



UvA-DARE (Digital Academic Repository)

Disciplinary writing

Four empirical studies on historical and philosophical literacy

Holdinga, C.C.

Publication date

2023

[Link to publication](#)

Citation for published version (APA):

Holdinga, C. C. (2023). *Disciplinary writing: Four empirical studies on historical and philosophical literacy*. [Thesis, externally prepared, Universiteit van Amsterdam].

General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

CHAPTER 2

STUDENTS' DISCIPLINARY WRITING PROCESSES: A THINK-ALOUD STUDY*

Source-based writing is a common, but difficult task in history and philosophy. Students are usually taught how to write a good text in language classes. However, it is also important to address discipline-specificity in writing, a topic likely to be taught by content teachers. In order to design discipline-specific writing instruction, research needs to identify which reading and writing activities during the source-based writing process affect students' thought process quality and text quality, as assessed by content teachers. We conducted a think-aloud study with 15 (11th grade) students who performed two source-based writing assignments, each representative of its discipline. From the data, we derived 11 activities which we analyzed for duration, frequency, and time of occurrence. Results showed that the disciplines required different approaches to writing. For philosophy, the writing process was dominant and influenced quality, leading us to conclude that philosophical thinking and writing are intertwined. For history, the planning process appeared to be paramount, but it influenced text quality only, and not the quality of the thought process. In other words, historical thinking and writing appear to be separate processes. Our findings can be used to develop strategy instruction that reinforces better writing, adapted to discipline-specific writing processes.

* Chapter 2 is based on: Holdinga, C. C., Janssen, T. M., & Rijlaarsdam, G. C. W. (2021). The relationship between students' writing process, text quality, and thought process quality in 11th-grade history and philosophy assignments. *Written Communication*, 38(4), 544-586. <https://doi.org/10.1177/07410883211028853>

1. INTRODUCTION

Over the past two decades, technological development has created a new digital world full of text, which has forced us to substantially expand our literacy skills. Essentially, this has turned us all into writers. Research has shown that writing is now a critical success factor in nearly all domains, including school and professional environments (Freedman et al., 2016). In the Netherlands, the geographical context of our study, writing skills are a requisite in secondary education. This is particularly true in the social sciences because they tend to rely heavily on written assignments (Van Drie et al., 2016). However, most content teachers tend to focus on disciplinary thinking rather than writing (De Oliveira, 2011). Teachers in the disciplines central to our study – history and philosophy – are no exception. The Dutch history exam program does not include teaching students how to communicate a path of reasoning in its formal objectives. The philosophy exam program does include the acquisition of argumentative skills among its learning objectives, but it makes no specific mention of writing skills. In general, discipline-specific writing functions mainly as a means of assessment; students are expected to demonstrate their mastery of the discipline-specific thought process through a writing assignment. Writing in this sense is not a learning activity, but rather a window to students' thought processes.

Since writing is used so extensively to judge students' progress, one would expect an emphasis on teaching writing properly and extensively in secondary education. However, content teachers appear to pay little attention to writing instruction, writing process coaching or writing task design (Gillespie et al., 2014; Mottart et al., 2009). Content courses therefore seem to contribute little to the development of students' writing skills. This seems like a lost opportunity.

Paying more attention to writing skills in content courses might have other benefits as well. Research has shown that writing promotes learning (Graham et al., 2020; Graham & Perin, 2007; Klein, 1999; Newell & Winograd, 1995) and thus may strengthen, expand, and deepen students' content knowledge (Graham et al., 2020). Hence, writing can not only be used as a tool for assessment or as a goal in itself (to learn the mechanics of writing), but also as a learning activity, since the act of writing may support the acquisition of content knowledge. However, it is important to carefully consider what type of writing assignment is appropriate (Gillespie et al., 2014), since different writing assignments can have different learning effects (Applebee, 1984).

Furthermore, rhetoric and reasoning are not only genre-specific, but also discipline-specific (Bazerman, 1981; Klein & Boscolo, 2016).

Writing in school settings often involves reading (source-based writing), particularly in disciplines that use source materials, such as history (frequent use of historical sources) and philosophy (frequent use of philosophical source texts). This is why we have focused our study on these two disciplines. Source-based writing is a 'hybrid' task. Students are asked to produce a new text on the basis of reading one or more source texts (Spivey & King, 1989). To make such assignments useful for both content learning and improving writing skills, some kind of instruction is necessary. Simply adding general writing instructions is not sufficient, since addressing discipline specificity is crucial (Bazerman, 1981; Carter, 2007).

In order to be able to design such discipline-specific writing instructions, we first needed to explore the role of discipline specificity in source-based writing. We did this by exploring the reading and writing activities of 11th grade students performing source-based writing assignments for history and philosophy, and by asking the following questions: To what extent are these activities and their distribution across the whole process task-specific and therefore discipline-specific? Can the variation in approaches be related to variation in the quality of the resulting written text (the product) and to the quality of thought about a given issue (the path of disciplinary reasoning)? Our objective was to use the answers to these questions as an aid in designing instruction for 10th and 11th grade students in both disciplines.

1.1 Theoretical background

Previous research about reading and writing processes has given us valuable insights into the strategies used by expert readers and writers. Generally, expert readers are active readers, activating prior knowledge, predicting upcoming text content, drawing inferences, answering questions, reflecting on main points, constructing personal interpretations and images, self-explaining, and problem monitoring (Bråten & Strømsø, 2011; Chi, 2000; Pressley, 2002; Pressley & Afflerbach, 1995; Pressley & Harris, 2006). Furthermore, expert readers often go back and forth through the text, trying to meet their own standard of coherence (Van den Broek & Helder, 2017).

Expert writers typically plan and revise more than novices (Flower & Hayes, 1981; Hayes & Flower, 1986), as they tend to have more topic-, discourse- and language knowledge. They are also more active in monitoring their writing process (Ferrari et al., 1998). Moreover, experts make more

recursive and flexible use of reading and writing (Mateos et al., 2008; Lenski & Johns, 1997; McGinley, 1992). Furthermore, they try to transform their knowledge, unlike novices, who tend to restrict themselves to the repetition of knowledge (Bereiter & Scardamalia, 1987). With regard to experts' texts, research has shown that good writers write longer texts (Ferrari et al., 1998).

Whether these strategies are also crucial in discipline-specific writing is an unresolved question in need of further exploration. Previous research on reading in history does tell us, however, that a core competence is so-called 'sourcing' (Brante & Strømsø, 2018). Based on a think-aloud study with novice and expert historians who were asked to read particular source texts, Wineburg (1991) distinguished three main aspects of sourcing as carried out by experts: (1) checking for corroboration (is this plausible or likely considering other sources?), (2) checking credentials by sourcing (who wrote the text and when?), and (3) checking the contextualization (when and where did this happen?). Wineburg also noted that novices tend to regard texts as 'bearers of information' and hardly notice the source's features, whereas experts interpret texts in light of the source's characteristics and regard texts as social entities. Therefore, we can conclude that reading history texts requires more than general reading skills, as sourcing is considered a core skill in the discipline of history (Brante & Strømsø, 2018).

