Synthesis writing

Teaching high school students how to read, plan, draft, and revise

van Ockenburg, L.

Publication date
2022

CHAPTER 5

CHOOSING HOW TO PLAN INFORMATIVE SYNTHESIS TEXTS

Effects of Strategy-Based Interventions on Overall Text Quality[*]

Abstract. The current two studies examined the effects of an instructional unit on synthesis writing for ninth-grade students on holistic text quality. Because students' writing routines have been shown to affect the effectiveness of writing instruction, we designed a unit that aimed to be equally effective for all writers, regardless of their routines, titled Learning to Compose Synthesis Texts, or LCST for short. On two occasions in the unit, we provided students with options to choose between a Preplanning or Drafting strategy. They could create an individual learning path within the unit. We tested the effects of LCST in two consecutive studies with independent samples, using a quasi-experimental pretest-posttest design with switching panels. LCST 1 included three teachers and 152 ninth-grade students from five classes in one school. We found a significant effect of the unit on text quality in both panels. However, analyses showed that this effect was moderated by writing routine. To generalize the effect across writing routines, we provided students in LCST 2 with an extra lesson on metacognitive knowledge about (synthesis) writing processes, enabling them to make better-informed strategy choices. LCST 2 included six teachers and 233 students from 10 classes in three schools, all different from LCST 1. Analyses again showed a significant effect of the experimental condition in Panel 1, which was replicated in Panel 2. Moreover, the effect was no longer moderated by writing routine. This indicates that the redesigned intervention was equally effective for students with different routines.
Keywords: Writing Education, Intervention Study, Synthesis Texts, Writing routines, Secondary education

1 INTRODUCTION

The attainment goals for secondary school students in The Netherlands who are preparing for higher education, state that students should be able to gather and process source information to present relevant information in a source-based texts (2009, Ministerie van Onderwijs, Cultuur en Wetenschap [Ministry of Education, Culture and Science], p. 15). This goal should ensure a good preparation for the transition to higher education, since higher education institutes usually require students to write source-based essays or research reports (Meestringa, 2011; Meestringa & Ravesloot, 2013; Oudakker & Groenendijk, 2020). Still, many first-year students in The Netherlands seem to lack adequate writing skills

[*] *This chapter is a somewhat adapted version of Van Ockenburg, L., Van Weijen, D., & Rijlaarsdam, G. (2021). Choosing how to plan informative synthesis texts: Effects of strategy-based interventions on overall text quality.* Reading and Writing*. https://doi.org/10.1007/s11145-021-10226-6*

(Bonset, 2010; Oudakker & Groenendijk, 2020). A panel of teachers from several Dutch universities identified problems with, among other things, writing coherently, revising, and handling sources (Bonset, 2010). While secondary school students are allowed to pick and mix information from their sources at will, in higher education students must synthesize relevant information in a way that accurately reflects the information provided in research articles (Meestringa, 2011; Oudakker & Groenendijk, 2020). The difference between these two representations of source-based writing seems to be too large for students to deal with independently. In addition, students generally receive little instruction on source-based writing, although they certainly need instruction to perform such a complex task successfully (Leijten et al., 2017). After all, writing skills do not improve just by 'doing it frequently' (Rijlaarsdam et al., 2005). Fortunately, a fair amount of research has already been done on learning to write synthesis texts. An analysis of various effective interventions (see Chapter 2) yielded many learning activities that have proven effective, which could guide the design of a unit aimed at learning the basics of writing informative synthesis texts in mid-secondary education. Teaching students to compose synthesis texts in secondary school could better prepare them for understanding writing tasks in higher education.

1.1 *Learning to Compose Synthesis Texts*

Synthesis texts present information from sources, which can complement or contradict each other. The aim of synthesizing is to compose a representative and well-integrated reflection of the source information, presented either as a stand-alone text or embedded in a larger expository or argumentative text. Synthesizing entails recursive interaction between reading and writing activities: exploratory source reading, understanding the sources, selecting information sensibly, arranging the information logically, and integrating the content from different sources while formulating a new text (Klein & Boscolo, 2016). Recursivity, i.e., continuously switching between the reader and writer role, is necessary to successfully perform the synthesis task (Martínez et al., 2015).

Most research reported on composing synthesis texts is situated in higher education (Boscolo et al., 2007; Daher & Kiewra, 2016; Darowski et al., 2016, 2020; Gil et al., 2008; Lundstrom et al., 2015; Luo & Kiewra, 2019; Zhang, 2013). Yet it is in secondary education that students reach the sensitive age for learning how to perform this task. Spivey & King (1989) discovered that learning gains were most evident between Grade 8 and Grade 10, roughly at the age at which Brown and Day (as cited in Spivey & King, 1989, p. 20) found special gains in metacognitive development, such as an increased sensitivity to textual importance, use of text

organization and ability to supply connective links. Consequently, it seems feasible to introduce synthesis writing early in secondary education.

For students from Grades 6 to undergraduate level, several effective interventions have already been developed and tested (Barzilai & Ka'adan, 2017; Britt & Aglinskas, 2002; Kirkpatrick & Klein, 2009; Martinez et al., 2015; Zhang, 2013). Often these interventions focus on one particular subprocess of the synthesis task, for instance integration (e.g., Barzilai & Ka'adan, 2017), outlining (e.g., Kirkpatrick & Klein, 2009), or sourcing (e.g., Britt & Aglinskas, 2002). However, a systematic review of intervention studies (see Chapter 2) showed that effective instruction should encompass all three key processes of synthesizing: selecting, organizing, and connecting source information (Spivey & King, 1989). Moreover, effective synthesizing instruction combines cognitive strategies with metacognitive knowledge about when to engage in which cognitive process and why (Segev-Miller, 2004, p. 8). In conclusion, an effective intervention should thus include metacognitive knowledge about the synthesis task and learning activities with cognitive strategies for all subprocesses.

1.2 *Writing Routines*

In addition to the cognitive load that recursive synthesizing processes already cause, multiple other subprocesses play a role during writing, including the most basic components: planning the text, converting the plan into written text, and revising it (Hayes, 2012). All these processes are cognitively demanding and together they may lead to working memory overload (Kellogg, 1996). Therefore, many writers develop a routine for decomposing the complex writing process into "steps" that they execute in a certain order. Such a writing routine reduces both task complexity and the risk of working memory overload (Kieft & Rijlaarsdam, 2005).

Students' writing routines may be understood as being guided by two dimensions of idea generation processes: preplanning and post-revising (Galbraith, 1992; Kieft & Rijlaarsdam, 2005). Preplanning concerns the extent to which students engage in outlining the content structure of their texts before starting to generate text, while post-revising (in this paper referred to as *Drafting*) concerns the extent to which students are inclined to engage in rethinking and changing their texts during and after writing a first draft.

