



UvA-DARE (Digital Academic Repository)

Choosing how to plan informative synthesis texts: Effects of strategy-based interventions on overall text quality

van Ockenburg, L.; van Weijen, D.; Rijlaarsdam, G.

DOI

[10.1007/s11145-021-10226-6](https://doi.org/10.1007/s11145-021-10226-6)

Publication date

2023

Document Version

Final published version

Published in

Reading and Writing

License

Article 25fa Dutch Copyright Act (<https://www.openaccess.nl/en/in-the-netherlands/you-share-we-take-care>)

[Link to publication](#)

Citation for published version (APA):

van Ockenburg, L., van Weijen, D., & Rijlaarsdam, G. (2023). Choosing how to plan informative synthesis texts: Effects of strategy-based interventions on overall text quality. *Reading and Writing*, 36(4), 997–1023. <https://doi.org/10.1007/s11145-021-10226-6>

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.

UvA-DARE is a service provided by the library of the University of Amsterdam (<https://dare.uva.nl>)



Choosing how to plan informative synthesis texts: Effects of strategy-based interventions on overall text quality

Liselore van Ockenburg¹ · Daphne van Weijen¹ · Gert Rijlaarsdam¹

Accepted: 5 November 2021 / Published online: 19 November 2021
© The Author(s), under exclusive licence to Springer Nature B.V. 2021

Abstract

The current two studies examined the effects of an instructional unit on synthesis writing for 9th 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. 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 the unit in two consecutive studies with independent samples, using a quasi-experimental pretest–posttest design with switching panels. Study 1 included three teachers and 152 9th 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 Study 2 with an extra session on metacognitive knowledge about (synthesis) writing processes, enabling them to make better-informed strategy choices. Study 2 included six teachers and 233 students from 10 classes in three schools, all different from Study 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

✉ Liselore van Ockenburg
L.vanOckenburg@uva.nl

¹ Research Institute for Child Development and Education, University of Amsterdam, Nieuwe Achtergracht 127, P.O. Box 15776, 1001 NG Amsterdam, The Netherlands

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 (Referentiekader Taal, 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 (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). Fortunately, a fair amount of research has already been done on learning to write synthesis texts. An analysis of various effective interventions (Van Ockenburg et al., 2019) yielded many learning activities that have proven to be 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.

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 and King (1989) discovered that learning gains were most

evident between 8 and 10th grade, 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 textual 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 focused 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 (Van Ockenburg et al., 2019) 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.

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: pre-planning and post-revising (Galbraith, 1992; Kieft & Rijlaarsdam, 2005). Pre-planning 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 after writing a first draft.

Based on earlier research, we assume that students have mixed routines, i.e., they each Preplan and Draft to some extent. Kieft et al. (2007) reported a correlation between the two dimensions for 10th grade ($r=0.38$, $p<0.001$) and Van Ockenburg et al., (2018) for 9th grade students ($r=0.24$; $p<0.001$). These correlations indicate that Preplanning and Drafting are two related but distinct dimensions of one's writing routine.

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 allowed 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. (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.

Current studies

We report here on two studies. Study 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 Van Ockenburg et al., (2021a). Here, we report on its effectiveness. Subsequently, Study 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 & Galbraith, 2018; Baaijen et al., 2014; Kieft et al., 2007). To our knowledge, no study to date has provided participants with the opportunity to choose between different preplanning 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 showed 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 would be equally effective for all students, regardless of their writing routine. Therefore, we let students choose between two planning strategies: Preplanning and Drafting. On two occasions they had to choose between these two strategies 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 will be considered to be effective when differences in learning paths do not affect the outcomes of the learning unit.

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 1 (Shadish et al., 2002, pp. 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 1 Quasi-experimental design with switching panels (identical for Study 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 |

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

Participants

Study 1

Participants were 152 9th 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; $\chi^2(2) = 0.097$, $p = 0.953$) or age ($m = 14.00$ years, $sd = 0.48$ in group EC; $m = 14.01$ years, $sd = 0.52$ in group CE; $t(150) = 0.128$, $p = 0.988$). Most of the participants (93%) were native speakers of Dutch (L1). The others came from families where parents did not speak Dutch as their first language. Relatively few participants (2.7%) reported having a speech or language disorder. No students from these two groups showed up in our outlier analyses, we therefore included them in the analyses. The parents of the students received an informed consent e-mail and could object to their child's participation, which one parent did.

Study 2

Participants were 233 9th 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; $\chi^2(2) = 2.558$, $p = 0.27$) or age (age $m = 13.93$, $sd = 0.45$ in group EC; $m = 13.92$, $sd = 0.54$ in group CE; $t(191) = 0.116$, $p = 0.908$). Most participants (92%) were L1 Dutch speakers. Only 3.8% of participants indicated having a speech or language disorder. No students from these two groups showed up in our outlier analyses, we therefore included them in the analyses. 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.

Materials

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.