The core competency in philosophy is philosophical thinking, which entails (1) problematization (formulating questions and problems), (2) conceptualization (reflecting on philosophical concepts), and (3) argumentation (proposing and defending one's own arguments) (Tozzi, 2012). These skills can be practiced by writing, and particularly by writing argumentative texts (Corcelles & Castelló, 2017). Because a knowledge base of philosophical concepts is crucial to philosophical thinking, source texts play a role here as well. With regard to reading such sources, previous research suggests that expert philosophers are able to chunk the information from source material into categories (different views or bodies of opinion), with the help of metacognitive skills (Concepción, 2004).

In intervention studies, educational researchers have tried to equip content teachers with a knowledge base about writing, providing guidance not only on how to teach and coach writing (supporting the process of learning to write), but also on textual genres and other writing task conditions (supporting the process of writing to learn) (Klein & Boscolo, 2016). In the field of history, considerable research has been done on the question of how to encourage students to think and write as historians (De La Paz & Felton, 2010; Van

Drie et al., 2014; Wiley et al., 2014; McCarthy Young & Leinhardt, 1998). Likewise, in science, numerous studies have explored how to best stimulate awareness of scientific discourse and practice (Hand et al., 2002; Moje et al., 2004). Such research is based on the tenet that teachers should be aware of how disciplinary epistemologies define their expectations and understanding of what they regard as 'good writing'. Furthermore, teachers should be explicit about their own understanding of discipline-specific genre conventions if they are to improve students' writing abilities (Freedman et al., 2016).

1.2 Aim of the Present Study

Given the importance of source-based writing in history and philosophy, our aim was to explore reading and writing activities for discipline-specific source-based writing in both disciplines and to identify effective patterns in these activities. These insights will serve as a basis for the development of writing instruction, aimed at improving students' discipline-specific writing skills. Our main questions were as follows: Which patterns in source-based reading and writing can be related to (a) the quality of text produced by students, and (b) the quality of students' thought processes? Are there differences with regard to the previous question depending on the discipline under consideration (history versus philosophy)?

2. METHOD

2.1 Research Design

We set up a descriptive, within-subject, think-aloud study followed by reflective interviews, with students performing source-based writing in history and philosophy. We combined online (think aloud) with off-line (reflective interviews) data.

2.2 Participants

Fifteen Dutch students (11th grade, pre-university level, 10 females and 5 males, mean age: 16.8) from seven different high schools in the Netherlands participated voluntarily. We selected relatively high achievers, as we reasoned that the processes used by high achievers could serve as models for instructional design. Since experienced teachers are believed to make accurate judgments on students' achievement level (Südkamp et al., 2012), we asked teachers of history and philosophy to select these high achievers for our study. Without further guidance, the teachers relayed our request to those who they

regarded as 'good students'. Because we wanted to include students who were high achievers in both disciplines, we checked whether the students' grades were above average for both disciplines. Informed consent was obtained from all students.

2.3 Procedure

Data was collected by means of two think-aloud assignments (one for each discipline) presented in random order and with no time limit. Students were instructed to verbalize their thoughts and feelings while performing the tasks. Subsequently, we conducted follow-up reflective interviews. All sessions took place in convenient and quiet locations determined in consultation with the student. To test the feasibility of the procedure, we conducted a trial session with two students.

At the beginning of a session, the researcher explained the procedure to the student. During the think-aloud tasks, the researcher would visualize the student's thought process by arranging sticky notes on a piece of paper. On these sticky notes, the researcher described the student's activity (e.g., "reading" or "revising"). The students were told not to pay attention to this, because the process schemes would be discussed (to check accuracy and to clarify uncertainties) with the students afterwards, in a reflective interview. When necessary, the researcher (first author) prompted the thinking-out-loud process by posing questions (e.g., "What are you thinking now?"). All sessions took 40-50 minutes and were audio recorded.

2.4 Task Construction

In consultation with history and philosophy teachers, we selected two assignments from a textbook (philosophy) and an exam (history), ensuring both assignments featured typical characteristics of their respective discipline. Other selection criteria were: (a) inclusion of a source text, (b) similarity in genre, difficulty level, and time required, and (c) performability without prior content knowledge.

Task selection occurred in several rounds. After exploratory interviews with 21 teachers to identify disciplinary focus (unpublished data), we selected two assignments per discipline and discussed these with panels of two teachers. On this basis, we decided which of the two tasks to assign for each discipline. Finally, the two selected assignments were tested with two students.

Since our aim was to use typical disciplinary assignments, the assignments themselves were not entirely similar. Again, we strove for assignments that

reflected their respective discipline; the history assignment included a question about the usability and reliability of the presented source text (“is source text A pertinent to topic X?”), and the philosophy assignment contained a statement open for discussion (“to what extent is statement X true?”). The teacher panels confirmed that the two assignments were representative of their disciplines. In this article, we will therefore refer to the differences between assignments and between the reading/writing processes when performing these tasks as differences between the disciplines.

Because the key questions in the two assignments differed, the source texts were used differently as well. The history assignment required students to use sourcing skills, while in the philosophy assignment, which focused on students’ philosophical thinking skills, the source text was merely an aid to generate arguments. Although both assignments contained a single question requiring two arguments, in the history assignment those two arguments were interdependent. By contrast, the philosophy assignment could be divided into several parts to be dealt with independently. Lastly, the philosophy assignment was somewhat shorter (279 words) than the history assignment (427 words). See Appendix A for both assignments.

2.5 Data Preparation

2.5.1 Coding Scheme

We used both theory and research data to develop a coding scheme for analyzing the think-aloud protocols. The research we used to inform the coding scheme included research on general reading strategies (Merchie & Van Keer, 2014; Rogiers et al., 2019; Schellings et al., 2006; Vandeveldel et al., 2015); research on general writing strategies (Breetvelt et al., 1994; Kuhn et al., 2015), and on general metacognition (Nelson, 1996; Veenman, 2011); research combining reading and writing (Martinez et al., 2015; Mateos et al., 2008); research on disciplinary reading and/or writing (Brante & Strømsø, 2018), history reading and/or writing (Hof, 2015; Van Drie & Van Boxtel, 2008), philosophy reading and/or writing (Corcelles & Castelló, 2015); and research on general learning and reasoning (Bisra et al., 2018; Chi, 2000; Kuhn, 1991; Kuhn et al., 2015). We started from three main activities: reading, writing and metacognition, and then further subcategorized each of these activities into more detailed activities. Reading was subdivided into initial reading, rereading and analyzing the assignment or the source. Writing was subdivided into planning, writing, reading text produced so far and reviewing. Metacognition was subdivided into self-instruction, monitoring and evaluating. The subsequent

refinement of the coding scheme was an iterative process; we developed the coding scheme further during the first rounds of coding, resulting in additions and deletions of categories within the main activities. The final scheme is presented in Table 2.1.

2.5.2 Coding

All audio-recorded sessions were transcribed into written protocols by the first author of this paper. These protocols were subsequently divided into units in which a phrase or group of sentences was identified as a single thought process or approach representing a single activity. Every unit's duration was noted in seconds. The number of uttered words was between 537 and 2063 (history: $M = 1117$, $SD = 420$; philosophy: $M = 1067$, $SD = 453$). The number of units per protocol ranged from 19 to 132 (history: $M = 64$, $SD = 27$; philosophy: $M = 64$, $SD = 34$).

Interrater reliability was established by having a second independent coder code two randomly selected protocols, containing 53 units in total ($\kappa = .78$). An intra-rater reliability of .89 was established by re-coding a data subset (20%) one week after the coding of all protocols was completed.

