Synthesis writing in science orientation classes: An instructional design study

Edith Alkema, Daphne van Weijen & Gert Rijlaarsdam

University of Amsterdam | The Netherlands

Abstract: This study tested an instructional design to improve students' synthesis performance in a specific academic subject, Science Orientation, which aimed to teach students how to critically evaluate scientific debates. The design included three components: 1) students construct a task definition via a learning strategy based on comparing and contrasting texts and processes, 2) students comprehend source information via a read-stop-think-note strategy, and 3) students connect source information critically via a semantic-textual transformation strategy.

After several design iterations, the instructional design was tested in a quasi-experimental experiment with a pretest-posttest. Seven 10th grade classes participated in the intervention (n=129), four in the control condition (n=86). The design seemed feasible for teachers, students completed most learning tasks as intended and evaluated the course positively. Furthermore, texts written in the experimental condition at posttest were rated significantly higher than those written in the control condition on the instructed aspects: representation of source information, intertextual integration, and critical stance. This instructional design appears to have potential for helping students improve their comprehension of scientific debates and comprehensive writing. In the discussion we propose that the instructional design might be a general format for learning to synthesize domain specific information from contrasting sources.

Keywords: synthesis writing, science orientation, critical thinking, instructional design, observational learning

journal of WRITING RESEARCH

Alkema, E., Van Weijen, D., & Rijlaarsdam, G. (2023). Synthesis writing in science orientation classes: An instructional design studio. *Journal of Writing Research, 15*(1), 133-165. https://doi.org/10.17239/jowr-2023.15.01.06

Contact: Edith Alkema, University of Amsterdam, Research Institute Child Development and Education, Nieuwe Achtergracht 127, 1018 WS Amsterdam | The Netherlands – edithalkema@gmail.com - ORCID: https://orcid.org/0000-0002-8016-5205

1. Introduction

In most school subject classes, in secondary education, students often have to write texts based on domain specific source materials. Instead of integrating, students tend to copy and paste parts of sources in the new text, so called patchwriting (e.g., Van Ockenburg et al., 2018, 2019). Such a representation of what integration requires, hinders the writing of synthesis texts, that accurately represent the sources and are comprehensible without readers having access to the source materials (Spivey, 1991; Van Ockenburg et al., 2019).

Students must be taught to master synthesizing skills (Mateos, 2008). Previous research has already been conducted on teaching synthesis writing strategies in different disciplines, including language classes (Casado-Ledesma et al., 2021; González-Lamas et al., 2016; Van Ockenburg et al., 2021a), history (Van Driel et al., 2022; Martinez et al. 2015) and psychology (Cuevas et al., 2016; Granado-Peinado et al., 2019; Luna et al., 2020; Mateos et al., 2008, 2018, 2020). The main aim of the current study was to design an introductory course to teach students how to synthesize in another specific discipline: Science Orientation. This school subject aims to prepare students to think critically about scientific issues, laying the foundation for scientific thinking and academic writing in all subjects.

In secondary education in the Netherlands, most source-based writing, including synthesis tasks, occurs in subjects such as History, Geography and Philosophy, with a focus on information quality. Science Orientation is a relatively new school subject, with a curriculum based on three pillars: 1) knowledge about scientific debates, 2) philosophy of science, and 3) academic skills such as scientific reading, academic writing, and critical thinking about well-known scientific issues (www.wetenschapsorientatie.nl). Science Orientation is an optional part of the curriculum; schools can choose to offer it as a school subject. Synthesis tasks' topics in Science Orientation classes are (semi-) scientific, related to all sciences (natural, social and formal) and current affairs. Students usually work with source texts focused on conflicting scientific perspectives, such as Aristotle's teleologic model (History), genetic engineering (Biology), and whether man has no free will (Psychology/Philosophy).

We set up this paper as a design study, focusing first on the instructional design's theoretical validation, followed by its empirical validation in terms of feasibility and effectiveness.

2. Theoretical background: design principles

Design principles can be seen as a set of guidelines which indicate the main components that should be included in an educational design, based on findings from earlier empirical studies (see Rijlaarsdam et al., 2017; Van Ockenburg, et al., 2021b). In this section we present the design principles for the intervention's three components.

2.1 Component 1: Students *construct a task definition* via a learning strategy based on comparing and contrasting texts and processes

When it comes to an entirely new, unfamiliar task for students, the first action should be to orient themselves, to form a task definition (Plakans & Liao, 2018): what is expected as outcome, and how can you create it? The instructional component must therefore provide students with a learning strategy, generally applicable in new task situations, but which they learn in a specific task situation.