Based on earlier research, we assume that students have mixed routines. They each Preplan and Draft to some extent. Kieft et al. (2007) reported a correlation between the two dimensions for 10th-grade (r = .38, p < .001) and Van Ockenburg et al. (2018) for ninth-grade students (r = .24; p < .001). These correlations

indicate that Preplanning and Drafting are two related, but distinct dimensions of one's writing routine.

1.3 *Interaction Between Writing Routines and Planning Strategies*

Instruction aimed at limiting cognitive overload during writing often focuses on offering content preplanning strategies. Despite the widely proven effectiveness of preplanning strategies such as outlining and graphic organizers (e.g., Limpo & Alvez, 2018; Kellogg, 1987, 1988, 2001, 2008), not all preplanning strategies are equally effective for all types of writers (Rijlaarsdam et al., 2012). Some writers may benefit more from planning-by-drafting/post-writing revising strategies (Galbraith & Torrance, 2004), in which writers are enabled to immediately start writing to boost the idea generation process and then facilitated in revising this rough, first draft.

Regarding this complex relationship between planning strategies and students' writing routines, there are currently three perspectives in the literature. First, Torrance and Galbraith (2006) argue that offering students strategies that suit their preferred writing routine might help reduce their cognitive load while writing. However, from a second perspective it could also be argued that relations between writing routines and instruction may be more complex. Kieft et al. (2007) found that students who strongly adhered to a Preplanning or a Drafting oriented routine, both benefitted more from writing instruction based on a Preplanning strategy than a Drafting strategy. Students who did not adhere to a specific strategy, on the other hand, benefitted more from writing instruction based on a Drafting strategy. This is partly in line with research by Baaijen et al. (2014) who investigated the effects of another student characteristic on writing instruction: transactional beliefs. These beliefs relate to the degree to which students are convinced that they can only generate ideas and understand a topic by writing (White & Bruning, 2005, p. 168). Students with high transactional beliefs are likely to thrive by using a Drafting routine, in which thinking interacts with text production. Baaijen et al.'s (2014) study showed that instruction based on outlining strategies had a positive effect for low transactional writers but no effect for high transactional writers.

The Strategic Content Learning (SCL) approach (Butler, 1998) represents a third perspective, which encourages students to generate personalized strategies instead of teaching them specific, predefined ones. The SCL-model aims to help students acquire metacognitive knowledge about tasks, strategies, and themselves as learners and about how these three factors interact to affect the course and outcome of their cognitive activities. Robledo-Ramon (2016) applied

the SCL-approach in an effective lesson series on synthesis writing for undergraduates, in which students developed personal synthesizing strategies.
The outcomes of the described studies do not provide unambiguous answers to the question which strategy is most effective and how that effectiveness is related to writing routines. Hence, it remains unclear what the best instructional approach is for students with and without a strong adherence to a particular routine.

1.4 *Current Studies*

We report here on two studies into the unit Learning to Compose Synthesis Texts (LCST). LCST 1, an efficacy study, was carried out in a controlled, single school setting (N = 152), to determine whether the learning unit we designed was able to produce the desired outcome (O'Donnell, 2008, p. 41). We reported a full description of the design principles, an evaluation of its efficacy, preliminary analyses of its main effects and suggestions for a redesign in Chapter 3. Here, we report on its effectiveness. Subsequently, LCST 2 was set up as an effectiveness study, in which a redesigned version of the unit was implemented on a larger scale (N = 233) in multiple schools.

In previous (synthesis) writing intervention studies, the learning activities offered were similar for all participants. Even in studies investigating the influence of students' writing routines on the effectiveness of instruction, groups of students were all offered the same strategies (Baaijen et al., 2014; Baaijen & Galbraith, 2018; Kieft et al. 2007). To our knowledge, no study has provided participants with the opportunity to choose between different planning strategies (organizing Post-it's or Direct Drafting) several times during an intervention and, consequently, create a personal learning path. That is why our research questions, for both studies, were:

1. To what extent does a synthesis writing unit influence the quality of the texts produced?
2. Does the effect depend on students' writing routines?
3. Do strategy choices that students make in the experimental condition relate to their writing routine, affect text quality, and reflect their knowledge of the writing process?

Regarding RQ 1, we hypothesized that students who participated in the unit would write synthesis texts of higher quality compared to the control group, given that previous strategy-based intervention studies have shown

improvements in students' writing (e.g., Barzilai & Ka'adan, 2017; Britt & Aglinskas, 2002; Kirkpatrick & Klein, 2009; Martinez et al., 2015; Zhang, 2013).

Regarding RQ 2, we expected that writing routines would moderate the effects of writing instruction (Baaijen et al., 2014; Kieft et al. 2007). Such a moderation effect is undesirable, because our goal was that instruction is equally effective for all students, regardless of their writing routine. Therefore, we let students choose between two planning strategies: Preplanning and Drafting. They had to choose between these two strategies on two occasions and thereby created their personal learning path. As a result, three different learning paths could occur within the same unit: a) all-Preplanning, b) all-Drafting, and c) Switching between the two options. We expected that offering choices might facilitate the generation of metacognitive task knowledge, because choosing requires goal setting, which is a condition for monitoring and evaluating. Based on a study by Robledo-Ramon (2016) that focused on students gaining metacognitive task knowledge, we expected that our approach would promote equal effectiveness for students with different writing routines.

The outcomes for RQ 3 should provide more insight into these considerations. We expected that the opportunity to make and explain strategy choices would increase students' insight in the writing process, and thus enable them to choose an optimal individual learning path, regardless of their initial writing routine. The experimental condition is considered to be effective when differences in learning paths do not affect the outcomes of the learning unit.

2 METHOD

We carried out two independent studies to test the efficacy and effectiveness of an instructional design in terms of the quality of students' synthesis texts. In both studies we implemented a quasi-experimental design with switching panels and three measurement occasions as shown in Table 5.1 (Shadish et al., 2002, p. 146-147). We assigned intact classes randomly to either the EC or CE group. Group EC participated in Panel 1 as experimental group (E), group CE as control group (C). For this group M1 was the pretest, M2 was the posttest, and M3 was the retention test. After Measurement occasion M2, the groups switched conditions: Group EC became the control condition, group CE received the intervention. For this group M1 was the pretest, M2 was the control test, and M3 was the posttest. An advantage of this design is that it provides the opportunity to replicate possible intervention effects (Panel 1 vs Panel 2). Other advantages are that the design is ethically justified (all participants receive the intervention), and valid: switching panels controls for teacher effects since all teachers are involved in

both conditions. The same applies to differences between participants in both conditions. If the effects of the intervention would have been affected by hidden differences in group composition, then replication would not be possible and a difference between conditions would be found at M3.