Table 2.1. Coding Scheme

Code	Process activity	Category	Explanation	Example
M	Metacognition	Self-instruction Monitoring Evaluating	Telling yourself what to do next Monitoring part of the process Evaluating part of the process	<i>I am going to read the text first.</i> <i>Hmm, I still don't get what is expected of me.</i> <i>Is this a good example? Yes, I think so.</i>
R	Reading		Reading (parts of the) text	<i>((reads))</i>
RR	Rereading		Rereading (parts of the) text	<i>((rereads))</i>
TA	Task analysis	Task representation	Orienting yourself to the task or interpreting the assignment	<i>So, I need to give two examples.</i>
		Taking notes	Writing down keywords	<i>((takes notes))</i>
		Marking key words	Underlining or marking key words	<i>((marks key words))</i>
SA	Source analysis	Identifying source features	Noting author-specific phrasing and intentions and putting the source text in a context	<i>I see the text was written by a monk.</i>
		Drawing inferences	Making inferences or deductions based on the source introduction and source text.	<i>It says here that a kingdom was destroyed; that implies they were fighting.</i>
		Evaluating source content	Evaluating source text content, e.g. deciding what is important	<i>I think this part of the text is important.</i>
P	Planning	Generating	Generating ideas, reasons, arguments, counterarguments, rebuttals	<i>An example of altruism is giving up your seat to a senior citizen on the bus.</i>
		Organizing	Organizing content, connecting content elements	<i>So, this is an example that counters the statement.</i>
		Goal-setting	Setting clear goals	<i>Now I have to explain why the source is less suitable for the other chapter.</i>
PAU	Pausing		Silence of five seconds or longer	<i>(0.05)</i>

Table 2.1. Coding Scheme (Continued)

Code	Process activity	Category	Explanation	Example
W	Writing	Translating Transcribing Text	Mental text construction Text stated out loud Actual text written down	
RT	Reading text produced so far		Reading (parts of the) text produced so far	<i>((reads text produced so far))</i>
REV	Reviewing	Evaluating Revising	Evaluating translations or written text Changing translations or written text	<i>Wait, no, that's not the right word.</i> <i><-deleted word(s) or sentence></i>
O	Other	Researcher Interaction Affective responses Other	Researcher is speaking to the participant Interaction with the researcher Positive or negative responses from the participant Unclear	<i>Try to think out loud.</i> <i>Is it ok to use a marker?</i> <i>((laughs))</i> <i>Uh.</i>

2.5.3 Assessment of Text Quality

The quality of the students' texts was assessed using a holistic comparative method (Comproved online tool, <http://www.comproved.com>) by four panels of judges: (1) philosophy teachers ($N = 6$), (2) history teachers ($N = 6$), (3) L1 teachers assessing history texts ($N = 10$) and (4) L1 teachers assessing philosophy texts ($N = 5$). Since our goal was to determine relative quality, we opted for a comparative assessment method that would reliably result in a ranking of texts and an ordinal scale. Comparative judgment has proven to be a valid, reliable assessment method (Verhavert et al., 2019). Assessors made 30 random comparisons of texts, each time comparing two texts and deciding which text they thought was better, without specific criteria. All assessors were qualified secondary education teachers. We decided to include panels of L1 teachers on the assumption that these teachers would focus mainly on general text quality, whereas content teachers would focus more on content. The text quality assessment resulted in four rankings of the students' texts, one by each panel of judges (judges' reliability between .80 and .86). The rankings by content teachers and L1 teachers correlated (history and L1: $r = .78, p = .001$; philosophy and L1: $r = .60, p = .019$), which led us to conclude that the teachers all used similar quality criteria (per discipline) to assess the students' texts. We decided therefore to continue our analysis using only the content teachers' rankings, as we believe that in a study on discipline-specific writing, content teachers comprise the most valid rater group. We now know that content teachers partially agree with language teachers but that there is domain-specific variance.

2.5.4 Assessment of Thought Process Quality

Our second quality assessment focused on the students' thought process, based on the assumption that there might be a divergence between students' thought process and the text they produced. Thought process quality was assessed by two panels of judges: a panel of philosophy teachers ($N = 3$) and a panel of history teachers ($N = 3$). All assessors were qualified teachers in secondary education and/or teacher trainers. Each rater independently ranked the 15 protocols on thought process quality, disregarding the resulting text, holistically, without having been given any specific criteria ("Rank the protocols in order of the quality of the student's thought process"). We asked the judges to explain their rankings. Ranking correlations varied from .62 to .89 for pairs of history raters, and from .56 to .74 for pairs of philosophy raters.

For further calculations, we created a definitive ranking using median rank scores.

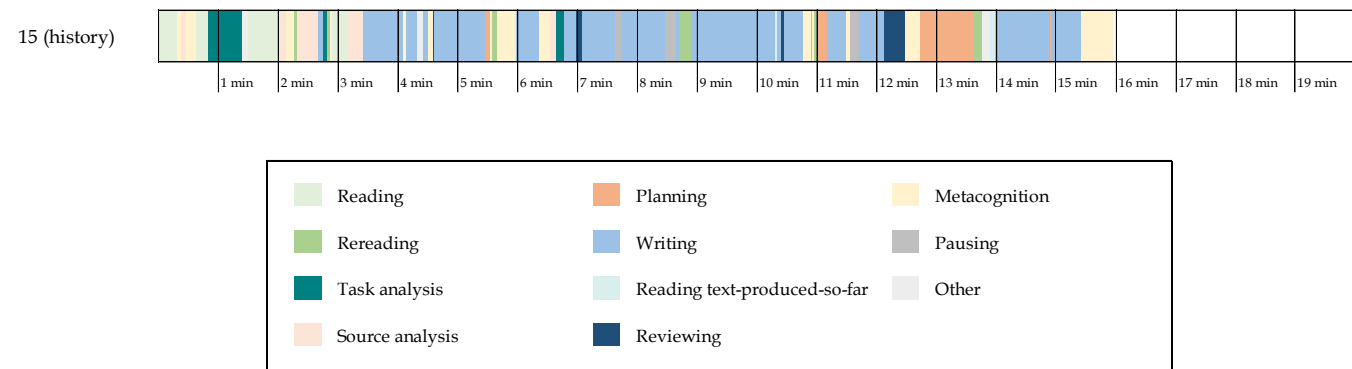
2.6 Data Analysis

To obtain an overall insight into each student's thought process, we analyzed the coded activities for duration and frequency, with duration referring to both absolute duration (in seconds) and relative duration (percentage of total time). Furthermore, we analyzed the number of words written (per writing spurt). As Figure 2.1 shows, we then visually represented the whole process (activities, duration) as process schemes, with colors representing the main activities from the coding scheme, and solid black lines indicating time units measuring one minute.

To refine our analysis, we subdivided the students' processes into three main segments, each covering approximately 33% of the total time spent on the process, and representing the beginning, middle and end of the process respectively. This allowed us to further explore at what point in the process a particular activity appeared to have an impact, since earlier research had shown that it is not only relevant what occurs, but also when in the process it occurs (Rijlaarsdam & Van den Bergh, 1996, 2006).

Furthermore, we calculated correlations between the process variables and the quality of the student's text and thought, both overall and for each main segment. These correlations served to spotlight aspects of the various processes that were sensitive to variation in text quality or thought process quality. When dealing with students who showed large discrepancies between text and thought process quality, we used these correlations as a starting point to conduct a qualitative analysis of their process and text.