To acquire new knowledge and skills in a motivated and fruitful way, Merrill (2002) formulated as his first Design Principle that students must experience the whole task, to experience how it works, and what they need to learn to improve task performance. Such a task experience lays the foundation for new information about task outcome and task processes.

To gain insight into how synthesizing works, the best option is to observe such processes and compare and evaluate them (Merrill, 2002; Rijlaarsdam et al., 2017). Observation must be a goal-directed learning activity, while guidance by evaluative questions can help learners focus (Bandura, 1986; Braaksma et al., 2006). Under specific circumstances observational learning can change both students' writing processes (Braaksma et al., 2004) and text quality (Braaksma, 2002). Braaksma et al. (2006) reported on students' learning paths during observational learning and showed that comparing and evaluating were two key activities.

If students are to gain insight as to what they should and can achieve, they must develop insight into the qualities of the intended outcome, the text. According to Hillocks' meta-analysis, comparing texts is more effective (Text scales ES .36) than analyzing model texts (ES .22; Hillocks, 1986). Working with text scales may induce comparing and contrasting, which provide insight into the relative position of texts, and a holistic insight into the dynamic contribution of text features to text quality. The basic learning activities applied in text scales are ranking texts, exploring and verbalizing differences by comparing and contrasting texts, positioning one's own written texts on the text scale, checking, extending and generalizing information from time to time in small groups and ultimately with the whole class. So, these learning activities are based on learning via stepwise inductive inquiry: initial 'naive' findings (task definitions) are sharpened using new information and sharing findings, as this may best reflect a natural learning strategy to gain insight in a certain task.

Based on the above, design principle 1 of this study is: If students must construct a task definition, they should:

- experience the task at hand by performing a synthesis task,

- observe and evaluate a model demonstrating the process of synthesizing given information in a new text,
- formulate a definition of what a good synthesis text entails, via a sequence of inductive inquiry actions including ranking several synthesis texts on quality, discussing the criteria for the ranking with others, reconsidering the ranking after criteria are provided by the teacher.

2.2 Component 2: Students *comprehend source information* via a read-stop-think-note strategy

Comprehending source information for writing a synthesis text may require a change in many students' attitude towards information. In the Netherlands, students are used to writing argumentative texts, presenting their own opinion about an issue, and adding arguments from sources to support that opinion. They may pick from the sources what they need to support their position or to attack a counter position. For a synthesis text, however, defined as a text that accurately represents the source materials, a more objective stance towards source materials is needed (Stein, 1989). At the same time, students must read (semi) scientific and mostly new information and may lack a frame of reference to distinguish main ideas from less important ones (Ter Beek et al., 2018). Therefore, a component that aims at comprehending source materials should focus on learning to summarize texts objectively. Adding the goal to write a comprehensible text may increase the cognitive load of the reading comprehension process, as it might influence the reader's standards of coherence which are the "criteria that a reader has for what constitutes adequate comprehension and coherence in a particular reading situation" (Van den Broek & Helder, 2017, p. 364). These criteria can be influenced by a reader's goals for reading. If a student wishes to write a comprehensible text based on several sources, then it is likely that their standards of coherence will be higher as they wish to comprehend the sources' content on a deeper level, which in turn requires more cognitive effort. An often-implemented instructional strategy is to break down a complex task into linear components – comprehension first, to some degree, then text production. Instruction must then first support students' comprehension process. This does not imply that writing should not play a role during the comprehension process. Escorcia et al. (2017) found a significant positive influence of investing in writing activities, such as note-taking, that support the comprehension process with regard to content. In addition, synthesis writing process research reports that switching between sources and the text-written-so-far contributed to greater text quality (Martinez et al., 2015; Solé, et al., 2013) and the more often this occurs during the initial stage of the process, the better the resulting text (Vandermeulen et al., 2020b). Thus, when designing a separate comprehension phase, this implies that writing activities must be part of such a comprehension strategy, to support reprocessing source information into comprehension.