Table 5.1 Quasi-experimental design with switching panels (identical for LCST 1 and 2)

	M1	Panel 1 Condition	M2	Panel 2 Condition	M3
Group EC	O	E	O	C	O
Group CE	O	C	O	E	O

Note. M = Measurement occasion (1, 2, 3);
C = Control (Regular Language Curriculum without writing instruction);
E = Experimental instruction; O = Observation

2.1 *Participants*

2.1.1 LCST 1

Participants were 152 ninth-grade pre-university students (*M* age = 14.0, *SD* = 0.50; 61.8% female) from five classes, all from the same school in a southern region of The Netherlands. Class sizes varied from 29 to 32 students. Intact classes were assigned to the EC or CE group. These groups did not differ in terms of gender (63% female in group EC, 64% in group CE; $\chi^2(2) = 0.10$, $p = .953$) or age (*m* = 14.00 years, *sd* = 0.48 in group EC; *m* = 14.01 years, *sd* = 0.52 in group CE; $t(150) = -0.13$, $p = .988$). Most of the participants (93%) were native speakers of Dutch (L1). The others came from families in which parents did not speak Dutch as their first language. A speech or language disorder was indicated by 2.7% of participants. The parents of the students received an informed consent e-mail and could object to their child's participation, which one parent did.

2.1.2 LCST 2

Participants were 233 ninth-grade pre-university students (*M* age = 13.93, *SD* = 0.49; 55% female) from ten classes from three secondary schools, one located in an eastern region of The Netherlands, one in a southern, and one in a central region. Class sizes varied from 17 to 31 students. Schools were assigned to the EC or the CE group. These groups did not differ in terms of gender (60% female

in group EC, 50% in group CE; $\chi^2(2) = 2.558$, $p = .27$) or age (age $m = 13.93$, $sd = 0.45$ in group EC; $m = 13.92$, $sd = 0.54$ in group CE; $t(191) = 0.12$, $p = .908$). Most participants (92%) were L1 Dutch speakers. A speech or language disorder was indicated by 3.8% of participants. The parents of the students received an informed consent e-mail and could object to their child's participation. Although some parents requested further information, no one objected to their child's participation.

2.2 *Materials*

2.2.1 Instructional Designs

To respond to the differences in writing routines and to support all students equally in learning to write informative synthesis texts, we designed a learning unit guided by three main principles (for more information, see Chapter 3):

Principle 1. Focus on cognitive strategies. Through a systematic review (see Chapter 2), we found that most effective synthesis writing interventions covered all three synthesizing processes: *selecting*, *organizing,* and *connecting* (Spivey & King, 1989). That is why we chose to offer learning activities for practicing each of these three processes.

Principle 2. Focus on Writing Routines. Earlier research (Baaijen et al., 2014; Kieft et al., 2007; Torrance & Galbraith, 2006) seems to indicate that writing routines interact with the effectiveness of planning strategies. That is why we pay attention to raising awareness of one's own writing process in the design.

Principle 3. Activate Metacognition by Choice. In a study by Robledo-Ramon (2016) the SCL-approach in which students developed metacognitive task knowledge proved to be very effective. We aimed to activate metacognition by encouraging students to choose between strategies. We assumed that when students were required to choose, and thus set goals for their writing task, their metacognition would be activated. This required designing flexible lessons, so students would be able to experiment with different strategies related to Preplanning and Drafting.

Table 5.2 shows an overview of the operationalization of the design principles in LCST 1.

Table 5.2 General lesson plan of the unit (LCST 1)

Phase	Specification
0 (Whole pre-task experience)	
i	Observing a 3-minute animated video providing information about integrating source information in synthesis texts.
ii	Performing a synthesis pre-task to experience the synthesizing process as a whole.
1 (*pre*-flection)	
i	Ranking four model texts (on the same topic of the pre-task) according to quality, to build task representation for synthesis texts, by noticing differences in text quality.
ii	Discussing text quality of these four model texts to refine and elaborate task representation for text quality.
2-4 (Acquisition of subskills; similarly structured lessons)	
i	Observing two strong peer models who perform a synthesis subtask by employing different strategies, while noticing similarities and differences.
ii	Discussing the modelled strategies to determine which strategy one prefers and why. In lesson 3, noting and motivating choice between the modelled planning strategies.
iii	Applying chosen strategy to a subtask to experience one's preferred strategy.
iv	Discussing written product(s) to evaluate the result of applying the chosen strategy and estimate whether the result was satisfactory.
5-6 (Integration: practice)	
5-i	Considering whether to keep or change the planning strategy chosen in lesson 3 and noting and motivating new choice in workbook.
5-ii	Performing a new synthesis task to experience the use of the chosen strategies in a whole-task assignment.
6-i	Considering the quality of each student's own text, using a text scale to determine to what extent the chosen strategies helped to meet the quality criteria agreed on.
6-ii	Giving peers feedback, based on previously discussed text quality criteria (in lesson 1) to experience the effect of one's text on a reader and determine to what extent the chosen strategy was adequate

Note. Adapted from Chapter 3, Table 3.1

This design comprised seven 50-minute lessons. The first two lessons aimed to create a task representation (Table 5.2, Lesson 0-1). Three lessons followed (Table 5.2, Lesson 2-4), each instructing one of the three basic synthesizing processes, respectively *selecting*, *organizing* and *connecting*, adding up to a final text version in Lesson 4. The last two lessons of the unit (Table 5.2, Lesson 5-6) provided students with the opportunity to practice the recently acquired strategies in a new, whole-task synthesis assignment.

Lesson 0 (pretest) provided students with a whole-task experience: writing a synthesis text. Because students had no prior experience with synthesis writing, they first watched a three-minute animated video, which provided them with a task definition of what synthesis writing entails. However, the video did not contain any instructions on how students could approach the synthesis writing task (see Chapter 4). Subsequently, Lesson 1 offered the opportunity for *pre*-flection: creating a shared definition of a well-written synthesis text before engaging in instructional lessons. Then, three instructional lessons (2-4) followed, all offering the same sequence of learning activities, activated by learning through observation. In Lesson 2, students compared-and-contrasted two peers, both relatively proficient in writing but preferring different writing routines. A film clip showed how they studied the assignment, read sources, and *selected information*. In Lesson 3, the same models *clustered and organized* source information. One model used a Preplanning strategy: she wrote the most important information from each source in key words on Post-it notes and then clustered and arranged the notes. Her approach was based on the Color-Coding Method (Darowski et al., 2016, 2020; Lundstrom et al., 2015). The other model used a Drafting strategy: she immediately started writing and linking information from different sources. In Lesson 4, the models *re-organized and connected* the source information and wrote the final draft of their synthesis text. In all instructional lessons (2-4) the students chose between different cognitive strategies, but they only made their choice explicit in Lessons 3 and 5, by noting and substantiating it in their workbooks. In Lesson 3, they could choose between two modelled planning strategies, Preplanning or Drafting, after having observed the two models. In Lesson 5, students again chose explicitly which planning strategy they wanted to apply, after having started a new writing task. They could either stick with their first choice, or switch strategies.