Principle 1. Focus on cognitive strategies. Through a systematic review (Van Ockenburg et al., 2019), 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 process. Earlier research (Baaijen et al., 2014; Kieft et al., 2007; Torrance & Galbraith, 2006) seemed to indicate that writing routines interact with the effectiveness of planning strategies. That is why we paid attention to raising awareness of one's own writing process in the design.

Principle 3. Focus on activating metacognition. 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 2 shows an overview of the operationalization of the design principles in Study 1. This design comprised seven 50-min sessions. The first two sessions aimed to enable students to create a task representation (Table 2, session 0–1). Three sessions followed (Table 2, session 2–4), each instructing one of the three basic synthesizing processes, respectively *selecting*, *organizing* and *connecting*, adding up to a final text version in session 4. The last two sessions of the unit (Table 2, session 5–6) provided students with the opportunity to practice the recently acquired strategies in a new, whole-task synthesis assignment.

Session 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 (Van Ockenburg et al., 2021b). Subsequently, session 1 offered the opportunity for *pre-flection*: creating a shared definition of a well-written synthesis text before engaging in instructional sessions. Then, three instructional sessions (2–4) followed, all offering the same sequence of learning activities, activated by learning through observation. In session 2, students compared-and-contrasted two peers, both relatively proficient in writing but who preferred different writing routines. A film clip showed how they studied the assignment, read sources, and *selected information*. In session 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 session 4, the models *re-organized and connected* the source information and wrote the final draft of their synthesis text. In all instructional sessions (2–4) the students chose between different cognitive strategies, but they only made their choice explicit in sessions 3 and 5, by noting and substantiating it in their workbooks. In session 3, they could choose between two modeled planning strategies, Preplanning or Drafting, after having observed the two models. In session 5, students again chose explicitly which planning strategy

Table 2 General lesson plan of the unit (Study 1)

| Session | Phase | Specification |
|---|-------|--|
| 0 (Whole pre-task experience) | i | Observing a 3-min animated video providing information about integrating source information in synthesis texts |
| 1 (<i>pre-reflection</i>) | ii | Performing a synthesis pre-task to experience the synthesizing process as a whole |
| 2–4 (Acquisition of subskills; similarly structured sessions) | 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 |
| 5–6 (Integration; practice) | 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 |
| | iii | In session 3, noting and motivating choice between the modeled planning strategies |
| | iv | Applying the chosen strategy to a subtask to experience one's preferred strategy |
| | v | Discussing written product(s) to evaluate the result of applying the chosen strategy and estimate whether the result was satisfactory |
| 5 | i | Considering whether to keep or change the planning strategy chosen in session 3 and noting and motivating new choice in workbook |
| 6 | ii | Performing a new synthesis task to experience the use of the chosen strategies in a whole-task assignment |
| | 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 |
| | ii | Giving peers feedback, based on previously discussed text quality criteria (in session 1) to experience the effect of one's text on a reader and determine to what extent the chosen strategy was adequate |

Adapted from Van Ockenburg et al., 2021a

they wanted to apply, after having started a new writing task. They could either stick to their first choice, or switch strategies.

In Study 2, the experimental instructional design was largely consistent with Study 1, except for one extra session in the task representation phase, between the first and the second session of intervention 1. This additional session 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 session and related their scores to the average scores of students from a national baseline study on synthesis writing (Vandermeulen et al., 2020). Next, the students watched an animation video showing different effective writing configurations for synthesis tasks, based on research by Vandermeulen (2020, 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?

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., 2020). 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. endangered wild animals). Content-wise, the sources complemented each other. Students had to write informative synthesis texts of approximately 200 words (min. 180, max. 220).

For Study 2, we decided to replace one measurement task (on the topic of self-driving cars), prompted by comments from several raters in Study 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 Study 1 (artificial food additives) at M2 and chose another task from Vandermeulen et al. and's (2020) study to administer on M3 (big earners). We think we solved the problem raters in Study 1 noticed: for Study 2 we did not receive comments from raters about differences between the assignments.

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 pre-planning 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 validated by Arias-Gundín et al. (2021). The current version was adapted and validated in the context of synthesis tasks (Vandermeulen, 2020). In the last version, it starts with a case description which asks students to imagine that they are writing a synthesis text, after which they are told that

Table 3 Studies 1 and 2: Participants and selected groups for assessing Text Quality

| Study | School | Location | 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 |

the researchers would like to know how the students think they would actually do this, not how they think they should.

Procedures

Implementation

Study 1 took place in five classes of school A (see Table 3). During the first panel, the first author, an experienced teacher in the school subject Dutch language and literature, taught the learning unit to her own two classes (group EC, $n=60$) for two consecutive weeks, as part of the regular curriculum, completing three 50-min sessions per week. The remaining three classes served as control group during this period and continued with their regular curriculum consisting of reading skills instruction, spelling, grammar, and literature, but no writing instruction. During the second panel the conditions were reversed, and the first author's two colleagues taught the unit in three classes (group CE, $n=92$) while panel 1's experimental classes served as the control group.

