie et al. (2015) found that domain-specific writing instruction had added value over general writing instruction (as provided in L1 classes). Within history, explicit strategy instruction has been most often investigated and yielded positive effects on students' writing in history (e.g., De La Paz, 2005; De La Paz & Felton, 2010; De La Paz, Ferretti, Wissinger, Yee, & MacArthur, 2012; Martinez, Mateos, Martin, & Rijlaarsdam, 2015; Nokes & De La Paz, 2018). For example, De La Paz and Felton (2010) used strategy instruction focused on reading historical documents (including considering the author, understanding the source and critiquing the source) as well as on the writing process, in particular generating ideas on both sides of an argument before taking a standpoint. Other instructional approaches have been investigated less often in history, although an example of using text models can be found in van Drie et al. (2015).

Despite these positive findings of domain-specific writing instruction, history teachers do not often incorporate writing instruction in their teaching and if they do they only provide some directions for the writing product, such as structure and layout. Hardly any attention is given to writing processes such as generating ideas, organizing, and revising, or specific characteristics of genres, goal and audience (De Oliveira 2011; Holdinga, 2013; McCarthy Young & Leinhardt, 1998; Mottart, van Brabant, & van de Ven, 2009; van der Leeuw & Meestringa, 2011). It is not yet clear what the effects are on students' knowledge of writing if history teachers incorporate writing instruction in their lessons.

2.3 Aims and research questions

This study aims to advance our insight in effects of teacher-designed domain-specific writing instruction on adolescents' metacognitive knowledge about writing and their writing performances in history. We are especially interested in what kind of metacognitive knowledge might be elicited, i.e. more product or process related knowledge and whether students' knowledge of writing is related to the quality of their writing.

The research questions are:

1) Does domain-specific writing instruction in history lessons have a positive effect on students' metacognitive knowledge of writing?
2) Does domain-specific writing instruction in history lessons have a positive effect on the quality of students' writing?
3) Is there a relationship between students' metacognitive knowledge of writing and writing performance after the intervention?

We hypothesized that (a) domain-specific writing instruction would positively affect students' knowledge of writing, in particular with respect to the number of product- and process related recommendations that students would provide to peers, and

that (b) domain-specific writing instruction would positively affect text quality. Furthermore, it was hypothesized that knowledge of writing and text quality would correlate after the intervention.

3. METHOD

3.1 Design

Data for this study derived from a larger study on effects of a professional development program on domain-specific writing in secondary education that included measures on students' outcomes (van Drie, Janssen, & Groenendijk, 2017). For the present study we selected the three writing interventions from this project that were specifically developed for the subject of history. In all three cases a quasi-experimental design was used with pretest, posttest and a control group. Data included student answers on a knowledge of writing test and their performance on a writing task in history.

3.2 Participants

Three history teachers from two schools in the Netherlands volunteered to participate in the study; two of them taught at the same school. One of the history teachers collaborated with a Dutch language (L1) teacher, who taught the same classes. They were all experienced teachers, with an academic degree. The three history teachers were male, the L1 teacher female. The history teachers had no experience beforehand in teaching writing to their students, and it was not yet part of their curriculum.

The teachers participated in an 18-hours professional development program aiming at developing teachers' competences of teaching domain-specific writing instruction. The program was centered around five design principles for effective writing instruction, which were derived from research literature (e.g., Graham & Perin, 2007) and were considered suitable for designing domain-specific writing instruction. These principles were: writing strategy instruction (including modelling), studying text models, prewriting activities, collaborative writing (including peer feedback), and using (semi-)authentic tasks. Within the PD program the teachers designed lessons for their own classes and addressed writing problems they encountered in these classes, based on the design principles. For a detailed description of the PD and its' effects on teacher professionalization see van Drie et al. (2017).

Each history teacher participated with two classes: an experimental and a control group. The first intervention took place in Grade 8 (experimental, N = 28; control, N = 28). The second intervention took place in Grade 11, with students in their final year of higher general secondary education (experimental, N = 19; control, N = 26). The third intervention took place in both the history and L1 lessons, and was conducted in Grade 10, pre-university level (experimental, N = 25; control, N = 13). Students who missed two of the four measurements were excluded from the analyses.

3.3 Writing interventions

All three writing interventions aimed at developing students' argumentative text writing in history, but highlighted different aspects of it (see the descriptions below). In designing the interventions, the teachers were free to choose which of the five design principles they would use and adapt to their specific teaching context. The interventions thus varied, both in historical topic, as in instructional aspects. Interventions 1 and 2, in Grade 8 and 11 respectively, each lasted five lessons of 50 minutes. Intervention 3 in Grade 10 also included L1 lessons and took in sum six lessons of 50 minutes. Students in the control condition received regular lessons. They worked on the same content and the same writing tasks, but did not receive domain-specific writing instruction.

Table 1 provides a general overview of the lessons. Below we describe the interventions in more detail and refer to the design principles in italics.

Table 1. Overview of the three interventions, designed by the teachers.

Intervention	Grade	Number of lessons	Historical topic	Writing task	Design principles
1	8	5	Industrial Revolution	Argumentative letter from the perspective of a historical agent to a committee	Authentic writing task Prewriting Modelling
2	11	5	Persecution of Jews in the Netherlands during WWII	Argumentative letter to a historian	Authentic writing task Prewriting Modelling Studying text models Peer feedback
3	10	4 (L1) 2 (history)	Dutch Golden Age	Answering open ended questions that call for a standpoint and arguments	Prewriting Collaborative writing Studying text models Peer feedback

Intervention 1 Grade 8. The main aim of this intervention was to develop students' historical empathy ability and writing, in particular with respect to text structure and argumentation. The topic of the lesson series was the Industrial Revolution and the social conditions in the Netherlands, in particular in the town the school was situated in. The intervention centred around a writing task asking students to take the perspective of a specific person of that time (i.e., a female worker, a pastor, a factory owner) and write a formal letter about their views on the social conditions to a committee investigating workers' living and working conditions. The length of the text was about 300 words. This task can be considered a semi-authentic task, as it was

situated in a more or less realistic setting and the text goal and audience were clear (*authentic writing task*).

In the first lesson the task was introduced to the students. Furthermore, students already had sufficient background knowledge about the Industrial Revolution. Next, the focus was on argumentation in history, which the teacher modelled and discussed with the students (*modelling*). Next, students practiced this on a new question using a worksheet *(prewriting activities*). The second lesson focused on finding, selecting and ordering information from historical sources from a local archive. A scheme was used in which students could list the arguments and counterarguments they planned to use (*prewriting activities*). At the end of the lesson there was a short reflection on what went well and what remained difficult. During the third lesson students could continue with this task and handed in their scheme at the end of the lesson. The teacher then checked the schemes and gave students feedback on them. Finally, lessons 4 and 5 were spent on writing the final text.

Students in the control condition worked on the same task and studied the same topic information, but did not receive the writing instruction, such as modelling of argumentation in history and prewriting activities.