In LCST 2, the experimental instructional design was largely consistent with LCST 1, except for one extra lesson in the task representation phase, between the first and the second lesson of LCST 1. This additional lesson aimed to provide students with metacognitive knowledge about (synthesis) writing processes and writing routines. They received their personal outcome from a questionnaire

about writing approaches that they had completed at the beginning of the first lesson and related their scores to the average scores of students from a national baseline study on synthesis texts (Vandermeulen et al., 2020a). Next, the students watched an animation video showing different effective writing configurations for synthesis tasks, based on research by Vandermeulen (2020b, pp. 95-134). The students compared and contrasted their approach with the effective configurations shown and then made a writing plan: which aspects of their approach did they want to change and how? And which aspects were they satisfied with?

2.2.2 Writing Measurement

To ensure the validity of the synthesis tasks for the measurement occasions, we chose to use tasks that were developed and tested as part of a national assessment study (Vandermeulen et al., 2020a). In the original tasks the number of source texts varied from three to five, but we limited the number of source texts to three due to the lack of experience our target group had with synthesizing. The mean number of words per source text was 188.9 (*SD* = 55) per topic (e.g., the human-wildlife conflict in Africa). Content-wise, the sources complemented each other. Students had to write informative synthesis texts of approximately 200 words (min. 180, max. 220).

For LCST 2, we decided to replace one measurement task (on the topic of self-driving cars), prompted by comments from several raters in LCST 1 who indicated that they had noticed that the sources for that task were more difficult to integrate than those in the other tasks. We administered task M3 from LCST 1 (food additives) at M2 and chose another task from Vandermeulen et al.'s (2020a) study to administer at M3 (pay gap). By doing so, we think we solved the problem raters in LCST 1 noticed: for LCST 2 we did not receive comments from raters about differences between the assignments.

2.2.3 Writing Style Inventory

The online Writing Style Inventory (WSI) was used to collect data about the students' writing routines. Initial versions of the WSI were designed by Kieft et al. (2007) to measure the extent to which writers tend to invest time and effort in the preplanning and/or post-draft revision phase in argumentative writing tasks. The WSI has previously been used in various contexts (e.g., Arias-Gundín & Fidalgo, 2017; De Smedt et al., 2016), and was validated by Arias-Gundín et al. (2021). The current version was adapted and validated in the context of synthesis tasks (Vandermeulen, 2020). The last version starts with a case description which

asks students to imagine that they are writing a synthesis text, after which they are told that the researchers would like to know how students think they would actually do this, not how they think they should.

2.3 *Procedures*

2.3.1 Implementation

LCST 1 took place in five classes of school A (see Table 5.3). During the first implementation, the teacher-researcher, an experienced teacher in the school subject Dutch language and literature, taught the learning unit to her own two classes (Panel 1 (EC), n = 60) for two consecutive weeks, as part of the regular curriculum, completing three 50-minute lessons per week. The remaining three classes formed Panel 2 and served as control group during this period. They continued with their regular curriculum consisting of reading skills instruction, spelling, grammar, and literature, except for any writing instruction. During the second iteration the conditions were reversed, and the teacher-researcher's two colleagues taught the unit in three classes (Panel 2 (CE), n = 92) while Panel 1's classes served as the control group.

In the week prior to the start of the first iteration, all students took part in the pretest (M1). The students watched a brief instructional video and received a synthesis assignment and three sources on paper in randomized order. They wrote their synthesis text on the computer and completed a digital questionnaire in which they answered questions about their background and writing routines. The pretest (M1) took place during a scheduled extra-curricular lesson hour, under the supervision of the teacher-researcher and three research assistants. The other tests were administered between iterations 1 and 2 (M2) and, finally, in the week following the completion of the lessons of the second iteration (M3). These tests took place during regular lessons under the supervision of their own teacher.

LCST 2 was carried out in 10 classes; all three schools that registered after an open call on social media participated (see Table 5.3). During the first panel, three teachers from schools B and C taught the unit as part of the regular curriculum to five classes (Panel 1 (EC), n = 121) in two to four weeks. Per school the number of available teaching hours and the duration of the teaching hours differed, which resulted in some variation in implementation. The classes from school D served as the control group during this period and continued with their regular curriculum, except for any writing instruction. During the second iteration the conditions reversed, and the two teachers at school D taught the unit in five

classes (Panel 2 (CE), *n* = 129) while the first iteration's experimental classes served as the control group.

We assessed text quality for all LCST 1 participants. For practical reasons (time and budget constraints) we assessed text quality for a sub-sample (*n* = 162) of the total group of LCST 2 participants (*N* = 233). First, we selected the students who had completed all lessons and measurements and then we randomly selected participants from this subsample while ensuring that all participating classes were represented (see Table 5.3).

Table 5.3 Participants. Total and selection for assessing Text Quality

Study	School	Location in the Netherlands	Teachers	Classes	Students	Selected for TQ
1	A	South-Central	3	5	152	152
2	B	Central	2	4	104	71
	C	South	1	1	17	14
	D	East	2	5	112	77

2.3.2 Writing Routines

Table 5.4 Student variables included in these studies as measured by WSI

Study	N	Scale	N Items*	α	*M*	*SD*	Highest loading items
1	126	Pre-planning	6	.68	2.34	.67	a) I always make notes before I start writing.
2	193		7	.76	2.65	.69	b) If I write a text, I spend a lot of time thinking about my approach.
1	126	Drafting	6	.71	2.77	.67	a) If I read and revise my text, its structure can still change considerably.
2	193		7	.74	2.88	.60	b) When I have finished writing my text, I must read it very carefully to delete superfluous information.

Note. *Both scales consist of 7 items in total. In LCST 1 we had to delete 1 item to ensure acceptable reliability.

We used the Writing Style Inventory (WSI) to obtain students' scores on two subscales: Preplanning and Drafting. The higher the average score on the subscales Preplanning and Drafting, the more intensively the students tend to plan before they start writing or revise after writing a first draft. The average scores (Table 5.4) indicated that in general students did not expect to plan their text comprehensively prior to writing or revise it extensively afterwards. Both scales correlated in LCST 1 (r = .36, p < .001), but not in LCST 2 (r = .08, p = .25).