In the week prior to the start of panel 1, all students took part in the pretest (M1). The students watched a brief instructional video, 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 first author and three research assistants. The other tests were administered between panels 1 and 2 (M2) and, finally, in the week following the completion of the sessions in panel 2 (M3). These tests took place during regular lessons under the supervision of their own teacher.

Study 2 was carried out in 10 classes at all three schools that had registered to participate after an open call on social media (Table 3). During the first panel, three teachers from schools B and C taught the unit as part of the regular curriculum to five classes (group 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 panel

the conditions reversed, and the two teachers at school D taught the unit in five classes (group CE, $n = 129$) while panel 1's experimental classes served as the control group.

We assessed text quality for all Study 1 participants. For practical reasons (time and budget constraints) we assessed text quality for a sub-sample ($n = 162$) of the total group of Study 2 participants ($N = 233$). First, we selected the students who had completed all sessions and measurements and then we randomly selected participants from this subsample while making sure that all participating classes were represented (see Table 3).

Writing routines

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 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 Study 1 ($r = 0.36$, $p < 0.001$), but not in Study 2 ($r = 0.08$, $p = 0.25$).

Text Quality

To assess text quality, 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 0.79 (Study 1) and 0.82 (Study 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 Study 1 21 raters were involved, in Study 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 $\rho = 0.71$ for Study 1 and $\rho = 0.70$ for Study 2.

Table 4 Student variables included in these studies as measured by WSI

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

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

Choices and motives

We analyzed students' workbooks for their strategy choices and whether they showed insight in their own writing process. Students could choose twice between a Preplanning and a Drafting strategy for organizing source information (see Table 2, sessions 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.

Treatment fidelity

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

In Study 1, the first author, who designed the learning unit and materials, taught her regular classes in the first panel. 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 panel, the first author familiarized the two panel 2 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 sessions each: session 1 in three classes; session 2 in three classes; session 4 in four classes; session 5 in two classes; and session 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 session, 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 sessions were taught as intended: 92.2% of the key lesson activities were fully completed, 7.8% partially.

Prior to the start of Study 2, the first author 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 session. The response rate was 93.5%. Sessions 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.

Data analysis

To answer Research Questions 1 and 2, we analyzed 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 analyzed students' strategy choices and their motives substantiating these choices to answer Research Question 3.

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 analyze 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.

RQ 2: moderator effects—writing routines

To explore moderator effects, we expanded the analyses with three additional models. In these models, the centered 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.

RQ 3: strategy choices and motives

In two sessions, students were explicitly invited to choose between two strategies: Preplanning or Drafting. They were free to switch at the second occasion. We analyzed 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, choices and motivation.

In study 1 we analyzed how choice affected TQ for all subjects of whom we had obtained choice data ($n=132$) during the first implementation and coded the motive for their choice based on a small sample (48 students from 2 classes from different teachers and panels) to explore the data. In Study 2 we assessed TQ for a sample that was sufficient based on a power analysis ($n=132$, total set: $N=233$). Of these 132 participants, 72 provided us with choice data and motives.

Table 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 | - | - |

RP=refers to Regular Process; UT=refers to Upcoming Task

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 (a) information about the students' regular writing process was provided, and/or (b) their process in the upcoming writing task. Table 5 shows examples of students' motives and how they were coded.

Based on a protocol with elaborations of the codes for *a* and *b*, and their application to various sample responses, the two coders first analyzed 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* ($k=0.615$) and *b* ($k: 0.744$).

Results

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 panel 1. Subsequently, if, after a significant difference at M2, this effect was absent at M3, that would indicate the effectiveness of the intervention in panel 2, because both groups, EC as well as CE, scored similarly again.

Table 6 Study 1. Mean Holistic Text Quality (and Standard Error), as estimated under model 3

| | M1 | M2 | M3 |
|-------|------------------------|------------------------|------------------------|
| Group | <i>M</i> (<i>SE</i>) | <i>M</i> (<i>SE</i>) | <i>M</i> (<i>SE</i>) |
| EC | 57.87 (1.32) | 68.88 (1.34) | 62.49 (1.33) |
| CE | 58.78 (1.08) | 60.22 (1.07) | 63.51 (1.07) |

EC=panel 1 Experimental Condition, panel 2 Control Condition;
CE=panel 1 Control Condition, panel 2 Experimental Condition

For Study 1, model 3 was significant at $p < 0.001$, 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 6 presents the means and standard errors for holistic text quality scores for Study 1, as estimated under model 3. The estimated effect size under model 3 is $R^2 = 0.321$, $ES = 10.3\%$. 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.

For Study 2, model 3 was significant at $p = 0.001$, 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 7 presents the means and standard errors for holistic text quality scores for Study 2, as estimated under model 3. The estimated effect size under model 3 is $R^2 = 0.467$, $ES = 20.8\%$.