Figure 2.1. Example of a Participant's Process Scheme



3. RESULTS

3.1 Correlations Between Process Activities and Quality

As Table 2.2 shows, we found significant correlations between process variables and text quality and thought process quality in both subjects, despite our limited number of participants. Process variables that appeared to influence quality included aspects of reading, source analysis, planning and pausing, writing, reading text produced so far, reviewing, metacognition, and total duration. We subsumed analyses of pausing under the planning category because pausing was associated with the idea generation process. This association was confirmed by significant correlations between pausing and planning variables (frequencies: $r = .46, p < .001$; percentage of time: $r = .30, p = .001$, duration: $r = .41, p < .001$).

Overall, it was clear that the students who spent more time thinking produced higher-quality texts (history: $r_s = .64, p = .010$; philosophy: $r_s = .54, p = .037$). As the students were not given a time limit, they took anywhere from 6 minutes 10 seconds to 20 min. 10 secs. (history: $M = 12.1, SD = 4.3$; philosophy: $M = 11.7, SD = 4.3$). We also noted different patterns for the two assignments. In the history assignment, we observed that 11 process variables were associated with text quality, three of which were also connected with thought process quality. None of the variables was uniquely connected with the thought process. Most of the variables that affected text quality were planning variables; students who planned more and spent more time on planning produced higher quality texts. In the philosophy assignment, text quality and thought process quality were more closely linked; eight variables had a bearing on text quality, four of which were also connected with thought process quality. Most of the variables affecting quality were writing variables.

Table 2.2. Overview of Correlations Between Writing Process Variables, Text Quality and Thought Process Quality

Variable	History				Philosophy				
	Text quality		Thought quality		Text quality		Thought quality		
	r_s	p	r_s	p	r_s	p	r_s	p	
Total	Duration	.64	.010**			.54	.037*		
M ^a	Frequency	.55	.033*	.53	.041*				
R	Duration					.52	.047*		
	% of time	-.68	.005**	-.55	.032*				
SA	% of time	-.53	.043*						
P	Duration	.64	.011*	.63	.012*				
	% of time					-.53	.044*		
	Frequency	.62	.013*						
PAU	Duration	.58	.024*						
	% of time	.55	.035*						
	Frequency	.57	.027*						
W	Duration	.55	.035*			.89	.000**	.63	.012*
	% of time					.54	.038*		
	Frequency					.56	.029*	.63	.012*
	Number of words	.67	.007**			.92	.000**	.61	.016*
RT	Duration							.57	.025*
	Frequency					.62	.014*	.67	.007**
REV	Duration							.52	.048*

Note. M = Metacognition, R = Reading, SA = Source Analysis, P = Planning, PAU = Pausing, W = Writing, RT = Reading Text produced so far, REV = Reviewing.

^aCategory: Evaluating.

3.2 Metacognition

Metacognitive activities were generally brief ($M = 4.3$ s, $SD = 2.1$ s). The frequency of metacognitive segments varied markedly from participant to participant. For example, Student 8 expressed 35 metacognitive thoughts while carrying out the history assignment, whereas Student 10 expressed such thoughts only three times (Figure 2.2).

Metacognition frequency correlated with the total number of segments ($r = .46, p = .011$), which indicates that students whose protocols took longer also engaged in more metacognitive activities. As Table 2.2 shows, in the history assignment, evaluation frequency correlated with both text quality ($r_s = .55, p = .033$) and thought process quality ($r_s = .53, p = .041$). This indicates that students who evaluated more frequently during the process, produced higher quality texts and had higher quality thought processes.

Table 2.3 depicts the correlations between aspects of metacognition and quality in the three main segments. The Table illustrates that students who produced higher quality texts and thought processes evaluated more often, particularly during Segments 1 and 2. For philosophy, metacognition frequency only correlated with thought process quality during Segment 1 ($r_s = .61, p = .016$).

Metacognitive activities mostly served as a catalyst for switching to another activity (80%) rather than an interruption of one and the same activity (20%). In the history assignment, metacognitive activities tended to prompt the student to switch to rereading. This was the case in all three of the main segments. However, each segment featured a different secondary focus: during Segment 1, the students' secondary focus was on task analysis; during Segment 2 on writing; and during Segment 3 on reading the text they had produced so far, and on reviewing and pausing.

In the philosophy assignment, metacognitive activities prompted fewer switches to rereading the assignment and source text, and more switches to writing, irrespective of the segment. Our analysis of the separate segments of the philosophy assignment yielded a pattern similar to the pattern we found in history, but with different secondary foci: during Segment 1, the students' secondary focus was on rereading and task analysis; during Segment 2 on planning; and during Segment 3, on reading what they wrote so far. These differences in main and secondary foci between the two disciplines indicate that the assignments may trigger different approaches.

3.3 Reading

The concept of 'reading' contained various kinds of reading activities: reading, rereading, task analysis and source analysis. The only correlations for quality we found were with reading and source analysis (see Table 2.2). Table 2.4 presents the correlations between reading variables and quality, analyzed for each of the three main segments.

Table 2.3. Correlations Between Metacognitive Activities, Text Quality and Thought Process Quality

Variable		History								Philosophy							
		Text quality				Thought quality				Text quality				Thought quality			
		0	1	2	3	0	1	2	3	0	1	2	3	0	1	2	3
Monitoring	Frequency		+					+									
Evaluating	Frequency	+	+	+				+	+								+

Note. 0 = total process, 1 = Segment 1, 2 = Segment 2, 3 = Segment 3.

Table 2.4. Correlations Between Reading Activities, Text Quality and Thought Process Quality

Variable		History								Philosophy							
		Text quality				Thought quality				Text quality				Thought quality			
		0	1	2	3	0	1	2	3	0	1	2	3	0	1	2	3
Reading	Duration									+	+						
	% of time	-	-			-	-										
Source analysis	% of time	-		-													

Note. 0 = total process, 1 = Segment 1, 2 = Segment 2, 3 = Segment 3.

3.3.1 Initial Reading

In terms of absolute reading time, there was a significant difference ($t = 6.82$, $p < .001$) between the two disciplines (history: $M = 106.47$ s, $SD = 15.55$ s; philosophy: $M = 80.93$ s, $SD = 6.79$ s). This difference can be explained by the length of the respective assignments: at 427 words, the history assignment was longer than the philosophy assignment (279 words). When comparing relative duration, the difference between the disciplines disappeared (history: $M = 16.6\%$, $SD = 6.6\%$; philosophy: $M = 13.0\%$, $SD = 4.4\%$; $t = 1.97$, $p = .069$). For history, the percentage of time spent on reading had a negative effect on text quality ($r_s = -.68$, $p = .005$) and thought process quality ($r_s = -.55$, $p = .032$). By contrast, in philosophy reading process duration appeared to be a quality indicator, that is, students who spent more time reading produced higher quality texts ($r_s = .52$, $p = .047$). Obviously, this was observed only during Segment 1 (Table 2.4), because initial reading only takes place at the beginning of the task.