2.3.3 Text Quality

To assess text quality (TQ), we provided each rater with a written instruction that included: (1) students' assignments and source texts, (2) a benchmark scale, showing sample texts of increasing quality, including explanations of their characteristics, and (3) four assessment criteria with an extended explanation: (a) representation of source content, (b) integration of source information, (c) structure, and (d) style/language. We asked raters to first focus on each of these four aspects separately, by providing a score on a scale from 1 to 5 for each aspect, and then provide a holistic score on a scale from 1 to 100. We assumed that assessing the criteria could support the raters to form a nuanced judgment about the holistic quality of the text. Indeed, the holistic score seems to represent all four aspects since the average correlation between aspect scores and the holistic score was .79 (LCST 1) and .82 (LCST 2). Therefore, we decided to use the holistic text quality score as the single dependent variable that represents four aspects of text quality of informative synthesis texts.

In LCST 1 21 raters were involved, in LCST 2 there were 25 raters. Raters were teachers, preservice teacher students, former teachers and writing researchers, who volunteered after a call through various social media channels. They received a reward in the form of a gift voucher. Each text was rated independently by three raters. We created overlapping teams of three raters, with each rater evaluating about 50 texts. Jury reliability was calculated using Van den Bergh and Eiting's (1989) method and the correlation coefficient was ρ = .71 for LCST 1 and ρ = .70 for LCST 2.

2.3.4 Choices and Motives

We analysed students' workbooks for their strategy choices and whether they showed insight into their own writing process. Students could choose twice between a Preplanning and a Drafting strategy for organizing source information (see Table 5.2, Lessons 3 and 5). On both occasions, they noted their chosen

strategy in their workbooks and explained why they chose it. We included their choices in the analyses as well as the motives for their choice on the first occasion.

2.4 *Treatment Fidelity*

To monitor the implementation and determine the learning unit's efficacy (LCST 1) and effectiveness (LCST 2) (O'Donnell, 2008), we used two data sources: time-on-task observations (LCST 1), and teacher logs (LCST 1 and 2).

In LCST 1, the teacher-researcher, who designed the learning unit and materials, taught her regular classes in the first iteration. She used her experiences to create a comprehensive teacher manual, including background information, detailed lesson plans, and answer keys to all activities. Before the start of the second iteration, the teacher-researcher familiarized the other teachers with the theoretical-empirical background of the unit, its design and the use of student and teacher materials. Prior to this training session, these teachers had studied the materials to prepare specific questions.

Pre-instructed research assistants carried out time-on-task classroom observations to obtain an indication of the proportion of realized learning time, that is the amount of available learning time that students actually devoted to the tasks they were required to perform (e.g., Karweit, 1984). They observed several experimental lessons each: Lesson 1 in three classes; Lesson 2 in three classes; Lesson 4 in four classes; Lesson 5 in two classes; and Lesson 6 in one class. Overall, students spent on average 78% of the observed time on the task they were supposed to work on. 19,8% of the time they were "off task", while 2,2% was coded as "unclear".

After each lesson, teachers completed a log, indicating the extent to which they had completed key lesson activities (fully, partially, or not). The response rate was high (98%). The lessons were taught as intended: 92.2% of the key lesson activities were fully completed, 7.8% partially.

Prior to the start of LCST 2, the teacher-researcher discussed guidelines for the implementation of the learning unit and the teacher's manual with each participating teacher on site. They filled in a logbook after each lesson. The response rate was 93.5%. Lessons were taught as intended: 91.4% of the key lesson activities were fully completed, 8.6% partially. Teachers reported lack of time as the main reason when failing to complete lesson activities.

2.5 *Data Analysis*

To answer Research Questions 1 and 2, we analysed the data in two steps; first we established the main effects of the intervention on holistic text quality (RQ 1), then we explored the extent to which these effects were moderated by the writing routine dimensions Preplanning and Drafting (RQ 2). Furthermore, we analysed students' strategy choices and their motives substantiating these choices to answer Research Question 3.

2.5.1 RQ 1: Effects of the Learning Unit

As data were nested in individuals (M1-2-3), and individuals were nested in intact classes, we used mixed models in SPSS to analyse the data. We compared the fit of four nested models with a likelihood ratio test. We started with a model without any explanatory variables, consisting of the intercept and student as a source of variance (Model 0), then added the effect of measurement occasion (Model 1), condition (Model 2) and the interaction between measurement occasion and condition (Model 3), which would point to differences in development over time due to condition. We ran these analyses separately for both panels.
Regarding effect sizes, there is much discussion about the formulas to calculate R-square from multilevel analyses. To provide a proxy indication of the explained variance, we correlated the predicted values, based on the models we report in the paper, with the observed scores. Squared correlations indicate the variance explained by the fixed parameter, i.e., the learning unit.

2.5.2 RQ 2: Moderator Effects - Writing Routines

To explore moderator effects, we expanded the analyses with three additional models. In these models, the centred Preplanning and Drafting scores were added as continuous variables. In Model 4, we added a moderator variable to test whether this variable affected the dependent variable. Then we tested whether the variable affected the outcomes differently on the three measurement occasions (Model 5), on both conditions (Model 6), or the interaction between condition and measurement occasion (Model 7). If Model 7 would best fit the data, this would indicate that writing routine moderated the effect of the experimental learning condition.

2.5.3 RQ 3: Strategy Choices and Motives

In two lessons, students were explicitly invited to choose between two strategies: Preplanning or Drafting. They were free to switch on the second occasion. We analysed students' choices for one of the two strategies on both occasions, to what extent they switched strategies, how their choices related to their WSI-scores, and how their choices affected text quality. These questions could only be answered if three types of data were all available for a participant: TQ scores, choices and motivation.

In LCST 1 we analysed how choice affected TQ for all participants of whom we had acquired choice data (n = 132) during the first implementation and coded the motive for their choice from just a small sample (48 students from 2 classes from different teachers and panels) to explore the data. In LCST 2 we assessed TQ for a sample sufficient conform a power analysis (*n* = 132, total set: *N*= 233). Seventy two of these 132 participants provided us with choice data and motives.

Two independent coders coded the motives with which students substantiated their choices and determined whether these motives showed insight into the writing process. Motives showed insight when information was provided about (a) the students' regular writing process, and/or (b) their process in the upcoming writing task. Table 5.5 shows examples of students' motives and how they were coded.

Table 5.5 Examples of students' responses and their codes

	Choice	Examples of students' responses	RP	UT
1	Preplanning	I would plan first because otherwise I would probably lose track.	-	+
2	Preplanning	I always like to list everything that belongs together first, so that writing is easier afterwards and I have an outline.	+	-
3	Drafting	Because I often don't make extensive writing plans and I am going to revise my first version very well.	+	+
4	Drafting	It takes less time / I like it better.	-	-

Note. RP = Refers to Regular Process; UT = refers to Upcoming Task

Based on a protocol with elaborations of the codes for *a* and *b*, and their application to various sample responses, the two coders first analysed ten motives step-by-step. After coding each sample independently, codes were compared and discussed. If needed, the protocol was supplemented. After coding ten

motives this way, the coders independently coded the motives in random order. Consistency between coders was substantial for *a* (*k* = .615) and *b* (*k* = .744).