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

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 Study 1 the interaction between condition and measurement occasion was moderated by Preplanning (Model 7: $p = 0.015$). 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 $p = 0.05$, $B = 1.91$, $SE = 0.96$), which indicates there was a general contribution of Drafting to text quality, but smaller than the general Preplanning effect. Figures 1 and 2 respectively visualize the complex interaction in a switching replication design for Preplanning and Drafting in Study 1. The estimated effect sizes

Table 7 Study 2. Mean Holistic Text Quality (and standard error), as estimated under model 3

| | M1 | M2 | M3 |
|-------|--------------|--------------|--------------|
| Group | $M (SE)$ | $M (SE)$ | $M (SE)$ |
| EC | 55.63 (1.51) | 63.66 (1.51) | 66.41 (1.51) |
| CE | 51.02 (1.48) | 53.89 (1.48) | 66.55 (1.48) |

EC=panel 1 Experimental Condition, panel 2 Control Condition;
CE=panel 1 Control Condition, panel 2 Experimental Condition

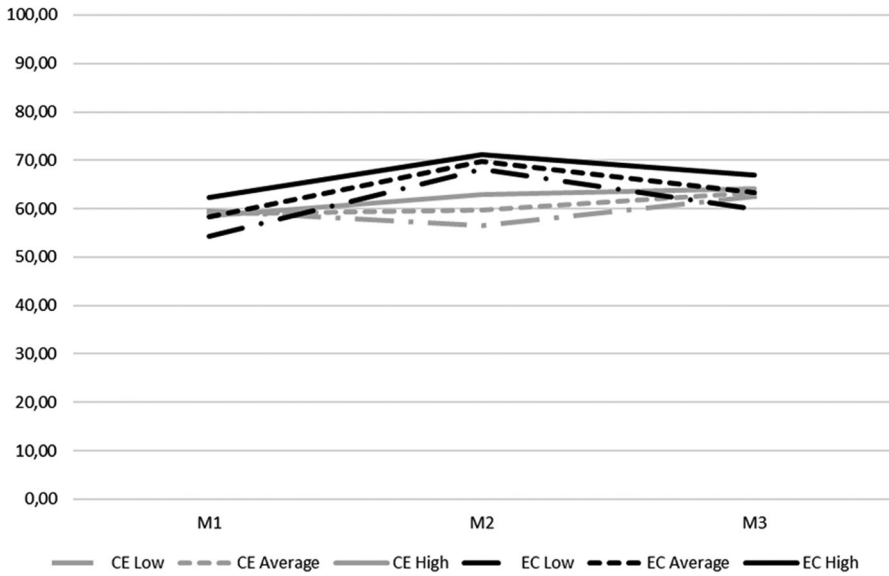


Fig. 1 Study 1. Effect of Preplanning on Holistic Text Quality. *Note.* CE=panel 1 Control Condition, panel 2 Experimental Condition; EC=panel 1 Experimental Condition, panel 2 Control Condition; Average=Level of Planning; High=average + 1 *sd*; Low = average-1 *sd*

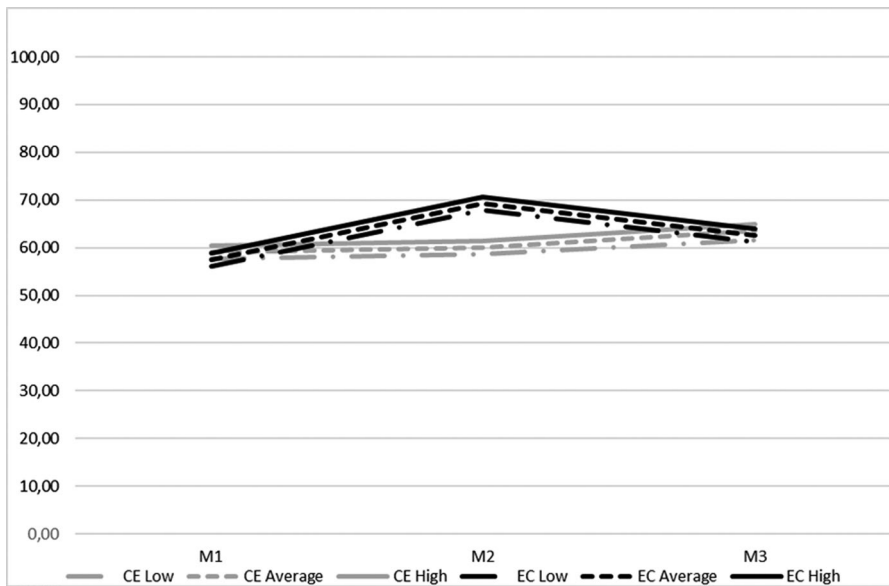


Fig. 2 Study 1. Effect of Drafting on Holistic Text Quality. *Note.* CE=panel 1 Control Condition, panel 2 Experimental Condition; EC=panel 1 Experimental Condition, panel 2 Control Condition; Average=Level of Planning; High=average + 1 *sd*; Low = average-1 *sd*

Table 8 Students' distribution over strategy choices (in percentages) for Studies 1 (n = 132) and 2 (n = 72)

| Study | | 1 | | 2 | |
|----------|-------------|-----|-----|-----|-----|
| Session | | a | b | a | b |
| Strategy | Preplanning | 44% | 50% | 50% | 58% |
| | Drafting | 56% | 50% | 50% | 42% |

Table 9 Students' distribution over learning paths (in percentages) for Studies 1 (n = 132) and 2 (n = 72)

| Study | | 1 | 2 |
|---------------|-----------------|--------|-------|
| Learning path | All-Preplanning | 38.6 % | 40.3% |
| | All-Drafting | 44.7% | 31.9% |
| | Switching | 16.7% | 27.8% |

for Preplanning under model 7 are $R^2 = 0.321$, $ES = 15.5\%$, and for Drafting under model 4 $R^2 = 0.339$, $ES = 11.5\%$.