The students dealt with the initial reading phase in various ways. Some approached this part of the task very straightforwardly, reading without pause or interruption and seemingly no goal orientation. After this first readthrough, some students struggled to understand the text and had to

reread it, but others seemed to grasp the ideas expressed in the text in a single reading and were able to start generating a response immediately afterwards. For example, Student 6 (Figure 2.3) read the text without any interruption, and then immediately started the planning phase. Other students took a more complex approach and alternated reading with rereading, task analysis or source analysis. This approach might point to a desire to fully understand the text, for example in the case of participants who immediately wanted to interpret and analyze what they read, as Student 15 did. Alternatively, a more complex approach might also point to reading problems, with participants failing to understand or misinterpreting parts of the text, as seemed to be the case for Student 8. However, we did not find a relationship between reading approach and text quality or thought process quality.

3.3.2 Rereading

Some students reread the assignment and the source text more than once and spent a relatively long time on this, while others did not reread at all, as the process schemes in Figure 2.4 illustrate. For example, Student 11 reread the text several times in both disciplines, whereas student 13 did not reread in either of the disciplines. Although these examples might indicate 'individual preference', other students took a different approach depending on the discipline. For example, Student 7 adjusted their rereading behavior to the discipline (or the assignment).

When we analyzed the rereading activities for the three segments of the process, we found that both duration and percentage of time spent were stable across all three segments in the history task, but were not stable in the philosophy task. In philosophy, students spent a significantly ($t = 3.34, p = .005$) lower percentage of the time rereading ($M = 0.91\%$; $SD = 1.06\%$) than in history ($M = 2.93\%$; $SD = 0.63\%$). However, in neither discipline (neither overall, nor in a specific segment) was the time they spent rereading connected to the quality of the text they produced or to the quality of their thought process.

We also noted that some students took more time to reread than to do a first readthrough. These students often did this for both assignments, which we took to indicate a 'personal strategy'. This assumption was supported by the correlation between the percentage of time spent rereading for history and for philosophy ($r = .51, p = .050$), which occurred mainly during Segment 3 ($r = .61, p = .016$).

Figure 2.3. Different Approaches to Initial Reading

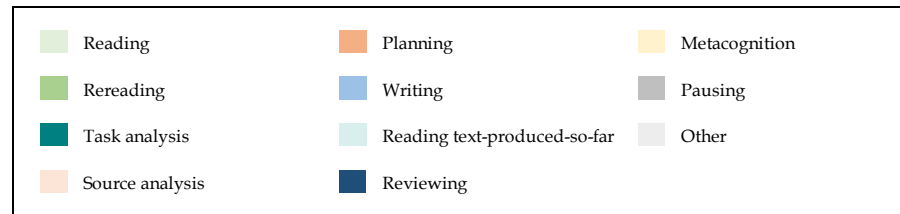
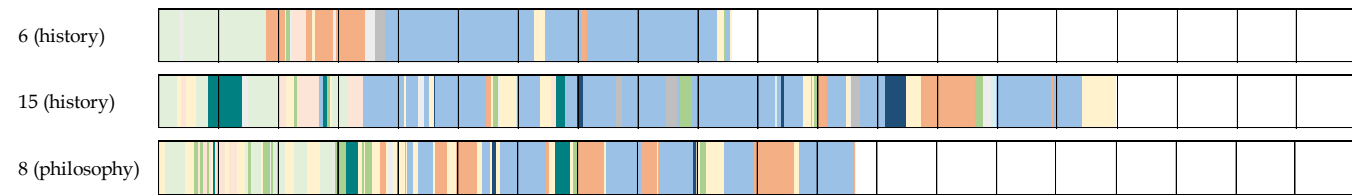
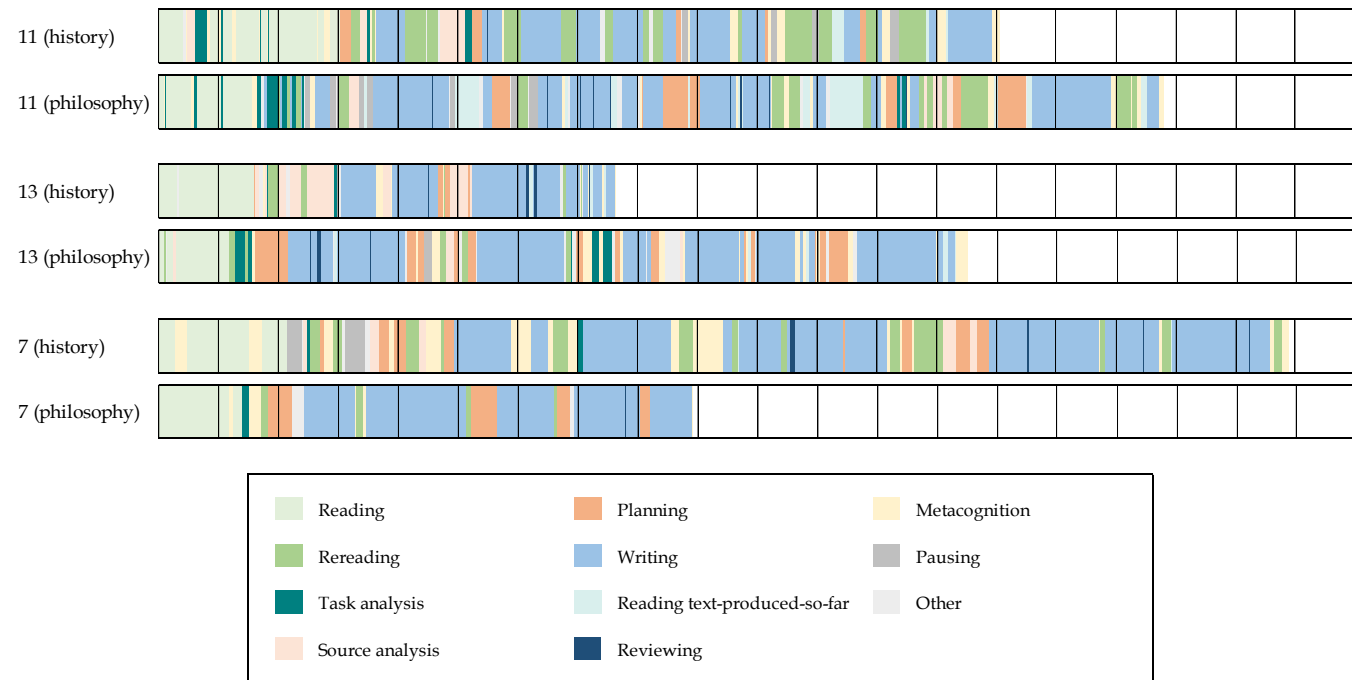


Figure 2.4. Differences in Rereading Behavior



3.3.3 Source Analysis

Some students analyzed the source text extensively, while others did not. Overall, students spent a significantly larger percentage of the time ($t = 7.94$, $p < .001$) analyzing the history source text ($M = 7.4\%$, $SD = 4.27\%$) than the philosophy source text ($M = 1.58\%$, $SD = 1.87\%$). This difference was observable in Segments 1 and 2.

The percentage of time spent on source analysis negatively correlated with text quality for history ($r_s = -.53$, $p = .043$). In other words, students who performed better on text quality seemed to have spent less time reading and analyzing the source. This applied only to Segment 2 (see Table 2.4).

3.4 Writing

For our research we defined the concept of 'writing' to include planning, pausing, writing, reading text produced so far and reviewing. Table 2.5 shows the correlations between writing variables and quality in the three main segments.