3 RESULTS

3.1 *RQ 1: Effects of the Learning Unit*

Although the tasks in all three measurement occasions were similar, they were not equal. This implies that score differences between measurement occasions are difficult to interpret as an improvement or a decline. To avoid possible task effects, we only determined whether scores between conditions on a specific measurement occasion differed significantly. If this was the case at M2, and the intervention group scored higher than the control group, that indicated the effectiveness of the intervention in the first iteration. Subsequently, if, after a significant difference at M2, this effect was absent at M3, that would indicate the effectiveness of the intervention in the second iteration, because both groups, Panel 1 (EC) as well as Panel 2 (CE), scored similarly again.

For LCST 1, model 3 was significant at $p < .001$, indicating an interaction between measurement occasion and condition. At M1, no difference between the two conditions was observed (*B* = 0.73, *SE* = 2.23, *B* being smaller than 1.96*SE), but at M2 the experimental group significantly outperformed the control group (*B* = -9.91, *SE* = 2.23, *B* being larger than 1.96*SE). At M3, after switching conditions, no differences between the two groups were observed. Table 5.6 presents the means and standard errors for holistic text quality scores for LCST 1, as estimated under model 3. The estimated effect size under model 3 is R^2 = .321, *ES* = 10.3%. However, because we found that the writing routines Preplanning and Drafting moderated holistic text quality, we present further outcomes in the elaboration of RQ 2.

Table 5.6 LCST 1. Mean Holistic Text Quality (and standard error), as estimated under model 3

Panel	M1 *M* (*SE*)	M2 *M* (*SE*)	M3 *M* (*SE*)
1 (EC)	57.87 (1.32)	68.88 (1.34)	62.49 (1.33)
2 (CE)	58.78 (1.08)	60.22 (1.07)	63.51 (1.07)

Note. Panel 1 = First Experimental then Control Condition; Panel 2 = First Control then Experimental Condition

For LCST 2, model 3 was significant at p = .001, which indicated an interaction between Measurement Occasion and Condition. At M1, no significant difference between the two conditions was observed (B = -4.75, SE = 2.51). At M2, the difference between the experimental and the control condition was significant (B = -9.91, SE =2.51). At M3, no differences between the two groups were observed. Table 5.7 presents the means and standard errors for holistic text quality scores for LCST 2, as estimated under model 3. The estimated effect size under model 3 is R^2 = .467, ES = 20.8%.

Table 5.7 LCST 2. Mean Holistic Text Quality (and standard error), as estimated under model 3

Panel	M1 *M* (*SE*)	M2 *M* (*SE*)	M3 *M* (*SE*)
1 (EC)	55.63 (1.51)	63.66 (1.51)	66.41 (1.51)
2 (CE)	51.02 (1.48)	53.89 (1.48)	66.55 (1.48)

Note. Panel 1 = First Experimental then Control Condition;
Panel 2 = First Control then Experimental Condition

Model comparisons for LCST 1 and 2, for the dependent variable holistic text quality, with parameter estimates for model fit are presented in Appendix I.

3.2 *RQ 2: Moderator Effects - Writing Routines*

Model comparisons for the moderation of the dependent variable holistic text quality by the Preplanning and Drafting variables, indicated that in LCST 1 the interaction between condition and measurement occasion was moderated by Preplanning (Model 7: p = .015). The interaction-effect with measurement occasion and condition was located at M2 (B = 6.62, SE = 3.14) in the CE-group: participants who scored relatively high on Preplanning, scored higher than expected at M2. Furthermore, there was a general contribution of Preplanning to text quality (B = 5.04, SE = 1.93). We observed no interactions for the Drafting variable but found a general effect independent of measurement occasion or condition (Model 4: marginally significant at p = .05, B = 1.91, SE = 0.96), which indicates there was a general contribution of Drafting to text quality, but smaller than the general Preplanning effect.

Figure 5.1 LCST 1. Effect of Preplanning on Holistic Text Quality

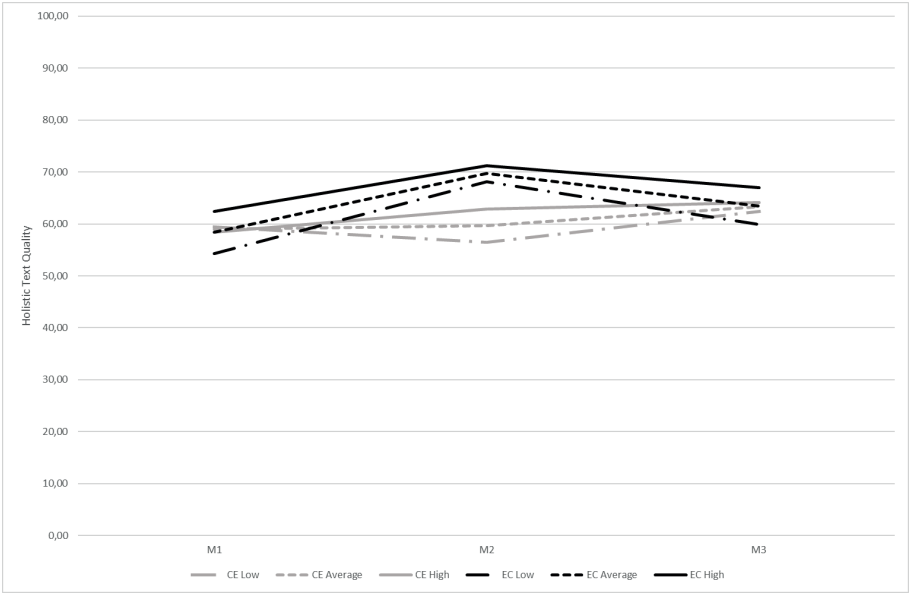
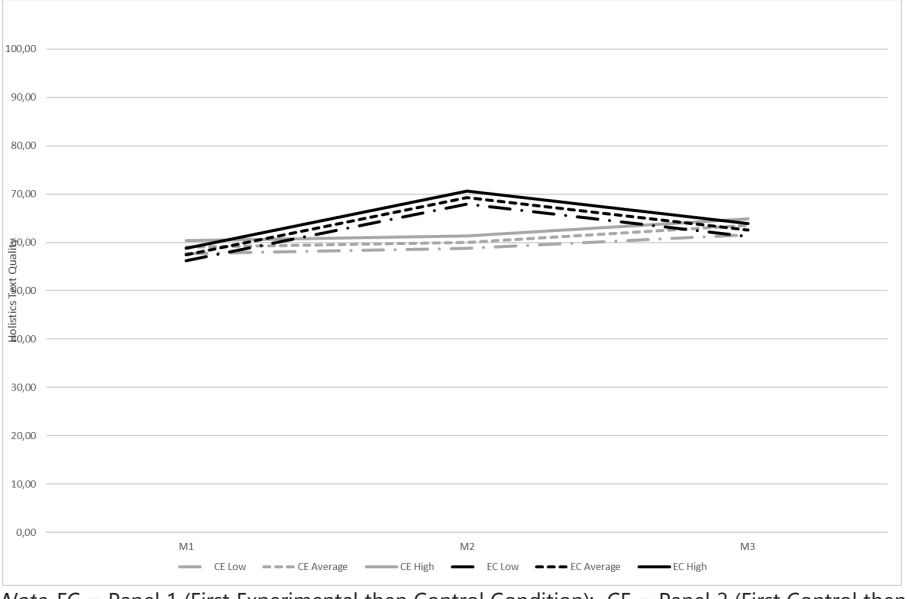