In Study 2, holistic text quality and interaction between condition and measurement occasion were not moderated by Preplanning or Drafting. Appendix B presents all model comparisons for Study 1 and 2.

RQ 3: strategy choices and motives

In two sessions, 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 8 shows the choice distribution for both studies.

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 9 shows that in Study 2 a larger proportion of students chose to switch strategies and that in Study 1 the all-Drafting path was more popular than in Study 2.

The all-Preplanning path was preferred when students scored relatively high on the Preplanning scale of the WSI (Effect size in Study 1 resp. 2, choice 1 resp 2: 0.28, 0.43, 0.45, 0.42). The score on the Drafting scale only had a small effect in Study 1, session 3: students with higher scores on the Drafting scale tended to choose the all-Preplanning path ($ES = 0.22$).

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 Study 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) ($B = 3.50$, $SE = 1.46$, $p = 0.018$). A general effect of Preplanning was also observed ($B = 3.30$, $SE = 1.42$, $p = 0.022$), 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 ($B = 4.56$, $SE = 1.46$, $p = 0.002$) only applied 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) ($B = 5.96$,

$SE=2.51, p=0.019$). The effect was marginally statistically negative for the all-Preplanning path ($B=-4.16, SE=2.17, p=0.058$). In Study 2 we found no effects of learning path or writing routine.

In students' motives substantiating their strategy choices, we did not find differences between Studies 1 and 2 in terms of references to the regular writing process ($\chi^2(2)=0.77; p=0.67$). However, there was a difference in references to the upcoming task ($\chi^2(2)=6.69, p=0.03$). In Study 2, both coders agreed in 36.5% of the cases that the motives referred explicitly to the upcoming task, while in Study 1 this was lower (21%).

Discussion

The current study was set up to examine the efficacy of a synthesis writing intervention (Study 1) and the effectiveness of the redesign on a larger scale (Study 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 related to their writing routine, affected text quality, and reflected their knowledge of the writing process (RQ3).

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 Studies 1 and 2. Furthermore, the average text quality at M3, the delayed post-test for group EC, suggests that the effect of the experimental condition in panel 1 was maintained in panel 2, while in both studies no difference between the conditions at M3 was observed. These results further support findings that strategy instruction is an effective approach for improving students' synthesis writing (e.g., Barzilai & Ka'adan, 2017; Britt & Aglinskias, 2002; Kirkpatrick & Klein, 2009; Martinez et al., 2015; Zhang, 2013) and that a systematic analysis of intervention studies (Van Ockenburg et al., 2019), provides a good starting point for developing effective synthesis writing instruction.

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 Study 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 Study 2, we added an extra session 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.

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 Study 1: those students tended to choose the all-Preplanning path on the first occasion. There are several possible explanations. 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 be a reason as well that students in our studies opted for an all-Preplanning path, despite having 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 Study 2, the Preplanning strategy was chosen more often than in Study 1. It is possible that the session on writing processes we added to the learning unit in Study 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 session may have reminded students of the importance of revising, because in this session the

same amount of attention is 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 creating ideas through the writing of the text), to subsequently ignore the important revision process: they only marginally reread and hardly revise 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 Study 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 ($ES=0.32$), 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 Planning-by-drafting strategy is most effective for students with a strong adherence to a Drafting routine. In Study 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 Study 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 Study 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., (2020) found in their national baseline study that more experienced writers approach the writing process differently than novices and they 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.

Limitations and future research

There are some limitations to this research. The first limitation is the relatively low scores of the students on text quality. 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. Further, there are limitations related to determining the effects of the additional session on writing processes in Study 2. The additional session was administered as part of the complete unit, and it is therefore impossible to determine what the effect of this individual session has been. Nevertheless, the moderation of TQ by writing routines does indeed seem to have disappeared in Study 2.

Future research could focus on determining the effects on metacognitive knowledge of a single session in which students are informed about (their own) writing processes. The results of the current research give rise to the assumption that such a session 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.

Educational implications and concluding remarks

The results of both studies demonstrate that a strategy-based learning unit is generally effective for teaching 9th 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).

Acknowledgements Special thanks to Huub van den Bergh (Utrecht University) for helping us with the statistical analyses, and to the students and colleagues of the participating schools for their great efforts.