Intervention 2 Grade 11. The main aim of this intervention was to develop students' writing of a text in which they argued whether or not they agreed with an historical account about the persecution of Jews in the Second World War. The writing task was to write an argumentative letter (about 500 words) to a historian who wrote a book in which he claimed that the Dutch people were partly responsible for the Holocaust, as they could have known at that time what was happening to the Jews. The students could use a set of seventeen historical sources, including primary sources such as diary fragments and secondary sources. This task can be considered a semi-authentic task, as it was situated in a more or less realistic setting (various readers actually wrote a letter in response to this book) and the text goal and audience were clear (*authentic writing task*).

In Lesson 1, the writing task was introduced and background information on the topic of the Jews in the Netherlands during WWII was provided. In addition, instruction on argumentation in history was given (use of arguments pro and contra and weighing arguments, use of contextual information) and students completed some short assignments (*prewriting activities*). In lesson 2 instruction was provided on the analysis of historical sources, with a step-by-step guide. Students practiced using this guide and analyzed the sources in dyads, using a worksheet (*prewriting activities*). In lesson 3 the students received instruction on argumentative text structure, in which text examples were also studied (*studying text models*). The teacher modelled writing an argumentative letter (*modelling*), and students then worked on their own writing plan (*prewriting activities*). Lesson 4 was spent on writing the text, based on their writing plan and using the historical background and the analysis of the sources. Finally, in lesson 5 students gave feedback on another student's text, using a rubric, and students then revised their own text based on this feedback (*peer feedback*).

Students in the control condition worked on the same task as the experimental group. They received the same information about the topic. However, they did not receive the writing instruction parts (i.e., instruction on argumentation in history, modelling, prewriting activities, studying text models and peer feedback).

Intervention 3 Grade 10. This intervention differed from the other two interventions as it focused on improving students' writing in the context of answering open-ended questions (that call for a standpoint and arguments) in the context of assessment and without reading additional sources. In addition, the intervention also included L1 lessons. The complete intervention included three text genres (compare and contrast, argumentation and explaining), but we focus here on the lessons on argumentative writing. The L1 teacher designed eight lessons, of which four lessons focused on argumentative writing (the other lessons focused on writing goals and use of sources in general and on other text genres). The L1 lesson preceded the history lessons, so students could build on their knowledge acquired in L1.

The first L1 lesson on argumentative writing focused on prewriting activities such as different kind of arguments that could be used and making a writing plan. Next, students wrote a text in pairs (*collaborative writing*). The second lesson focused on criteria of a good argumentative texts by discussing one of the written texts from the first lesson (*studying text models*). Students reviewed a text of another pair based on the criteria discussed before and provided feedback (*peer feedback*). They revised their own text and in the next lesson they received feedback of the teacher. In the third lesson, students had to write another text. After a short introduction, they made a writing plan which they discussed in pairs (*prewriting*). In pairs they wrote one text (*collaborative writing*). At the end of the lesson, the teachers discussed two students' text with the whole-class (*studying text models*). In the final lesson students discussed texts from a newspaper and checked whether all criteria discussed in the lesson applied to these texts and which texts were best and why (*studying text models*). Students gathered all their texts in a portfolio.

Within the history class one lesson was dedicated to argumentative writing. The topic was the Dutch Golden Age and students practiced answering open ended questions that call for a standpoint and arguments. Students made a formative task in which they had to argue to what extend they agreed with the statement that the province Holland had most power in the Dutch Republic. Next, the teacher provided instruction on the structure of a correct answer and the different parts in it (which was discussed beforehand during L1 lessons). Different colors were used to make different parts of the texts visible. In addition, a step to step guide was provided for answering these kind of argumentative questions (*strategy instruction*). He modelled the use of this guide (*modelling*). Next, students practiced using the guide on three new questions. They checked each other's answers and ways of formulating and they provided each other feedback (*peer feedback*). The next two lessons focused on other text genres (compare and contrast, explanation). In the final lesson students practiced with the three different genres.

The control condition worked in the history lessons with the same content, but did not receive writing instruction (e.g., strategy instruction, modelling, peer feedback). Nor did they receive the instruction on text-genres in the L1 lessons.

3.4 Instruments

Knowledge of writing. We measured students' knowledge of writing with a task based on Schoonen and de Glopper (1996). In the original task students were asked to write a letter of advice to a friend who is planning to attend their school and who asked them to explain how to write a composition that will be considered good by teachers in their school. This turned out to be a powerful task for eliciting students' declarative and procedural knowledge of writing (Schoonen & de Glopper, 1996; Bouwer & Koster, 2016). As it is an open task, students can include different kinds of recommendations. In our adapted version we asked students to write a short email to a friend who needs advice on how to write an argumentative text in history (see Appendix 1). We specifically included the domain and the text genre in the assignment, as we were interested in students' knowledge about genre-specific writing in the domain of history. Students had 15 minutes to complete this task.

All recommendations provided in the advice texts were coded; parts of the texts that did not contain a writing recommendation (i.e., greetings) were left out of the analysis. First, we counted all recommendations and identified whether they were genre-specific (yes/no) and domain-specific (yes/no). Genre-specific refers to recommendations that were specifically related to the genre of argumentative writing, instead of writing in general. Domain-specific recommendations refers specially to the domain of history. Next, a distinction was made between product related and process related recommendations, which are mutually exclusive. Product related refers to recommendations related to information about text characteristics (i.e., structure, main components, text-type). Process related refers to recommendations about the process of writing or writing strategy (i.e., orientating, generating content, revising). Thus, each recommendation received three codes: for genre-specific, for domain-specific, and either product or process related. For example, the recommendation 'You should back your claim with arguments from historical sources.' was coded as genre-specific, domain-specific and product related. The recommendation 'First, make a list of the information you want to write about.' was scored negative on genre-specific and domain-specific, but positive on process related. The codes genre-specific, domain-specific and product related reflect students' declarative knowledge about writing, whereas process related measures students' procedural knowledge. The categories product related and process related were further divided into subcategories, based on a version of the coding scheme of Schoonen and de Glopper (1996). Table 2 provides an overview of the categories and their examples. Interrater agreement was calculated by having 40 texts (16% of the total number of 247 letters) coded by two independent raters (first author and a research-assistant). The texts were equally spread over pre or post measurement, condition and grade

and chosen randomly within each group. Agreement between the two coders was high; Cohens' Kappa varied between .92 and 1.00.

Table 2. Overview of categories of knowledge of writing.