Table 2.5. Correlations Between Writing Activities, Text Quality and Thought Process Quality

Variable	History						Philosophy						
	Text quality			Thought quality			Text quality			Thought quality			
	0	1	2	3	0	1	2	3	0	1	2	3	
Planning	Duration % of time	+			+							+	
	Frequency	+							-			+	
Pausing	Duration % of time	+	+										
	Frequency	+	+			+							
Writing	Duration % of time	+		+					+	+	+	+	+
	Frequency	+							+	+	+	+	+
	No. of words	+		+					+	+	+	+	+
	Duration % of time								+	+		+	+
Reading text produced	Duration % of time					+			+	+		+	+
	Frequency					+			+	+		+	+
Reviewing	Duration % of time					+		+	+	+		+	+
	Frequency					+			+				

Note. 0 = total process, 1 = Segment 1, 2 = Segment 2, 3 = Segment 3.

3.4.1 Planning and Pausing Time

We analyzed the planning and pausing phases for duration, percentage of total time spent, and frequency. We saw significant differences between the

disciplines ($t = -2.35, p = .034$) in terms of the percentage of time spent on planning (history: $M = 9.27, SD = 4.77$; philosophy: $M = 14.35, SD = 7.68$). Students spent relatively more time planning for philosophy than for history. For philosophy, the relative time spent on planning correlated negatively with text quality ($r_s = -.53, p = .044$). Another significant difference ($t = -2.28, p = .039$) between the two disciplines pertains to how frequently the students paused (history: $M = 1.80, SD = 1.86$; philosophy: $M = 4.07, SD = 4.65$).

Most of the planning and pausing aspects seemed to influence text quality in history, but not in philosophy. In the history assignment, the students who planned more often and for longer periods of time, and who paused more frequently and for longer periods of time, produced better texts. The duration of planning also correlated positively with the quality of their thought process ($r_s = .63, p = .012$).

In our analysis of the interaction between generating ideas and writing, we noticed that some of the students seemed to generate while writing, instead of writing down what they had first generated. Student 14 remarked upon this, saying: "This pops up while I'm writing." As a consequence, their process included very few idea-generating sections (Figure 2.5). It seems the writing process fostered their thought process. Student 15 also displayed no obvious idea-generating process starting the writing process (see Figure 5).

There was no connection between this strategy and quality, although we did find a negative correlation between percentage of time spent on planning and text quality for philosophy ($r_s = -.53, p = .044$). Students who spent relatively less time on planning wrote better texts. This was not connected to a specific segment.

3.4.2 Writing Time

The time the students spent on the writing process varied widely in both disciplines (history: $M = 297.5$ s, $SD = 144.1$ s; philosophy: $M = 287.1$ s, $SD = 109.5$ s) and correlated strongly with the length of the text they produced (history: $r = .88, p < .001$; philosophy: $r = .94, p < .001$). In other words, the students who spent more time writing also wrote longer texts. We also observed that the students who produced higher quality texts seemed to spend more time writing (history: $r_s = .55, p = .007$; philosophy: $r_s = .89, p < .001$) and produced longer texts (history: $r_s = .67, p = .007$; philosophy: $r_s = .92, p < .001$). Text length ($r_s = .61, p = .016$) and writing time ($r_s = .63, p = .012$) also correlated with thought process quality for philosophy, but not for history.

In 16 out of 30 cases, the students started writing as early as Segment 1, although in six of the cases this only involved writing 'starter sentences' (e.g., rephrasing the question). Conversely, in 10 out of 16 cases this early writing contained actual content (history: three students; philosophy: seven students). For philosophy, students who spent a larger percentage of the time writing in Segment 1 scored better on text quality ($r_s = .72, N = 15, p = .003$) and thought process quality ($r_s = .59, N = 15, p = .022$).

3.4.3 Writing spurts

In our analysis of the students' writing in terms of frequency of spurts and number of words per spurt, we found that most students wrote in spurts of 13-15 words on average (history: $M = 12.8, SD = 5.5$; philosophy: $M = 15.4, SD = 6.4$). Some students managed to write the full text in only a few (longer) spurts of writing. This happened in both disciplines. It seems that these students had a personal approach to writing. For philosophy, the frequency of writing spurts correlated with text quality ($r_s = .56, p = .029$) and thought process quality ($r_s = .63, p = .012$), meaning that the students who wrote in more spurts, produced higher quality texts and had higher quality thought processes. This was visible in Segments 1 and 3. These findings can be substantiated by the negative correlation between the number of writing spurts and mean number of words per spurt ($r = -.60, p < .001$), i.e., the more spurts, the shorter their length. This pattern of numerous short spurts of writing might be an indicator of recursivity.

3.4.4 Reading Text Produced So Far

Some students read the text they produced so far quite often and some did not reread what they wrote at all. This finding was true in both disciplines. Reading text produced so far could be categorized as a strategy to either (1) generate new content or continue writing (76% of the cases) or (2) review or evaluate what had been written so far (24% of the cases). Students who read what they wrote so far early on in the process (in Segments 1 and 2) mainly did this as a strategy to create content. For philosophy, the frequency of reading text produced so far correlated to text quality ($r_s = .62, p = .014$) and to thought process quality ($r_s = .67, p = .007$). For history, there was no overall correlation between reading text produced so far and quality. However, in Segment 2, there was a correlation between duration, percentage of total time spent, and frequency of reading text produced so far with thought process quality (see Table 2.5).

Figure 2.5. Process Schemes of Students Who Generated Ideas While Writing



3.4.5 Reviewing

Some students did not review (revise/edit) their text at all, while others reviewed it quite often. This occurred in both disciplines. Reviewing occurred both during the writing process (in Segment 1 or 2) (history: 46%, philosophy: 54%), and at the end of the process (in Segment 3) (history: 56%, philosophy: 44%). Overall, the reviewing process was connected to thought process quality in philosophy but not in history (see Table 2.2). The correlation between the duration of the reviewing process and thought process quality ($r_s = .52, p = .048$) might indicate that the students' thought process was aided by the reviewing process. In the history task, reviewing was connected to thought process quality in Segment 1 (see Table 2.5). In the philosophy task, reviewing was also connected to text quality, but again, only in Segment 1. The latter two findings were linked to early writing, as students who only started writing in Segment 2 obviously would have had nothing to review in Segment 1.

3.5 Indicators of Good Writing

Since text quality and thought process quality correlated for both disciplines (history: $r_s = .81, p < .001$; philosophy: $r_s = .61, p = .015$), students with higher quality thought processes generally produced higher quality texts. Figure 2.6 and Figure 2.7 show the data points for both correlations, with 15 being the highest and 1 the lowest possible quality rank score. Next, we marked the outliers, i.e., the students who 'underachieved' in the sense that they showed high-quality thinking, yet produced a relatively weak text. We compared their process schemes with those of students who performed better on the writing task (Figure 2.8). The underachievers' performance will be discussed in more detail in Sections 3.5.1 (History) and 3.5.2 (Philosophy).