Authors' contributions All authors contributed to the studies' conception and design. Material preparation, data collection and analysis were performed by Liselore van Ockenburg, Daphne van Weijen and Gert Rijlaarsdam. The first draft of the manuscript was written by all three authors and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Funding This work is part of the research program ‘Promotiebeurs voor leraren’ with project number 023.007.011, which is financed by the Netherlands Organization for Scientific Research (NWO).

Data availability The datafiles will be published via figshare, the University of Amsterdam’s repository.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

Ethical approval Study 1: ERB_Approval_9673; Study 2: ERB_Approval_10908.

Reference

- Arias-Gundín, O., & Fidalgo, R. (2017). Writer profile as a modulating variable of processes involved in written composition in undergraduate students. *European Journal of Investigation in Health, Psychology and Education*, 7(1), 59–68. <https://doi.org/10.3390/ejihpe7010005>
- Arias-Gundín, O., Real, S., Rijlaarsdam, G., & López, P. (2021). Validation of the writing strategies questionnaire in the context of primary education: A multidimensional measurement model. *Frontiers in Psychology*. <https://doi.org/10.3389/fpsyg.2021.700770>
- Baaijen, V. M., Galbraith, D., & de Gloppe, K. (2014). Effects of writing beliefs and planning on writing performance. *Learning and Instruction*, 33, 81–91. <https://doi.org/10.1016/j.learninstruc.2014.04.001>
- Baaijen, V. M., & Galbraith, D. (2018). Discovery through writing: Relationships with writing processes and text quality. *Cognition and Instruction*, 36(3), 199–223. <https://doi.org/10.1080/07370008.2018.1456431>
- Barzilai, S., & Ka’adan, I. (2017). Learning to integrate divergent information sources: The interplay of epistemic cognition and epistemic metacognition. *Metacognition and Learning*, 12(2), 193–232. <https://doi.org/10.1007/s11409-016-9165-7>
- Bonset, H. (2010). Nederlands in het voortgezet en hoger onderwijs: Hoe sluit dat aan? Deel 1 [The subject of Dutch in secondary and higher education: How does that connect? Part 1.]. *Levende Talen Magazine*, 97(3), 16–20.
- Boscolo, P., Arfé, B., & Quarisa, M. (2007). Improving the quality of students’ academic writing: An intervention study. *Studies in Higher Education*, 32(4), 419–438. <https://doi.org/10.1080/03075070701476092>
- Britt, M. A., & Aglinskias, C. (2002). Improving students’ ability to identify and use source information. *Cognition and Instruction*, 20(4), 485–522. https://doi.org/10.1207/S1532690XCI2004_2
- Butler, D. (1998). The strategic content learning approach to promoting self-regulated learning: A report of three studies. *Journal of Educational Psychology*, 90(4), 682–697. <https://doi.org/10.1037/0022-0663.90.4.682>
- Daher, T. A., & Kiewra, K. A. (2016). An investigation of SOAR study strategies for learning from multiple online resources. *Contemporary Educational Psychology*, 46, 10–21. <https://doi.org/10.1016/j.cedpsych.2015.12.004>
- Darowski, E. S., Helder, E., & Patson, N. D. (2020). Explicit writing instruction in synthesis: Combining in-class discussion and an online tutorial. *Teaching of Psychology*. <https://doi.org/10.1177/0098628320979899>