Category	Description	Examples
1.Genre specific	Related to argumentative writing	In the middle part you present your arguments on the issue. Refuting a counter argument can convince your readers
2.Domain specific	Related to the domain of history	Consider the reliability of the historical source, that is important in history. In history you have to look at situations from different perspectives, so add them in the text.
3. Product related		
3a. Goal / text type	Goal of the task (i.e., convincing), in relation to the readers	Refuting a counter argument can convince your readers. An attractive introduction makes that the reader wants to continue reading. Start your letter with the name and address of whom you are writing to.
3b. Main components	Main parts of text (i.e., arguments)	Take care that you always have at least two arguments in favor of your opinion. Add facts in your argumentation. In history you have to look at situations from different perspectives, so add them in the text.
3c. Text structure / connectives	Ordering of the text, introduction, main part, use of connectives	Start your text with the statement. Lastly, you add a conclusion. You can use connectives to relate the different parts of the text, such as first, second, in sum.
3d. Formulation/style	Directed to formulation and style (i.e., short sentences, spelling)	Try to stick to the subject and not to stray from your subject. As texts in history are often long it is good to pay attention to mistakes in spelling and grammar.

4. Process related		
4a. Orientation	Orientation on the task requirements	You start by reading the assignment very closely. Read the task carefully, so that you do not miss anything.
4b. Generation	Generation of content, reading about content	Read about the subject on the internet. You have to decide whether you are pro or against the statement.
4c. Preparation	Prewriting activities (i.e., make a scheme, a draft)	It is most easy to make a writing plan beforehand. Make a mind-map with the most important information from the texts.
4d. Evaluation / revision	Evaluation and revision of the text	After that I would reread the text I have written thus far. Revise mistakes and wrong sentences.
4e. Other	About the writing process, but does not fit in the other categories	Lastly, take care not to start the day before handing in, but take your time. Think before you write, so you do not have to improve yourself all the time.

Writing performance. To measure the quality of students' writing a writing task was administered at pretest and posttest. For each intervention a different writing task was designed by the teachers, adapted to the particular content and similar to the writing task of the intervention. The prompts for the pre- and posttests were similar, although the topics differed. They corresponded to topics that had been taught previously, to ensure that a lack of content-knowledge did not influence the results. Similar as in the writing task in the intervention, the writing tests in Grade 8 and 11 included the reading of some background information (Grade 8) or historical sources (Grade 11), and the writing of an argumentative text (a letter). In contrast, students in Grade 10 did not have to do additional reading; they wrote answers to open ended questions. This explains the shorter text length as well as the shorter time length for carrying out the writing test in Grade 10: 10 minutes versus 50 minutes in Grades 8 and 11. The pre- and post-writing tests were administered during the history lessons. Table 3 provides details of the pre- and post-writing tests for the three interventions. Appendix 2 contains two examples of writing tasks and student writing.

Table 3. Overview of the pre- and post-writing tests for the different interventions.

	Intervention 1 Grade 8	Intervention 2 Grade 11	Intervention 3 Grade 10
Genre writing task	Argumentative letter from perspective of historical person	Argumentative letter in response to a historical account	Argumentative open-ended question
Topic pretest	Dutch Golden Age	Imperialism	Slavery
Topic posttest	Industrialism	Cold war	Absolutism
Length text	Minimum 300 words	Minimum 300 words	Maximum 200 words
Time	50 minutes	50 minutes	10 minutes

The quality of students' texts was assessed using a rubric that contained three main criteria: Genre-specific writing quality, General writing quality, and Subject matter quality. Genre-specific writing quality (maximum score 14) contained sub criteria for the introduction (i.e., standpoint is mentioned), the middle (i.e., mentions several arguments pro, provides evidence to support the arguments, includes counter arguments and refute these) and the closing of the text (i.e., summarizes viewpoint and main arguments, does not include new arguments). General writing quality encompassed the sub criteria audience-orientation, coherence, language use and spelling (maximum score 18). And lastly Subject-matter quality referred to the adequate use of subject-specific concepts and content (maximum score 12).

Teachers were trained using this rubric and scored their own students' texts. To check the interrater reliability, 62 texts (about 25% of the total number of 248 texts) were rated by an independent second rater (one of the authors). The texts were equally chosen from grade and pre- and post-measurement. Correlations between raters over the three criteria (Spearmans *R*) were acceptable, ranging from .68 to .91 ($p < .01$).

3.5 Analyses

For each of the different categories of knowledge of writing and criteria for writing quality means and standard deviations were calculated, per intervention. To determine effects of the intervention, univariate analyses with pretest scores as co-variate were used to evaluate changes in students' knowledge of writing and writing performance. To analyze whether students' knowledge of writing and writing performance were related after the intervention, we used the posttest scores of the students in the experimental groups, and calculated the correlations (Spearmans' *R*) between students' number of recommendations in the knowledge of writing test and their overall score on text quality (measured as the sum score of the three criteria in the rubric).

4. RESULTS

First, we will report the outcomes on students' knowledge of writing, and next the results on text quality. As the three interventions differed in historical subject matter, writing tasks as well as in types of instructional support for writing (see Table 3), we report the outcomes for the different interventions separately. Subsequently, we will report the outcomes of the correlational analyses.

4.1 Effects on knowledge of writing

Table 4 presents the mean number of writing recommendations students provided per grade and condition, and the mean number of genre-specific and domain-specific recommendations at pretest and posttest. In general, the outcomes show that

students in Intervention 2 and 3 (Grade 10 and 11) provided more recommendations compared to students in Intervention 1 (Grade 8).

Table 4. Mean frequencies and standard deviations for total number of recommendations, number of genre-specific and domain-specific recommendations for the three interventions.

	Pretest		Posttest	
	Experimental	Control	Experimental	Control
Intervention 1 Gr 8	*N* = 25	*N* = 27	*N* = 22	*N* = 28
Total recommendations	4.12 (1.42)	4.56 (2.22)	4.95 (1.62)	3.96 (1.75)
Genre-specific	2.04 (1.49)	2.56 (1.80)	2.64 (1.68)	1.32 (1.12)
Domain-specific	0.04 (0.20)	0.11 (0.42)	0.32 (0.48)	0.25 (0.56)
Intervention 2 Gr 11	*N* = 15	*N* = 24	*N* = 12	*N* = 22
Total recommendations	7.67 (2.26)	6.83 (2.94)	10.00 (2.73)	6.45 (3.16)
Genre-specific	3.87 (1.36)	2.88 (2.52)	6.58 (2.27)	3.23 (2.05)
Domain-specific	1.00 (1.20)	0.38 (0.65)	1.58 (1.44)	0.18 (0.66)
Intervention 3 Gr 10	*N* = 25	*N* = 12	*N* = 25	*N* = 10
Total recommendations	8.04 (2.76)	10.92 (2.97)	9.12 (2.86)	6.50 (3.81)
Genre-specific	5.20 (2.78)	6.17 (2.82)	5.40 (1.89)	4.60 (4.43)
Domain-specific	0.48 (0.77)	1.33 (2.42)	0.24 (0.52)	1.00 (2.21)

To answer the question whether the three interventions resulted in more knowledge of writing Ancovas were used with pretest scores as covariate. Students in the experimental condition in all grades gave significantly more recommendations compared to students in the control groups: Intervention 1 $F(1, 46) = 7.602$, $p = .008$; Intervention 2 $F(1, 31) = 10.252$, $p = .003$; and Intervention 3 $F(1, 32) = 7.049$; $p = .012$. Furthermore, a significant effect was found for the number of genre-specific recommendations for Intervention 1 $F(1, 46) = 14.401$, $p = .000$, and Intervention 2 $F(1, 31) = 17.539$; $p = .000$, but not for Intervention 3. Students in Intervention 2 also gave more domain-specific recommendations, $F(1, 31) = 13.072$, $p = .001$.