Figure 5.2 LCST 1. Effect of Drafting on Holistic Text Quality



Note. EC = Panel 1 (First Experimental then Control Condition); CE = Panel 2 (First Control then Experimental Condition); Average = Mean level of Preplanning (Fig 5.1) or Drafting (Fig 5.2; High = average + 1 sd; Low = average - 1 sd

Figures 5.1 and 5.2 respectively visualize the complex interaction in a switching replication design for Preplanning and Drafting in LCST 1. The estimated effect sizes for Preplanning under model 7 are R^2 = .321, *ES* = 15.5%, and for Drafting under model 4 R^2 = .339, *ES* = 11.5%. In LCST 2, holistic text quality and interaction between condition and measurement occasion were not moderated by Preplanning or Drafting. Appendix J presents all model comparisons for LCST 1 and 2.

3.3 *RQ 3: Strategy Choices and Motives*

In two lessons, students were explicitly invited to choose between a Preplanning and a Drafting strategy. They could maintain or switch strategies between the first and second occasion. Table 5.8 shows the choice distribution for both studies. The reasonably balanced distribution of the choices between the two strategies shows that both strategies were feasible choices for the students.

Table 5.8 Students' distribution over strategy choices (in percentages) for Studies 1 (n = 132) and 2 (n = 72)

		LCST			
		1		2	
	Lesson	a	b	a	b
Strategy	Preplanning	44	50	50	58
	Drafting	56	50	50	42

Table 5.9 Students' distribution over learning paths (in percentages) for Studies 1 (n = 132) and 2 (n = 72)

		LCST	
		1	2
Learning path	All-Preplanning	38.6	40.3
	All-Drafting	44.7	31g9
	Switching	16.7	27.8

When we examined the effect of strategy choices, we distinguished three learning paths: all-Preplanning (PP-PP), all-Drafting (D-D) and Switching (PP-D/D-PP). Table 5.9 shows that in LCST 2 a larger proportion of students chose to switch

strategies and that in LCST 1 the all-Drafting path was more popular than in LCST 2.

The all-Preplanning path was preferred when students scored relatively high on the Preplanning scale of the WSI (Effect sizes LCST 1: choice 1 .28, choice 2 .43; LCST 2: choice 1 .45, choice 2 .42). The score on the Drafting scale only had a small effect in LCST 1, Lesson 3: students with higher scores on the Drafting scale tended to choose the all-Preplanning path (*ES* = .22).

We examined the effect of learning paths on text quality with writing routines as moderator variables, to test whether the effects depended on writing routines. In LCST 1 an effect was observed for the all-Preplanning path (PP-PP), which led to significantly higher text quality scores than the all-Drafting path (D-D) (*B* = 3.50, *SE* =1.46, *p* = .018). A general effect of Preplanning was also observed (*B* = 3.30, *SE* =1.42, *p* = .022), which means that higher Preplan routine scores were associated with higher TQ scores, regardless of condition, measurement occasion and learning path. However, the effect of the Drafting routine score (*B* = 4.56, *SE* =1.46, *p* = .002) only applies to the students who followed an all-Drafting path. When students switched between strategy options, the effect of the Drafting routine score was significantly negative compared to that of the all-Drafting path (D-D) (*B* = - 5.96, *SE* =2.51, *p* = .019). The effect was marginally statistically negative for the all-Preplanning path (*B* = - 4.16, *SE* =2.17, *p* = .058). In LCST 2 we found no effects of learning path or writing routine.

We did not find differences in students' motives substantiating their strategy choices between Studies 1 and 2 in terms of references to the regular writing process ($\chi^2(2) = .77$; $p = .67$). However, there was a difference in references to the upcoming task ($\chi^2(2) = 6.69$, $p = .03$). In LCST 2, both coders agreed in 36.5% of the cases that the motives referred explicitly to the upcoming task, while in LCST 1 this was lower (21%).

4 DISCUSSION

The current study was set up to examine the efficacy of a synthesis writing intervention (LCST 1) and the effectiveness of the redesign on a larger scale (LCST 2). In both cases we investigated whether students' participation in the unit influenced the quality of the texts they produced (RQ1). Furthermore, we also investigated whether students' adherence to writing routines moderated the unit's effects (RQ2) and to what extent strategy choices students made in the experimental condition were related to their writing routine, affected text quality, and reflected their knowledge of the writing process (RQ3).

4.1 *RQ 1: Effects of the Learning Unit*

Regarding the efficacy and effectiveness of the unit, we found that both Studies 1 and 2 showed the expected effects to a large extent. As a result of the switching replication research design, we observed an effect of the learning unit four times, at M2 and at M3, both in LCST 1 and 2. Furthermore, the average text quality at M3, the delayed posttest for Panel 1 (EC), suggests that the effect of the experimental condition in the first iteration was maintained in the second iteration, while in both studies no difference between the conditions at M3 was observed. These results further support earlier findings that strategy instruction is an effective approach for improving students' synthesis writing (e.g., Barzilai & Ka'adan, 2017; Britt & Aglinskas, 2002; Kirkpatrick & Klein, 2009; Martinez et al., 2015; Zhang, 2013) and that a systematic analysis of intervention studies (see Chapter 2), provides a good starting point for developing effective synthesis writing instruction.

4.2 *RQ 2: Moderator Effects - Writing Routines*

Because we aimed to design a learning unit that would be equally effective for writers with different writing routines, we investigated whether students' writing routines moderated the unit's effects. Based on the study by Robledo-Ramon (2016) which focused on developing metacognitive task knowledge, we expected that offering strategy choices would contribute to generalizability of the effects across writing routines. Results indicated that in LCST 1 students' adherence to a Preplanning and Drafting routine moderated text quality in both the experimental and control condition. This is consistent with the outcomes of earlier research (Baaijen et al., 2014; Kieft et al., 2007), which showed that a strong adherence to Preplanning or Drafting routines interacts with the effectiveness of writing instruction based on Preplanning strategies.