- Darowski, E. S., Patson, N. D., & Helder, E. (2016). Implementing a synthesis tutorial to improve student literature reviews. *Behavioral and Social Sciences Librarian*, 35(3), 94–108. <https://doi.org/10.1080/01639269.2016.1243437>
- De Smedt, F., Van Keer, H., & Merchie, E. (2016). Student, teacher and class-level correlates of Flemish late elementary school children's writing performance. *Reading and Writing*, 29(5), 833–868. <https://doi.org/10.1007/s11145-015-9590-z>
- Galbraith, D. (1992). Conditions for discovery through writing. *Instructional Science*, 21(1–3), 45–71.
- Galbraith D., & Torrance M. (2004). Revision in the context of different drafting strategies. In: L. Allal, L. Chanquoy, & P. Largy (Eds.), *Revision Cognitive and Instructional Processes. Studies in Writing* (Vol 13). Springer. https://doi.org/10.1007/978-94-007-1048-1_5
- Gil, L., Vidal-Abarca, E., & Martínez, T. (2008). Eficacia de tomar notas para integrar información de varios textos. [Effectiveness of Notetaking for Integrating Information from Various Texts]. *Infancia y Aprendizaje*, 31(2), 259–272. <https://doi.org/10.1174/021037008784132905>
- Hayes, J. R. (2012). Modeling and remodeling writing. *Written Communication*, 29(3), 369–388.
- Karweit, N. (1984). Time-on-task reconsidered: Synthesis of research on time and learning. *Educational Leadership*, 41(8), 32–35.
- Kellogg, R. T. (1996). A model of working memory in writing. In C. M. Levy & S. Ransdell (Eds.), *The science of writing: Theories, methods, individual differences, and applications* (pp. 57–72). Lawrence Erlbaum Associates.
- Kellogg, R. T. (1987). Writing performance: Effect of cognitive strategies. *Written Communication*, 4, 269–298. <https://doi.org/10.1177/0741088387004003003>
- Kellogg, R. T. (1988). Attentional overload and writing performance: Effects of rough draft and outline strategies. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 14, 355–365. <https://doi.org/10.1037/0278-7393.14.2.355>
- Kellogg, R. T. (2001). Competition for working memory among writing processes. *American Journal of Psychology*, 114, 175–191. <https://doi.org/10.2307/1423513>
- Kellogg, R. T. (2008). Training writing skills: A cognitive developmental perspective. *Journal of Writing Research*, 1, 1–26. <https://doi.org/10.17239/jowr-2008.01.01.1>
- Kieft, M., & Rijlaarsdam, G. (2005). Schrijftaken en schrijverstypen. [Writing tasks and writer types.]. *Levende Talen Magazine*, 92(2), 9–12.
- Kieft, M., Rijlaarsdam, G., Galbraith, D., & van den Bergh, H. (2007). The effects of adapting a writing course to students' writing strategies. *British Journal of Educational Psychology*, 77(3), 565–578. <https://doi.org/10.1348/096317906x120231>
- Kieft, M., Rijlaarsdam, G., & Van den Bergh, H. (2008). An aptitude–treatment interaction approach to writing-to-learn. *Learning and Instruction*, 18(4), 379–390. <https://doi.org/10.1016/j.learninstruc.2007.07.004>
- Kirkpatrick, L. C., & Klein, P. D. (2009). Planning text structure as a way to improve students' writing from sources in the compare-contrast genre. *Learning and Instruction*, 19(4), 309–321. <https://doi.org/10.1016/j.learninstruc.2008.06.001>
- Klein, P. D., & Boscolo, P. (2016). Trends in research on writing as a learning activity. *Journal of Writing Research*, 7(3), 311–350. <https://doi.org/10.17239/jowr-2016.07.03.01>
- Leijten, M., Van Waes, L., Schrijver, I., Bernolet, S., & Vangehuchten, L. (2017). Hoe schrijven masterstudenten synthesesetksten? Het brongebruik van gevorderde schrijvers in kaart gebracht. [How do master students write synthesis texts? Source use of advanced writers mapped.] *Pedagogische Studiën*, 94(4), 233–253. <https://pedagogischestudien.nl/download?type=document&identificer=641006>
- Limpo, T., & Alves, R. A. (2018). Effects of planning strategies on writing dynamics and final texts. *Acta Psychologica*, 188, 97–109. <https://doi.org/10.1016/j.actpsy.2018.06.001>
- Lundstrom, K., Diekema, A. R., Leary, H., Haderlie, S., & Holliday, W. (2015). Teaching and learning information synthesis: An intervention and rubric based assessment. *Communications in Information Literacy*, 9(1), 4. <https://doi.org/10.15760/comminfolit.2015.9.1.176>
- Luo, L., & Kiewra, K. A. (2019). Soaring to successful synthesis writing. *Journal of Writing Research*, 11(1), 163–209. <https://doi.org/10.17239/jowr-2019.11.01.06>
- Martínez, I., Mateos, M., Martín, E., & Rijlaarsdam, G. (2015). Learning history by composing synthesis texts: Effects of an instructional programme on learning, reading and writing processes, and text quality. *Journal of Writing Research*, 7(2), 275–302. <https://doi.org/10.17239/jowr-2015.07.02.03>