Table 5, 6, and 7 present the results on the product and process related recommendations for the three interventions respectively. Overall, students in Intervention 2 and 3 gave more product related recommendations compared to process related recommendations, and both are low in Intervention 1. Again, Ancovas were used to determine effects of the intervention. First, with respect to Intervention 1 (Table 5), after controlling for pretest scores, a significant effect was found for the total number of product related recommendations $F(1, 46) = 7.629$, $p = .008$, and for the subcategory Main components $F(1, 46) = 5.191$, $p = .027$.

Table 5. Mean frequencies and standard deviations of product and process recommendations for pre- and posttest in Intervention 1 Grade 8.

	Pretest		Posttest	
	Experimental	Control	Experimental	Control
	N = 25	*N* = 27	*N* = 22	*N* = 28
	M (SD)	*M (SD)*	*M (SD)*	*M (SD)*
Product related				
-Goal text/text type	0.08 (0.40)	1.93 (1.30)	0.36 (0.58)	0.14 (0.45)
-Main components	1.40 (1.32)	1.70 (1.14)	1.77 (1.27)	1.07 (1.36)
-Structure/connectives	0.08 (0.40)	0.04 (0.19)	0.41 (0.67)	0.14 (0.36)
-Formulation/style	0.24 (0.60)	0.15 (0.46)	0.46 (0.22)	0.11 (0.31)
Total	*1.80 (1.76)*	*1.93 (1.30)*	*2.59 (1.53)*	*1.46 (1.62)*
Process related				
-Orientation	0.32 (0.48)	0.22 (0.42)	0.27 (0.46)	0.32 (0.67)
-Generation	1.08 (0.95)	1.30 (1.41)	1.64 (1.05)	1.43 (1.03)
-Preparation	0.36 (0.86)	0.30 (0.61)	0.32 (0.55)	0.32 (0.55)
-Evaluation/revision	0.00 (0.00)	0.22 (0.51)	0.14 (0.35)	0.25 (0.52)
-Other	0.00 (0.00)	0.15 (0.36)	0.09 (0.29)	0.18 (0.39)
Total	*1.76 (1.59)*	*2.19 (1.30)*	*2.18 (1.26)*	*2.50 (1.48)*

As for Intervention 1, also in Intervention 2 (see Table 6) positive effects of the condition were found for Product related recommendations $F(1, 31) = 12.946$, $p = .001$ and for Main components $F(1, 31) = 18.438$, $p = .000$.

Table 6. Mean frequencies and standard deviations of product and process recommendations for pre- and posttest in Intervention 2 Grade 11.

	Pretest		Posttest	
	Experimental	Control	Experimental	Control
	N = 15	*N* =2 4	*N* = 12	*N* = 22
	M (SD)	*M (SD)*	*M (SD)*	*M (SD)*
Product related				
-Goal text /text type	0.40 (0.74)	0.33 (0.64)	0.33 (0.65)	0.27 (0.46)
-Main components	4.73 (1.75)	3.46 (1.72)	6.92 (3.26)	2.68 (1.86)
-Structure/connectives	0.53 (0.83)	0.83 (1.34)	1.42 (1.31)	0.91 (1.06)
-Formulation/style	0.47 (0.92)	0.71 (1.27)	0.17 (0.39)	0.59 (0.91)
Total	*6.13 (2.59)*	*5.33 (3.58)*	*8.83 (4.41)*	*4.45 (2.76)*
Process related				
-Orientation	0.13 (0.35)	0.13 (0.34)	0.17 (0.58)	0.09 (0.29)
-Generation	1.13 (1.41)	0.75 (0.90)	0.67 (1.56)	1.27 (1.42)
-Preparation	0.00 (0.00)	0.21 (0.51)	0.00 (0.00)	0.27 (0.55)
-Evaluation/revision	0.07 (0.26)	0.00 (0.00)	0.33 (0.89)	0.09 (0.29)
-Other	0.20 (0.77)	0.08 (0.28)	0.00 (0.00)	0.23 (0.69)
Total	*1.53 (1.96)*	*1.17 (1.20)*	*1.17 (2.29)*	*1.95 (2.06)*

With respect to Intervention 3, students in in the experimental group (see Table 7) produced significantly more product related recommendations compared to the control group $F(1, 32) = 19.640$, $p = .000$. Students especially produced more recommendations related to Main components of the text $F(1, 32) = 16.547$, $p = .000$ and to Structure/connectives $F(1, 32) = 8.688$, $p = .006$. With respect to process related recommendations we see a different pattern. At posttest, the control group gave significantly more process related recommendations compared to the experimental group $F(1, 32) = 15.422$, $p = .035$. However, compared to the pretest scores, the mean score on the total number of process recommendations was about the same, whereas the mean score of the experimental group dropped.

Table 7. Mean frequencies and standard deviations of product and process recommendations for pre- and posttest in Intervention 3 Grade 10.

	Pretest		Posttest	
	Experimental	Control	Experimental	Control
	N = 25	*N* = 12	*N* = 25	*N* = 10
	M (SD)	*M (SD)*	*M (SD)*	*M (SD)*
Product related				
-Goal text /text type	0.52 (0.71)	0.33 (0.49)	0.56 (1.12)	0.30 (0.67)
-Main components	4.20 (2.08)	5.42 (3.06)	5.12 (2.30)	2.30 (1.77)
-Structure/connectives	0.68 (1.14)	1.67 (1.67)	1.76 (1.48)	0.30 (0.67)
-Formulation/style	0.92 (0.95)	0.83 (0.94)	0.72 (0.89)	0.80 (1.32)
Total	*6.32 (2.94)*	*8.25 (3.60)*	*8.16 (3.05)*	*3.70 (2.36)*
Process related				
-Orientation	0.48 (1.42)	0.00 (0.00)	0.12 (0.33)	0.30 (0.48)
-Generation	0.96 (1.34)	1.50 (1.45)	0.44 (1.19)	1.20 (1.40)
-Preparation	0.24 (0.83)	0.75 (1.76)	0.28 (0.89)	0.90 (1.37)
-Evaluation/revision	0.17 (0.39)	0.17 (0.39)	0.12 (0.33)	0.30 (0.48)
-Other	0.80 (0.28)	0.25 (0.62)	0.00 (0.00)	0.00 (0.00)
Total	*1.84 (2.64)*	*2.67 (2.15)*	*0.96 (1.74)*	*2.70 (2.31)*

To summarize the results, students in the experimental groups of all three interventions showed more knowledge of writing compared to the control group after the intervention, as they provided more writing advice. Furthermore, students in Intervention 1 (Grade 8) also produced more genre-specific recommendations and more product related recommendations, especially with respect to Main Components. The same pattern was found for Intervention 2 (Grade 11), but these students also gave more domain-specific recommendations. Students in Grade 10 gave more product related recommendations, in particular with respect to Main components and Structure/connectives. Students in the control condition gave more process related recommendations.