In addition, we found an interaction effect of the Preplanning routine with condition and measurement occasion. Remarkably, this effect was located at M2, at which the students from the control group with a high adherence to Preplanning performed better than expected based on the models.

In LCST 2, we added an extra lesson to the learning unit to provide students with metacognitive knowledge about (synthesis) writing processes and encourage them to compare their own writing process to their peers'. No moderator effects for Preplanning or Drafting were found. This suggests that stimulating the acquisition of metacognitive knowledge of (their own) writing processes, as we expected based on the results of Robledo Ramon's research (2016), was

effective in supporting students in making the most effective strategy choices for themselves.

4.3 *RQ 3: Strategy Choices and Motives*

The third research question was to what extent students' strategy choices were related to their writing routine, affected text quality, and reflected their knowledge of the writing process. When students scored high on the Preplanning scale of the WSI, they tended to choose an all-Preplanning path, choosing the Preplanning strategy in both instances. A high score on the Drafting scale only had a small effect in LCST 1: those students tended to choose the all-Preplanning path on the first occasion. There are several possible explanations for these outcomes. First, Kieft et al. (2007) hypothesized that the revising scale of the WSI (measuring the Drafting routine) predominantly reflects reactive revision: evaluating the extent to which the text satisfies the writer's pre-established goals. According to Galbraith and Torrance (2004) reactive revision is intrinsically related to a planning strategy, because this kind of revision focuses on restructuring a preconceived content plan. Second, Kieft et al. (2008) found that the higher students scored on Preplanning writing strategy, the more they appreciated the revising condition. This may indicate that students with a high score on Preplanning in fact did not have a natural preference for Preplanning but, instead, they scored high on the planning scale for other (unknown) reasons, possibly because they had been taught to plan before writing. This may also explain why students in our studies opted for an all-Preplanning path, despite high scores on the Drafting scale: in Dutch writing instruction Preplanning is the dominant strategy and students are used to working according to this strategy.
In LCST 2, the Preplanning strategy was chosen more often than in LCST 1. It is possible that the lesson on writing processes we added to the learning unit in LCST 2, made the students aware of the advantages of the Preplanning strategy when composing synthesis texts. This could have minimized the advantage of students who already showed a high level of Preplanning. In addition, the lesson may have reminded students of the importance of revising, because in this lesson the same amount of attention was paid to (substantial) revision of a text as part of the writing process, as to the Preplanning of the same text, which is usually the focus in Dutch writing education. This unilateral focus on Preplanning might lead to a lack of attention for revising in regular writing lessons, causing students who appear to apply a Drafting routine (quickly generating ideas through the writing of the text), to subsequently ignore the important revision

process: they only marginally reread and hardly revised their texts at all (Van Ockenburg et al., 2018).

Concerning the effects of learning paths on text quality, we found that the all-Preplanning path led to significantly higher text quality scores than the all-Drafting path in LCST 1. We also found an effect of the Preplanning routine: the greater the adherence to this routine (as measured by the WSI), the higher the quality of the text (*ES* = .32), independent of the chosen learning path. The effect of adherence to the Drafting routine only applied to students who followed an all-Drafting path. When students switched between strategies (PP-D or D-PP), the effect of the Drafting routine score was significantly negative compared to the all-Drafting path and marginally significant for the all-Preplanning path. These results support the findings of earlier research (Baaijen et al., 2014; Kieft et al., 2007; Torrance & Galbraith, 2006), indicating that a Drafting strategy is most effective for students with a strong adherence to a Drafting routine. In LCST 2 we found no effects of writing routine or learning path. This indicates that metacognitive knowledge of the (own) writing process is essential for making the strategic choices that fit a personal writing routine.

Finally, we expected that students would demonstrate this metacognitive knowledge of the writing process through the motives they gave for their choices. Our analysis focused on whether they provided insight into their own writing process, either by comments about their regular writing process, and/or their process approach in the subsequent writing task. Unfortunately, the results for LCST 1 were somewhat limited. Students' motives for their strategy choices showed hardly any signs of insight in their own writing process. The analysis of students' motives in LCST 2 revealed that relatively more students appeared to give motives related to possible advantages they might expect due to their choice. This may indicate a goal setting strategy. Still, the differences in metacognitive awareness were small. However, Vandermeulen et al. (2020a) found in their national baseline study that more experienced writers approach the writing process differently than novices and suggest that it is important to take writing processes into account in writing instruction. By focusing on the process aspects, students will become aware of their writing routine and will be able to relate it to the routines of other writers, which can help them become better writers.

4.4 *Limitations and Future Research*

There are some limitations to this research. The first limitation is related to students' relatively low text quality scores. After completing the unit, they still scored only an average of 65 out of 100 points on holistic text quality. This could have

been caused by the fact that this was the first time these students became acquainted with the genre of informative synthesis texts. Furthermore, there are limitations related to determining the effects of the additional lesson on writing processes in LCST 2. The additional lesson was administered as part of the complete unit, and it is therefore impossible to determine what the effect of this individual lesson has been. Nevertheless, the moderation of TQ by writing routines does indeed seem to have disappeared in LCST 2.

Future research could focus on determining the effects on metacognitive knowledge of a single lesson in which students are informed about (their own) writing processes. The results of the current research give rise to the assumption that such a lesson will influence students' strategy choices and their text quality. Furthermore, qualitative research could be used to investigate in depth students' motives for choosing a matching, compensatory or switching learning path and what the effects of these choices are on text quality for students with different writing routines.

4.5 *Educational Implications and Concluding Remarks*

The results of both studies demonstrate that a strategy-based learning unit is generally effective for teaching ninth-grade students the basics of composing synthesis texts. Nonetheless, students appear to differ greatly in the way they approach a synthesis task, which may be partly related to their writing routine. Therefore, offering a choice of different strategies to suit different writing routines could help increase the effectiveness of synthesis writing instruction for all types of writers. However, it is essential that students make a choice based on their knowledge of the writing process, and not just based on convenience or their feelings. Consequently, prior to making a choice, students should acquire sufficient knowledge of the different components of the writing process. In addition, they should make the implicit knowledge about their own writing routine explicit, so they can compare it with other writers' routines and evaluate it. We have argued that learning to compose synthesis texts could help improve the connection between secondary and higher education, because in the Netherlands the way in which source-based tasks should be performed in secondary education is currently not in line with what is expected of students in higher education. Therefore, performing synthesis tasks could help students to deal with these differences more easily. In addition, students generally receive little instruction in source-based writing. We recommend teaching students the basic skills of composing synthesis texts as early as mid-secondary education. In order to offer every school the opportunity to put the results of this research into

practice, the learning unit from this research is freely available, as is an automated version of the Writing Style Questionnaire (both in Dutch).