Figure 2.6. Relationship Between Text Quality and Thought Process Quality for History

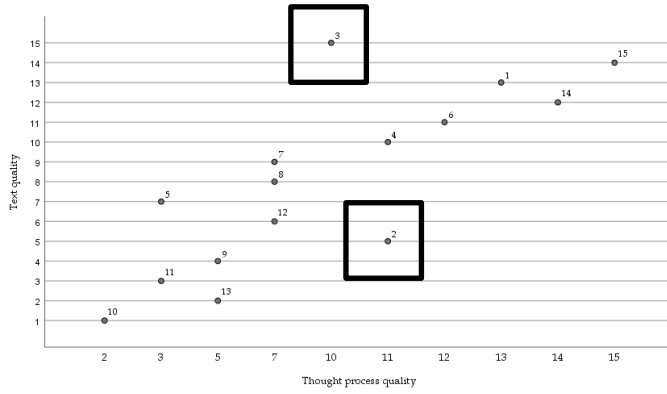


Figure 2.7. Relationship Between Text Quality and Thought Process Quality for Philosophy

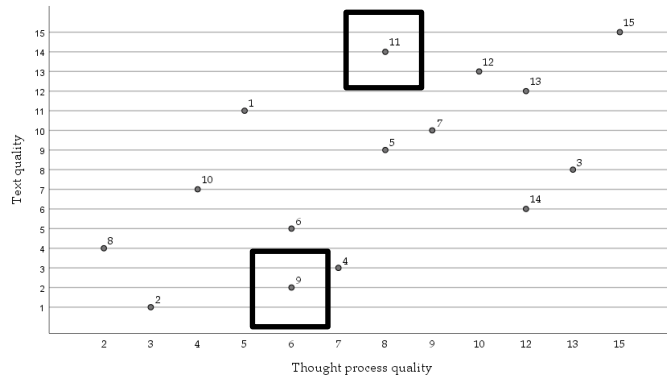
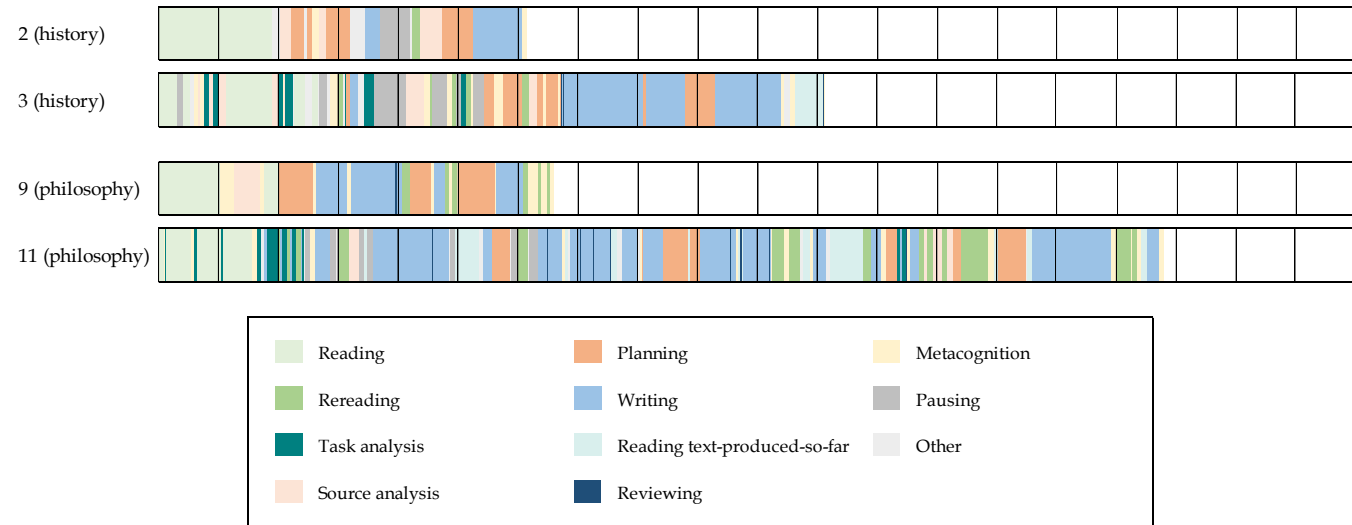


Figure 2.8. Process Scheme Comparison: Underachievers (2, 9) vs. Good Writers (3, 11)



3.5.1 History

Student 2 wrote a relatively weak text (scored 5 out of 15), although the quality of their thought process was considered relatively high (scored 11 out of 15). After a reading phase without interruptions, this student analyzed the source text, alternating with generating ideas. Next, the student wrote down the phrase "Chapter 1", then paused, thought about and analyzed the source text more extensively, and eventually wrote down their complete answer in a single 21-word spurt of writing. After that, the student said: "I think that's it". Their process showed decent procedural knowledge of source analysis as acquired in history class. For example, immediately after reading the assignment they started identifying source features:

Student 2 *First, I would look at who wrote the source text. I see that the text was written by a monk, so the second part of the question applies here.*

However, the student's text referred to this source analysis very succinctly, which resulted in a relatively short text that was not representative of the arguments the student developed in their thought process:

Student 2's text *Chapter 1: less usable, written down almost 100 years after the event, biased.
Chapter 2: usable because the source text was written by a monk.*

Student 3's thought process quality (scored 10 out of 15) was similar to that of Student 2, but they wrote the best text in this sample. As Figure 2.8 shows, this student's process was more complex. Their reading phase was interrupted by note taking and source analysis, which continued during the second segment. According to one of the assessors of thought process quality, the thought process score suffered from the student's relatively drawn out reading process. The student struggled to make sense of the assignment and the source text. Consider the following excerpt, recorded five minutes into the process:

Student 3 *Um, honestly, I still don't understand the assignment
((rereads)) So, for which chapter is the source more useful.
(0.05)
Ah, so I need to choose, yes, I have to make a choice.*

Furthermore, Student 3 did not use explicit procedural knowledge in their process, for example, they did not explicitly mention any of the source text's features. However, when this student appeared to start feeling confident

about their reading and reached their ‘standard of coherence’ (Van den Broek & Helder, 2017), they managed to generate and formulate an answer relatively fast in five spurts of writing, interspersed with revising errors. At the end of the process, they reread their entire text and revised only a minor error. The resulting text was correct, concise and well-structured.

Student 3’s text *The source is useful for the chapter on a monk’s world view, because he does not objectively relate the course of the war, he says "Oswiu appealed to God’s grace and help" and it turns out that his army wins. So, while this information it is not useful to determine the actual course of the war, it is useful for representing the monk’s view of the world, namely that you have to believe in God and if you do not, you will lose, no matter how strong your army is.*

3.5.2 Philosophy

In the philosophy task, Student 9 showed average thought process quality (scored 6 out of 15) and produced one of the weakest texts (scored 2 out of 15). Student 9 analyzed the source text during the reading phase, immediately after reading it. Subsequently, they displayed a three-step, non-recursive process, generating an answer to the first part of the question and phrasing this in three spurts of writing. The student tried to generate two separate arguments and wrote these down in one spurt of writing each. Lastly, they reread the assignment and checked the text for completeness. For this student, the discrepancy between thought quality and text quality appeared to be caused by the transformation of their thought process into writing. Consider the following example:

Student 9 *((reads)) An example that supports the statement that people are selfish. An example is people who donate 100 euro to a charity for orphans, although they don’t have any connection to those children. They think they do something good, but nine times out of ten, people are only concerned with making themselves feel better. So, that’s what I will write down.
((transcribes)) Example: people who donate money.
((writes)) e.g., people who donate money.*

Student 9’s text ended up being short and lacking any of the good explanations they gave while thinking out loud:

Student 9’s text *People are not always selfish. I think that many actions are selfish, but if an individual has ties with another ‘individual’, it is not just for themselves. e.g., people who donate money.
example: parents who take care of their children are not selfish.*

By contrast, Student 11, who also displayed an average thought process (scored 8 out of 15), produced a much higher quality text (scored 11 out of 15). During the reading phase, this student marked key words and analyzed the assignment by taking notes. Student 11 switched between activities often, sighed from time to time, and reread the source text over and over, which we took to mean they found the assignment difficult. This student started writing early (in Segment 1) and the writing process itself was fragmented (18 spurts). The result was a well-structured and quite lengthy text. At the end of the process, Student 11 evaluated their text based mainly on content.