- Mateos, M., Martín, E., Villalón, R., & Luna, M. (2008). Reading and writing to learn in secondary education: Online processing activity and written products in summarizing and synthesizing tasks. *Reading and Writing*, 21(7), 675–697. <https://doi.org/10.1007/s11145-007-9086-6>
- Meestringa, T. (2011). *Propedeutische schrijftaken. Analyse van 18 schrijftaken uit Nederlandse en Vlaamse opleidingen hoger onderwijs*. [Propaedeutic writing tasks. Analysis of 18 writing tasks from Dutch and Flemish higher education courses.] SLO.
- Meestringa, T., & Ravestloot, C. (2013). Schrijven in de Tweede Fase [Writing in Grades 10 to 12.]. *Levende Talen Tijdschrift*, 14(6), 6–10.
- Ministerie O. C. W. (2009). *Referentiekader Taal en Rekenen*. [Frame of reference Language and Mathematics.] SLO. https://www.slo.nl/publish/pages/5901/referentiekader_taal_en_rekenen_referentieniveaus.pdf
- O'Donnell, C. L. (2008). Defining, conceptualizing, and measuring fidelity of implementation and its relationship to outcomes in K–12 curriculum intervention research. *Review of Educational Research*, 78(1), 33–84. <https://doi.org/10.3102/0034654307313793>
- Oudakker, M., & Groenendijk, R. (2020). Kan schrijfvaardigheid in het voortgezet onderwijs (nog) beter aansluiten bij het hoger onderwijs? [Can writing skills in secondary education match higher education (even) better?]. *Levende Talen Magazine*, 107(8), 10–14.
- Rijlaarsdam, G. & Van den Bergh, H. (2004). Effective learning and teaching of writing: Student involvement in the teaching of writing. In G. Rijlaarsdam, H. van den Bergh, & M. Couzijn, (Eds.). *Studies in Writing*. (Vol. 14, pp. 1–16). Springer. https://doi.org/10.1007/978-1-4020-2739-0_1.pdf
- Rijlaarsdam, G., Van den Bergh, H., Couzijn, M., Janssen, T., Braaksma, M., Tillema, M., Van Steendam, E., & Raedts, M. (2012). Writing. In: Harris K. R., Graham S., Urdan T., Bus A. G., Major S., & Swanson H. L. (Eds.), *APA handbooks in psychology*[®]. *APA educational psychology handbook, Vol. 3. Application to learning and teaching* (p. 189–227). American Psychological Association. <https://doi.org/10.1037/13275-009>
- Robledo-Ramón, P. (2016) Eficacia de un programa de instrucción estratégica para la mejora de las síntesis escritas en alumnado universitario. [Efficacy of a strategic instruction program for the improvement of written syntheses in university students.] In *Variables Psicológicas y Educativas para la intervención en el ámbito escolar: Volumen II* (pp. 45–52). ASUNIVEP. https://doi.org/10.1787/ayuda_sintesis-2013-9-es
- Segev-Miller, R. (2004). Writing from sources: The effect of explicit instruction on college students' processes and products. *L1-Educational studies in language and literature*, 4(1), 5–33. <https://doi.org/10.1023/b:esll.0000033847.00732.af>
- Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference* (2nd ed.). Houghton, Mifflin and Company. <https://doi.org/10.1086/345281>
- Solé, I., Miras, M., Castells, N., Espino, S., & Minguela, M. (2013). Integrating information: An analysis of the processes involved and the products generated in a written synthesis task. *Written Communication*, 30(1), 63–90. <https://doi.org/10.1177/0741088312466532>
- Spivey, N., & King, J. R. (1989). Readers as writers composing from sources. *Reading Research Quarterly*. <https://doi.org/10.1598/rrq.24.1.1>
- Torrance, M., & Galbraith, D. (2006). The Processing demands of writing. In C. A. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (pp. 67–80). The Guilford Press.
- Van den Bergh, H., & Eiting, M. H. (1989). A method of estimating rater reliability. *Journal of Educational Measurement*, 26(1), 29–40.
- Vandermeulen, N. (2020). Synthesis writing in upper-secondary education: From a baseline of texts and processes to process-oriented feedback. [Doctoral Thesis, University of Antwerp]. <https://lirias.kuleuven.be/3070762>
- Vandermeulen, N., De Maeyer, S., Van Steendam, E., Lesterhuis, M., Van den Bergh, H., & Rijlaarsdam, G. (2020). Syntheseschrijven in het hoger secundair onderwijs in Nederland in kaart gebracht. Een nationale peiling naar tekstkwaliteit, schrijfproces en schrijverskenmerken. [Mapping synthesis writing in various levels of Dutch upper-secondary education A national baseline study on text quality, writing process and students' perspectives on writing]. *Pedagogische Studiën*, 97(3), 187–236. <https://pedagogischestudien.nl/search?identificer=717755>
- Van Ockenburg, L., van Weijen, D., & Rijlaarsdam, G. (2018). Syntheseteksten leren schrijven in het voortgezet onderwijs; Het verband tussen schrijfplan en voorkeur voor leeractiviteiten. [Learning to write synthesis texts in secondary education; The association between writing routine and preference for learning activities.] *Levende Talen Tijdschrift*, 19(2), 3–14.

- Van Ockenburg, L., van Weijen, D., & Rijlaarsdam, G. (2019). Learning to write synthesis texts: a review of intervention studies. *Journal of Writing Research, 10*(3), 401–428. <https://doi.org/10.17239/jowr-2019.10.03.01>.
- Van Ockenburg, L., van Weijen, D., & Rijlaarsdam, G. (2021a). Learning how to synthesize: The design and evaluation of a reading-writing learning unit for high-school students. *L1 Educational Studies in Language and Literature, 21*(running issue). <https://doi.org/10.17239/l1esll-2021.21.01.06>.
- Van Ockenburg, L., van Weijen, D., & Rijlaarsdam, G. (2021b). Een mini-interventie: het effect van een animatievideo van drie minuten op de integratie van broninformatie in syntheseseteksten. [A brief intervention: does a three-minute animation video promote the integration of source information in synthesis texts?] *Pedagogische Studiën, 98*(1), 46–66. <https://pedagogischestudien.nl/search?identifier=721616>.
- White, M. J., & Bruning, R. (2005). Implicit writing beliefs and their relation to writing quality. *Contemporary Educational Psychology, 30*, 166–189. <https://doi.org/10.1016/j.cedpsych.2004.07.002>
- Zhang, C. (2013). Effect of instruction on ESL students' synthesis writing. *Journal of Second Language Writing, 22*(1), 51–67. <https://doi.org/10.1016/j.jslw.2012.12.001>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.