4.2 Effects on text quality

Table 8 presents students' mean scores for text quality at pretest and posttest. Ancovas were used to determine whether the quality of students' texts was influenced by the intervention lessons. For Intervention 1 (Grade 8) a significant positive effect of condition was found on students' posttest scores after controlling for pretest scores for Genre specific aspects $F(1, 45) = 30.421$, $p = .000$ and General writing aspects $F(1, 45) = 5.575$, $p = .023$. No effect was found for Domain specific aspects of students' writings. For Intervention 2 (Grade 11) Ancovas also yielded an effect of condition for Genre specific aspects $F(1, 34) = 14.164$, $p = .001$ and for General writing aspects $F(1, 34) = 29.969$, $p = 0.00$, with students in the experimental condition scoring higher than students in the control condition. With respect to Intervention 3 (Grade 10), after controlling for pretest scores significant effects of condition were found for categories Genre specific $F(1, 31) = 10.078$, $p = .003$ and Domain specific $F(1, 31) = 5.674$, p = .024. Students in the experimental condition scored higher on these aspects than students in the control condition.

Table 8. Mean scores and standard deviations for text quality at pretest and posttest for the three interventions.

	Pretest		Posttest	
Aspect	Experimental *M* (*SD*)	Control *M* (*SD*)	Experimental *M* (*SD*)	Control *M* (*SD*)
Intervention 1 Gr 8	*N* = 25	*N* = 25	*N* = 24	*N* = 27
Genre specific	6.48 (1.61)	6.44 (1.71)	9.46 (1.79)	7.33 (1.47)
General writing	10.28 (2.51)	9.68 (2.01)	10.33 (1.81)	9.00 (1.73)
Domain specific	7.32 (1.03)	7.00 (1.71)	6.79 (1.59)	6.52 (.94)
Intervention 2 Gr 11	*N* = 14	*N* = 24	*N* = 14	*N* = 24
Genre specific	7.79 (1.93)	7.25 (2.00)	10.14 (2.18)	7.71 (2.14)
General writing	10.21 (.89)	9.17 (2.20)	10.43 (1.40)	7.54 (1.61)
Domain specific	5.14 (.36)	4.88 (1.30)	4.86 (.36)	4.79 (1.02)
Intervention 3 Gr 10	*N* = 25	*N* = 11	*N* = 24	*N* = 11
Genre specific	5.48 (1.23)	5.55 (1.57)	8.00 (1.56)	5.91 (1.81)
General writing	8.68 (2.14)	7.82 (.98)	9.58 (1.74)	8.55 (1.97)
Domain specific	4.72 (1.96)	5.55 (2.01)	6.21 (2.32)	4.18 (1.25)

In sum, we found that all three interventions had a beneficial effect on text quality with respect to the genre-specific aspects of students' texts. In Grade 8 and 11 the intervention also led to higher scores on general writing aspects, and in Grade 10 a positive effect on domain-specific aspects of the texts was found.

4.3 Correlations between knowledge of writing and text quality

Our third research question was whether there was a relation between students' metacognitive knowledge of writing and their writing performance after the intervention. Correlations (Spearmans' *R*) were calculated between the overall text quality and the three criteria of text quality on the one hand and the mean total number of recommendations and genre specific, domain specific, product related and process related recommendations on the other hand. The scores of overall text quality were for Intervention 1 *M* = 26.58, *SD* = 3.97, for Intervention 2 *M* = 25.43, *SD* = 3.52 and for Intervention 3 *M* = 18.88, *SD* = 3.38.

For Intervention 1 (Grade 8) significant correlations were found, but not for Intervention 2 (Grade 11) and 3 (Grade 10). Significant correlations for Intervention 1 are presented in Table 9. A high correlation was found between the overall text quality and the total number of recommendations at posttest. Furthermore, correlations were found between overall text quality and the number of genre specific and product related recommendations. With respect to the three criteria, we found that General writing quality correlated positively with the mean number of recommendations, with genre specific recommendations and product related recommendations. Domain specific writing quality correlated positively with the mean number of recommendations, but not with particular types of recommendations. Additionally, we checked whether text quality and knowledge of writing correlated at pretest for this experimental group. No significant correlations were found, suggesting that the intervention might be of influence here.

Table 9. Significant correlations (Spearmans' R) between aspects of text quality and knowledge of writing at posttest for the experimental condition for Intervention 1 (Grade 8).

	Total recommendations	Genre specific	Domain specific	Product related	Process related
Overall text quality	.598**	.441*	-	.450*	-
Genre specific quality	-	-	-	-	-
General writing quality	.575*	.463*	-	.483*	-
Domain specific quality	.580*	-	-	-	-

* Correlation is significant at the 0.05 level; ** Correlation is significant at the 0.01 level.

5. DISCUSSION

In this quasi-experimental study, we investigated the effects of three domain-specific writing interventions on adolescents' knowledge of writing and the quality of their written texts. The interventions were designed and conducted by three history teachers and one Dutch language teacher in three different grades: Grade 8, 10 and 11. The lesson series varied with respect to the historical subject matter the teachers addressed and the instructional support they provided. All three interventions focused on argumentative writing.

5.1 Effects on knowledge of writing

Our first research question was whether domain-specific writing instruction would have a positive effect on students' metacognitive knowledge of writing. To examine this, we asked students to write a letter of advice to a peer on how to write an argumentative text in history, at pretest and posttest. Results showed that the writing interventions had a beneficial effect on students' knowledge of writing, as students in the experimental groups for all three interventions produced significantly more writing advice compared to the control groups. Students especially reported more genre specific (Intervention 1 and 2), and product related recommendations (Intervention 1, 2 and 3). Students of all interventions gave more recommendations about the main ingredients of the text. This supports our first hypothesis that domain-specific writing instruction may have a positive effect on students' knowledge of genre specific and product related aspects of writing. After the interventions, students seemed to have a better idea of the demands set by the specific genre and how the text should look like. This finding is in line with earlier research (e.g., Klein & Kirkpatrick, 2010; Bouwer & Koster, 2016).

It turned out the interventions did not result in more knowledge of writing processes. For Intervention 3 (Grade 10) we found that the control group produced more process related recommendations compared to the experimental group. Although the small *N* in the control group might be at stake here, closer inspection of the data showed that the scores of the control group at pretest and posttest stayed more or less the same, whereas the scores of the experimental group for the pretest were already lower compared to the control group and even dropped at posttest. On the other hand, the experimental group produced significantly more product related recommendations. This suggests that students in the experimental group after the intervention focused more on product related aspects. A possible explanation for the more general finding that the interventions did not result in more process related knowledge of writing is that, although the teachers did include process related aspects in their lessons such as prewriting and (peer) revision activities, these aspects were not explicitly taught as strategies students could use when writing a text. For our history teachers, who taught writing for the first time, direct and explicit writing strategy instruction might have been a step too far. After all, writing strategy instruction is still relatively uncommon in Dutch writing education (Mottart et al., 2009; Rietdijk, van Weijen, Janssen, van den Bergh, & Rijlaarsdam, 2018), and difficult to implement in a normal classroom practice (De La Paz, 2007). A next step for these teachers might be to include explicit process related writing instruction in their teaching. Effects of this approach could be subject of further research.