Student 11's text *People are just selfish, because when you leave money to someone, for example, you only leave this to your children, because, out of self-interest, you don't want this to go to waste (as in: your hard work has been for nothing / your money goes to someone else). By giving your money to your children, you protect your own reputation (this is your only way of maintaining control over this after your death). So, this is pure self-interest.*
People are not just selfish, because sometimes people work hard especially for their children. They don't leave their money to their children to protect their own reputation, but to give their children opportunities and so indirectly pass on your genes.
People are not just selfish, because we are social animals. We need each other to survive and we are willing to help each other. This is evident, for example, from the fact that we have children and are willing to put money and effort into someone else's life, as described in the source text.

4. DISCUSSION AND CONCLUSIONS

The present study explores which reading and writing activities during source-based writing about history and philosophy affect (text and thought process) quality as assessed by content teachers. We approached this topic by conducting a think-aloud study with 15 participants (11th grade students) who performed a source-based writing assignment in both disciplines. From all the data gathered, we derived 11 activities, which were subsequently analyzed for relative and absolute duration as well as for frequency of occurrence and time of occurrence (beginning, middle or end of the process, i.e., during Segment 1, 2 or 3).

The results show that students approached the two assignments somewhat differently in terms of the three main activities: metacognition, reading and writing. For history, metacognition often acted as a bridge to rereading the assignment or source text, whereas for philosophy, metacognition more often segued into writing. In terms of reading, students spent more time rereading (in Segment 2) and more time analyzing the source text when doing the history assignment than when working on the philosophy assignment. With

regard to writing, the students seemed to tackle the history and philosophy assignments in similar ways, although they spent more time planning and pausing more frequently when working on the philosophy assignment.

Furthermore, the students' activity patterns and these patterns' relationship to quality differed for the two disciplines. A possible explanation for this could be that the two assignments require different dispositions of the writer towards the topic (Galbraith & Baaijen, 2018). For philosophy, the writing process was dominant and influenced both text and thought process quality, indicating interwovenness. Yet for the history assignment, the planning process seemed to be most important, but influenced only text quality. Despite a correlation between the two quality measures as a whole, the quality of text and thought process appeared unconnected. Thought process quality (and its correlation with text quality) thus depends on other factors than the ones we included in our analysis.

Overall, the metacognition and reading processes seemed to have only marginal impact on quality, even though other research has shown that expert readers tend to have a more active, recursive reading style (Pressley & Harris, 2006). However, our study did not support such a correlation between the students' reading approach and the quality of their writing or thought process. However, this might be due to the relatively short reading phase. Another possible explanation for our findings could be that all the students eventually managed to reach their 'standard of coherence' (Van den Broek & Helder, 2007), since it would have been almost impossible to write down an answer without reaching a level of understanding first. Monitoring and evaluating were influential in the beginning and the middle of the process (i.e., in Segments 1 and 2), when reading takes place. This indicates that an active reading style might be beneficial anyway. With regard to rereading, we saw that the need to reread the source text was greater in history than in philosophy. This was probably due to the nature of the assignment, since source analysis was a prerequisite for history, whereas it was not for philosophy. However, the rereading behavior in history could also be interpreted as an indicator of recursivity in the process. Some students spent more time rereading than they spent reading initially; this could suggest that those students were using rereading as a coping strategy to generate ideas, for instance, when their thinking or writing process stagnated (Bereiter & Scardamalia, 1987).

Regarding planning, it is known that writers can use a variety of methods to retrieve and generate ideas. These methods include the generation of ideas prompted by reading the assignment, the activation of new ideas when ideas

are translated into text, or the development of ideas while structuring or revising the text (Van den Bergh & Rijlaarsdam, 1999). Looking at the negative association between time spent on planning and text quality for philosophy, we therefore recommend that students working on a philosophy assignment try to generate ideas while writing. However, for students carrying out a history assignment, we recommend planning carefully and writing down answers after a thorough planning phase in order to reinforce a recursive pattern of planning and writing.

Overall, the writing process seems to be crucial for philosophy. Students who wrote relatively high-quality texts also produced longer texts, spent more time writing, and reviewed more frequently, which are all characteristics of expert writers (Ferrari et al., 1998). Moreover, the writing process seemed to enhance the thought process; these processes were closely intertwined. This is corroborated by the finding that starting the writing process during the first segment was associated with better texts: a finding that suggests writing and thinking co-develop.

To explore indicators for good writing, we also conducted a qualitative analysis of 'underachievers', i.e., students whose thinking and reasoning was sound, but who wrote relatively weak texts. Our analyses showed that these students tended to write very succinctly and that their texts lacked proper linking phrases and essential explanations that these students had voiced in their thought process.

Our conclusions must be viewed in light of the limitations of the present analysis. In the explanation we gave of our methodology, we discussed our reasons for presenting the differences between the assignments as differences between the two disciplines. We carefully selected assignments that were representative of their respective disciplines, implying that the differences in the assignments were indicative of actual disciplinary differences. However, since we used only one assignment per discipline, our results might also be attributable to the characteristics of these assignments. We are aware that this limits the validity of study, but we still believe that our results are very useful because our research design explored students' intra-individual differences. These differences can be used to inform instructional design, which was our end goal.

A second limitation has to do with the scale of our study. Our small-scale design (i.e., 15 students) with high-achieving students for both disciplines allowed us to make an in-depth analysis of activities, but prevented us from drawing more general conclusions. Also, the think-aloud method has side-

effects that can play a role when researching metacognitive activities. Despite our precautions, students might have felt unsafe, and this in turn might have affected the data we collected and thus induced a possible bias. This is one of the well-known limitations of the think-aloud method (e.g. Ericsson & Simon, 1980).

Still, we believe our results are valuable, as our study corroborates the conclusions of other research regarding process characteristics in source-based writing (e.g. Mateos et al., 2008). Moreover, by including two different disciplines, we were also able to obtain an understanding of discipline-specific writing. And finally, our reflexivity as researchers (Berger, 2015) also played a role. The research team do not consider themselves true experts in the field of history or philosophy, because these fields were not the main scope of our study. However, our position as writing experts provided us with an opportunity to analyze discipline-specific writing for either discipline from a more distant point of view.

As far as classroom practice is concerned, the differences between history and philosophy we identified in this study shows the importance of addressing discipline-specificity in writing tasks in secondary education. Philosophical thinking and writing are intertwined and parallel processes, which suggests that writing is an appropriate tool for learning in philosophy, and discipline-specific writing instruction might foster writing to learn. However, historical thinking and writing turned out to be separate and partly serial processes. The students' texts often did not include their historical reasoning; the texts merely reflected the outcome of the students' reasoning. Hence, it seems valid to ask whether writing instruction in history is likely to contribute to students achieving our teaching goals. We therefore recommend further exploration of the relationship between what we aim to achieve with discipline-specific writing assignments and what we actually accomplish.