Overall, the number of domain-specific recommendations was quite low in all grades and conditions. Only for Intervention 2 (Grade 11) an increase in domain specific recommendations was found. A possible explanation might be that students considered an argumentative writing task in history more as a general writing task than as a *history* writing task. Another explanation could be found in the instrument

we used. As already mentioned by Schoonen and de Glopper (1996), the question is whether a letter of advice indeed elicits all students' knowledge. Other instruments, such as an additional interview, might be used to gain more information. An additional explanation for the low number of domain-specific recommendations might be related to our task directions. Although it was explicitly stated to give advice on how to write an argumentative text in history, the added example was about a general topic, which might have confused students. The choice for a general example was made as the interventions focused on various historical topics, and it was decided to use the same neutral examples across these topics.

5.2 Effects on text quality

Our second research question focused on the effects on text quality. Here, our hypothesis was also confirmed; the interventions had a positive effect on text quality. This outcome is consistent with earlier findings in history (e.g., Klein & Boscolo, 2016). Students in all interventions improved significantly on genre-specific aspects, compared to the control groups. Students in Intervention 1 and 2 also improved on general writing aspects, and students in Intervention 3 improved on domain-specific aspects of their texts. That the patterns of improvement found for Intervention 1 and 2 seemed more similar and somewhat different from the patterns found for Intervention 3, can be explained by the fact that Intervention 1 and 2 were more similar in goals, writing task and instruction. In contrast, Intervention 3 in Grade 10 focused on writing in the context of answering open-ended questions and the intervention also included Dutch language lessons. This intervention is a nice example of how L1 teachers and content area teachers could collaborate in order to overcome the problems of compartmentalization of subjects, and lack of transfer between subjects in secondary education.

5.3 Relation between knowledge of writing and text quality

Our third question was whether there was a relationship between students' metacognitive knowledge of writing and their writing performance after the intervention. Correlational analyses did only confirm this for Intervention 1, but not for Intervention 2 and 3. This result is not in line with earlier findings of, for instance, Klein and Kirkpatrick (2010), who found that effects of instruction on text quality were mediated by genre knowledge. Bouwer and Koster (2016) found in their study in the upper elementary grades that knowledge of lower order aspects such as punctuation, capitals, spelling and grammar and process knowledge contributed to text quality, whereas higher order aspects (e.g., content, style) did not. We did not analyze these kinds of lower order knowledge aspects, as students seldom reported them. A possible explanation might be the different age group involved in our study. Perhaps text quality in the higher grades is influenced more strongly by other factors than knowledge of writing. More research is needed to investigate the mediating role of

knowledge of writing in upper secondary education and within the domain of history and other subjects.

5.4 Strengths and limitations

We believe that the current study has several strengths. First, we investigated the effects of three different writing interventions in the content area of history, instead of just one, which benefits the generalizability of our findings. Second, these interventions were designed by teachers themselves, based on a set of design principles derived from meta-analyses of effective writing interventions. The teachers could tailor the writing instruction and writing tasks to their own context, and carried the interventions out in their regular classrooms, which contributed to the ecological validity of the study. Third, we not only examined the effects on the quality of students' texts, but also on their metacognitive knowledge of writing, by which we mean students' knowledge of what writing a good text in the domain of history entails, in terms of product features, writing processes and strategies. Since previous studies in the domain of history predominantly examined effects on text quality and/or content learning, this study adds to the existing knowledge base of writing-to-learn research. Lastly, the present study demonstrated that students' writing and knowledge of writing may improve after relatively short interventions of only five to six lessons. This finding may be reassuring for content area teachers who shy away from teaching disciplinary writing, because they believe that would be too time consuming and would go at the expense of content teaching.

However, the present study has limitations as well. One limitation is formed by the relatively small sample sizes, in particular in the control condition of Grade 10 (N = 13). With larger sample sizes smaller effects may have been found. Furthermore, in each of the three interventions students' writing performance was measured with just one writing task at pretest and posttest. This is problematic, because previous studies have shown that students' writing scores tend to differ strongly across tasks (Schoonen, 2012; Rietdijk et al., 2018). In future studies, more writing tasks are needed to estimate students´ writing skill more precisely. Finally, the effects on students' knowledge of writing and text quality were measured by posttests administered immediately after the interventions. This raises the question whether the effects we found are lasting. In future studies a delayed posttest should be used to investigate this.

5.5 Conclusion

All in all, the current study contributes to existing research on writing instruction in content-area classrooms, in particular history classrooms, by demonstrating that such instruction may not only improve the quality of students' writing, but may also have an impact on students' knowledge of writing.

ACKNOWLEDGEMENTS

This study was funded by the Netherlands Organization for Scientific Research (grant number 405-14-505). We would like to thank Arie Westerhout, Johan van Driel, Lieke Holdinga and Roderik Egberink and their students for participating in this study.

REFERENCES

Bangert-Drowns, R. L., Hurley, M. M., & Wilkinson, B. (2004). The effects of school-based writing to learn interventions on academic achievement: A meta-analysis. *Review of Educational Research*, 74(1). 29 – 58. doi: 10.3102/00346543074001029

Biancarosa C. & Snow C. E. (2006). *Reading next—a vision for action and research in middle and high school literacy: a report to Carnegie Corporation of New York* (2nd ed.). Washington, DC: Alliance for Excellent Education.

Bouwer, R. & Koster, M. (2016). *Bringing writing research into the classroom. The effectiveness of Tekster, a newly developed writing program for elementary students*. Dissertation, University of Utrecht.

Coffin, C. (2006). *Historical discourse: The language of time, cause and evaluation*. New York, NY: Continuum.

De La Paz, S. (2005). Effects of historical reasoning instruction and writing strategy mastery in culturally and academically diverse middle school classrooms. *Journal of Educational Psychology, 97*(2), 139-156. https://doi.org/10.1037/0022-0663.97.2.139

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

De La Paz, S., & Felton, M. K. (2010). Reading and writing from multiple source documents in history: Effects of strategy instruction with low to average high school writers. *Contemporary Educational Psychology, 35*(3), 174-192. https://doi.org/10.1016/j.cedpsych.2010.03.001

De La Paz, S., Ferretti, R., Wissinger, D., Yee. L., & MacArthur, C. (2012). Adolescents' disciplinary use of evidence, argumentative strategies, and organizational structure in writing about historical controversies. *Written Communication, 29*(4), 412-454. https://doi.org/10.1177/0741088312461591

Expert Group Learning Trajectories (2009). *Referentiekader taal en rekenen: De referentieniveaus*. [Reference framework language and arithmetic: The referential levels]. Enschede, The Netherlands: Ministerie van OCW.

De Oliveira, L. C. (2011). *Knowing and writing school history. The language of students' expository writing and teachers' expectations*. Charlotte, NC: Information Age Publishing.

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry. *American Psychologist, 34*(10), 906-911. https://doi.org/10.1037/0003-066X.34.10.906

Flower, L., & Hayes, J. (1980). The cognition of discovery. *College Composition and Communication, 31*(1), 21-32. https://doi.org/10.2307/356630

Graham, S., & Harris, K. (2018). Evidence-based writing practices: A meta-analysis of existing meta-analyses. In R. Fidalgo & T. Olive (Series Eds.) & R. Fidalgo, K.R. Harris, & M. Braaksma, (Vol Eds.), *Design principles for teaching effective writing*. Studies in Writing Series: Vol. 34 (pp. 13-38). Leiden, The Netherlands: Brill.

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

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

Hayes, J. R. (1996). A new framework for understanding cognition and affect in writing. In C. M. Levy & S. Ransdell (Eds.), *The science of writing: Theories, methods, individual differences, and applications* (pp. 1-27). Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.

Hillocks, G. (1986). *Research on written composition: New directions for teaching*. Urbana, IL: ERIC Clearinghouse on Reading and Communication Skills, National Institute of Education.

Janssen, T. & Overmaat, M. (1990). *Tekstopbouw en stelvaardigheid. Een onderzoek naar de effecten van twee experimentele methoden voor tekstopbouw* [Text structure and writing; A study of the effects of two experimental methods to teach text structure]. Amsterdam/Lisse, The Netherlands: Swets & Zeitlinger, Forum 8.

Kellogg, R. T. (2008). Training writing skills: A cognitive developmental perspective. *Journal of Writing Research, 1*(1), 1 – 26. https://doi.org/10.17239/jowr-2008.01.01.1

Klein, P. D. & Boscolo, P. (2016). Trends in research on writing as a learning activity. *Journal of Writing Research, 7*(3), 311-350. https://doi.org/10.17239/jowr-2016.07.3.01.

Klein, P. D., & Kirkpatrick, L. C. (2010). A framework for content area writing: Mediators and moderators. *Journal of Writing Research, 2*(1), 1-46. https://doi.org/10.17239/jowr-2010.02.01.1

Koster, M., Tribushinina, E., de Jong, P. F., & van den Bergh, H. (2015). Teaching children to write: A meta-analysis of writing intervention research. *Journal of Writing Research, 7*(2). 300-324. https://doi.org/10.17239/jowr-2015.07.02.2

Martínez, I., Mateos, M. D. M., Martín, E., & Rijlaarsdam, G. (2015). Learning history by composing synthesis texts: Effects of an instructional programme on learning, reading and writing processes, and text quality. *Journal of Writing Research, 7*(2), 275-302. https://doi.org/10.17239/jowr-2015.07.02.03

McCarthy Young, K., & Leinhardt, G. (1998). Writing from primary document. A way of knowing in history. *Written communication, 15*(1), 25-68. https://doi.org/10.1177/0741088398015001002

McCutchen, D. (1986). Domain knowledge and linguistic knowledge in the development of writing ability. *Journal of Memory and Language, 25*(4), 431-444. https://doi.org/10.1016/0749-596X(86)90036-7

McCutchen, D. (2011). From novice to expert: Implications of language skills and writing relevant knowledge for memory during the development of writing skill. *Journal of Writing Research, 3*(1), 51-68. https://doi.org/10.17239/jowr-2011.03.01.3

Miller, D. M., Scott, C. E., & McTigue, E. M. (2018). Writing in the secondary-Level disciplines: a systematic review of context, cognition, and content. *Educational Psychology Review, 30*(1), 83–120. https://doi.org/10.1007/s10648-016-9393-z

Monte-Sano, C. (2010). Disciplinary literacy in history: An exploration of the historical nature of adolescents' writing. *Journal of the Learning Sciences, 19*(4), 539-568. https://doi.org/10.1080/10508406.2010.481014

Monte-Sano, C. & De La Paz, S. (2012). Using writing tasks to elicit adolescents' historical reasoning. *Journal of Literacy Research, 44*(3), 273-299. https://doi.org/10.1177/1086296X12450445

Moje, E. (2008). Foregrounding the disciplines in secondary literacy teaching and learning: A call for change. *Journal of Adolescent and Adult Literacy, 52*(2), 96-107. https://doi.org/10.1598/JAAL.52.2.1

Mottart, A., van Brabant, P., & van de Ven, P-H. (2009). Schrijven bij diverse schoolvakken. Een verkenning. [Writing in the disciplines; An exploration]. *Levende Talen Tijdschrift 10*(4), 14-20.

Nokes, J. D. & De La Paz, S. (2018). Writing and argumentation in history education. In S. Metzger & L. Harris (Eds). International handbook of history teaching and learning (pp. 551-602). New York, NY: Wiley. https://doi.org/10.1002/9781119100812.ch21

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*. Advance online publication. https://doi.org/10.1037/edu0000237

Rijlaarsdam, G., Braaksma, M., Couzijn, M., Janssen, T., Kieft, M., Broekkamp, H., & van den Bergh, H. (2005). Psychology and the teaching of writing in 8000 and some words. In P. Tomlinson, J. Dockrell & P. Winne (Eds.), *Pedagogy—teaching for learning*. British Journal of Educational Psychology Monograph Series II: Psychological Aspects of Education—Current Trends, Number 3 (pp. 127-153). Leicester, UK: The British Psychological Society.

Schoonen, R. (2012). The validity and generalizability of writing scores: the effect of rater, task and language. In Van Steendam, E., Tillema, M., Rijlaarsdam, G., & Van den Bergh, H. (Eds.), *Measuring*

writing: recent insights into theory, methodology and practices (Studies in writing, 27) (pp. 1-22). Leiden, The Netherlands/Boston, MA: Brill.

Schoonen, R., & de Glopper, K. (1996). Writing performance and knowledge about writing. In G. Rijlaarsdam, H. van den Bergh & M. Couzijn (Eds.), *Theories, models & methodology in writing research* (pp. 87-107). Amsterdam, The Netherlands: Amsterdam University Press.

Schoonen, R., van Gelderen, A. V., de Glopper, K. D., Hulstijn, J., Simis, A., Snellings, P., & Stevenson, M. (2003). First language and second language writing: The role of linguistic knowledge, speed of processing, and metacognitive knowledge. *Language learning, 53*(1), 165-202. https://doi.org/10.1111/1467-9922.00213

Shanahan, T. & Shanahan, C. (2008). Teaching disciplinary literacy to adolescents: rethinking content-area literacy. *Harvard Educational Review, 78*(1), 40-59. https://doi.org/10.17763/haer.78.1.v62444321p602101

Torrance, M. (1996). Strategies for familiar writing tasks: Case studies of undergraduates writing essays. In G. Rijlaarsdam, H. van den Bergh & M. Couzijn (Eds.), *Theories, models & methodology in writing research* (pp. 283-298). Amsterdam, The Netherlands: Amsterdam University Press.

van der Leeuw, B., & Meestringa, T. (2011). Eisen aan schrijfvaardigheid in de bovenbouw havo/vwo. [Writing requirements in the higher grades of secondary education]. *Levende Talen Tijdschrift, 12*(2), 14-24.

van Boxtel, C. & van Drie, J. (2018). Historical reasoning: Conceptualizations and educational applications. In S. Metzger & L. Harris (Eds), International handbook of history teaching and learning (pp.149-176). New York, NY: Wiley. https://doi.org/10.1002/9781119100812.ch6

van Drie, J., Braaksma, M, & van Boxtel, C. (2015). Writing in history: Effects of writing instruction on historical reasoning and text quality. *Journal of Writing Research, 7*(1), 123-156. https://doi.org/10.17239/jowr-2015.07.01.06

van Drie, J., Janssen, T., & Groenendijk, T. (2017). Improving writing in social studies through professional development: Effects on teachers' beliefs, classroom practice and students' writing. European *Journal for Applied Linguistics, 5*(2), 273-308. https://doi.org/10.1515/eujal-2017-0009

van Drie, J. & van Boxtel, C. (2008). Historical reasoning: towards a framework for analyzing students' reasoning about the past. *Educational Psychology Review, 20*(2), 87-110. https://doi.org/10.1007/s10648-007-9056-1

van Drie, J., van Boxtel, C. & Braaksma, M. (2014). Writing to engage students in historical reasoning. In P. D. Klein, P. Boscolo, L. Kirkpatrick & C. Gelati (Eds.), *Writing as a learning activity* (Studies in writing, 28) (pp. 94-119). Leiden, The Netherlands: Brill. https://doi.org/10.1163/9789004265011_006

van Drie, J., van Boxtel, C., & van der Linden, J. L. (2006). Historical reasoning in a computer-supported collaborative learning environment. In A. M. O'Donnell, C. E. Hmelo, & Erkens, G. (Eds.), *Collaborative learning, reasoning and technology* (pp. 266-297). Mahwah NJ: Lawrence Erlbaum Associates.

Wineburg, S. S. (1991). Historical problem solving: A study of the cognitive processes used in the evaluation of documentary and pictorial evidence. *Journal of Educational Psychology, 83*(1), 73-87. https://doi.org/10.1037/0022-0663.83.1.73

APPENDIX

1. Knowledge of writing test

Dear student,
A good friend of yours really needs a good grade for history. This friend therefore asks you for help. (S)he has to write an argumentative text for history.
An example of a prompt for writing an argumentative text is: Argue for or against the statement …. (for example, It should be forbidden to sell sweets in schools).
Write an email to your friend in which you explain how to write a good argumentative text in history. So, describe how (s)he can get a high mark.
Give your friend as many recommendations as possible.
Available time: 15 minutes.

2. Examples of writing tasks and student texts at posttest

Intervention 3, Grade 10

Writing task
Statement: "During the reign of Louis XIV the king had absolute power". To what extent do you agree with this statement?
Please, write a comment on this statement in which you provide arguments. Use a maximum of 200 words.

Example of a student text
Louis XIV was a king with absolute power. In the provinces he appointed agents. They served as the eyes and ears of the monarch in the provinces. This enabled a politics of centralization, which increased the absolute power of the king.

Louis also brought the army under his control. He was no longer dependent on the nobility, and as a consequence his power grew. Furthermore, Louis also determined the religion of his people, which is characteristic of a king with absolute power.

On the other hand, the cities and regions kept their own legal regulations and other privileges, which was not characteristic of absolute power. Louis also did not succeed in pressing the nobility and the church to pay taxes.

So, in general, Louis XIV had absolute power, although a few aspects can prove the opposite. But ultimately, one may call Louis XIV a good example of absolutism.

Intervention 2, Grade 11

Writing task
Shortly after the Second World War, the Cold War started. Both the United States and the Soviet Union accused the other of being responsible for the outbreak of the

Cold War. Up to the present day, there is discussion among historians on the question who was responsible for the start of the Cold War.
The Historical Journal wishes to publish an issue fully dedicated to the Cold War. To increase the involvement of the general public, the journal organizes a writing contest. Readers are being asked to write an argumentative text in response to the claim: "The Soviet Union is solely responsible for the outbreak of the Cold War".
The winner wins a trip to Berlin and his/her essay will be published in the *Historical Journal*.
Please, use the following sources [four historical sources were added, e.g. part of the Iron Curtain Speech of Churchill, and the Truman Doctrine Speech].
Length: 300 words
Time: 50 minutes

Example of a student text
[Place & Date]
To the Editors of The Historical Journal
[Address]
[Name & address student]

Concerning: Text for writing contest

Dear Editors,

In response to your contest, I have written an essay about the claim: "The Soviet Union is solely responsible for the outbreak of the Cold War." I do not agree with this statement. I have several reasons for my conclusion.

The Soviet Union is not the only one to blame. The United States are equally guilty. The US and the Soviet Union feared each other. Both had atom bombs. With those bombs, they could harass each other. If they started a war they would both perish. The US were the first to have an atom bomb and to use it. That showed the destructive power of the bomb. Later on, the Soviet Union also had an atom bomb. Together these two countries were the most powerful countries on earth. Both wanted to be more powerful than the other was. This idea made them fearsome of each other. That's when the Cold War started.

Yet, the Soviet Union is to a large extent responsible for the outbreak of the Cold War. When WWII had ended, they overtook many countries in Eastern Europe, presumably because the Soviet Union feared a new war against Western Europe. Therefore, the Soviet Union deemed it necessary to have a strip of countries that could absorb the attack. Those countries became communist too. Western Europe felt threatened by so many communist countries nearby. The Red Army could easily reach Western Europe in no time. Something had to be done about this, according to the United States and Western Europe. Therefore, stronger and better weapons were developed. When the Soviet Union discovered this, they also started develop-

ing stronger and better weapons. The threat of an (atomic) war increased more and more.

Because of the Cuba crisis the Cold War almost erupted. The Soviet Union had installed missiles on communist Cuba. From Cuba, the Soviet Union could easily bomb the east coast of America. America wanted to prevent that at all cost and demanded that the Soviet Union would remove the missiles. This happened at the last moment. Most people were convinced that the Soviet Union was to blame for nearly starting a war. What people did not know, was that the US also had missiles near the Soviet Union, namely in Turkey. The US and the Soviet Union had agreed to both remove their missiles. So, both countries were too blame.

All-in all, the Soviet Union is not the only one to blame. The US also played a large role in the Cold War. Both countries had atom bombs and threatened each other with these. The Soviet Union possessed a large part of Eastern Europe. Both countries blamed each other, but participated just as much in the war.

With best regards,
[Name student]