The foundation for our design was Carr and Ogle's *reading-thinking*/K-W-L reading method, "a strategy for comprehension and summarization" (Carr & Ogle, 1987, p. 626). The K stands for what the reader already knows (*Known*), the W for what the reader would like to know (*Want*) and the L for what has been learned

(*Learned*) (Carr & Ogle, 1987). While reading the text, students had to pause after one or two paragraphs, to think about the content, "monitor their comprehension" (Carr & Ogle, p. 627), and then make notes in the K-W-L matrix with three columns, for information about already existing knowledge (K, asking oneself what do I know?), desired knowledge (W, what do I want to know?) and learned knowledge (L, what have I learned?). This question-based summary method proved successful for selecting question-related relevant information from the text, beneficial for students' concentration, and which promoted their self-confidence. The strategy then consists of three key-elements: self-questioning, note-taking and using a summary scheme, each of which is described briefly below.

Self-questioning starts with retrieving prior topic knowledge, a necessary first step for formulating questions and reading the sources with a somewhat critical stance (Salmeron et al., 2018; Surma et al., 2019; Van Gelder, 2005; Yeh, 2009; see also Meneses et al. (2023) and Törmälä & Kulju (2023) in this special issue). It is a goal-setting activity which has proven to be an effective comprehension strategy (Hattie, 2009). It steers the search for sources' main ideas, via guidance and intentional content searches (Carr & Ogle's 1987 'the want to know'). It also supports comprehension monitoring, against the standard of coherence that the goal induces (Van den Broek & Helder, 2017). Questioning oneself works better than rereading the text (Surma et al., 2019), especially when the answers to one's own 'internal dialogue' are written down: what is written down, is remembered better (Venneker, 2017), and what is remembered better, is reprocessed better mentally, because the brain is activated during text formulation.

Alternating read-stop-think-note urges students to focus on comprehension and connecting information from external and internal sources and to monitor comprehension. Making notes while summarizing is an effective supporting action (Kiewra, 1989). The stop creates room for comprehension activities. Especially when students must write about relatively difficult issues or new material, as is the case in the school subject Science Orientation, slowing down the comprehension process to reprocess and reformulate the content is essential. Notes are a form of self-explanation and are likely to activate critical reading and thus, facilitate the successful selection of main ideas (Mason et al., 2012; Surma et al., 2019; Weston-Sementelli et al., 2016). If the strategy requires writing notes in *one's own words*, it may ease the generation of related knowledge from one's own memory (Galbraith & Baaijen, 2018).

Sorted and connected source information in a summary scheme is the hinge between the comprehension and the text production process (Barzilai et al., 2018; Hammann & Stevens, 2003; Kirkpatrick & Klein, 2016; Reynolds & Perin, 2009; Van Ockenburg et al., 2018). The scheme must be adapted to the task at hand and must be flexible so that students can easily change orders and notes. We will opt for a scheme that implicitly guides the search for main elements: the issue at stake, standpoints, arguments, using post-it notes for each note per source and distinct colors per source. This scheme must provide room for activities in the next phase when students reprocess their notes semantically and critically.

Strategies are best instructed via direct instruction, modeling, and forms of practice (Dewitz et al., 2009; Merrill, 2002). A condition for the successful implementation of the read-stop-think-note strategy is that texts are relatively short to execute the strategy, and to experience progression; the summary scheme is filled in step-by-step, in relatively short read-stop-think-note cycles.

Based on these considerations design principle 2 is: If students must acquire a typical strategy for comprehending information from multiple sources, they should

- generate topic knowledge via a pre-prepared prompt,
- observe a competent model demonstrating the *read-stop-think-note* strategy for comprehending and summarizing a source, reconstructing arguments and sub arguments concisely and in the model's own words, all of this shown in a flexible scheme based on post-its,
- practice the strategy, and
- evaluate their content schemes by comparing these in pairs, and eventually revise their scheme.

The result of this activity is a summary for each paragraph on post-it notes, all pasted in a scheme, following the source text's structure.

2.3 Component 3: Students *connect source information critically* via a semantic-textual transformation strategy.

Before the actual composing phase, one more conceptual activity must take place. Luo & Kiewra (2019) found that reading and summarizing through their SOAR method (*Selection, Organization, Association, Regulation*) yielded better selection and organization of main ideas, but not necessarily better integration of source content. They proposed that integration requires relationship prompting, which should trigger an internal memory search for overarching concepts and labels. To teach students this process, we had to design a conceptual strategy, that creates intertextual relations put into words, as a bridge between the conceptual content and the structure of the text-to-be-produced. For this strategy, we studied materials from another research group (Casado-Ledesma et al., 2021; Martinez et al., 2015; Mateos et al., 2020) and designed a specific strategy that we trialed in two earlier iterations of the instructional unit. We will now briefly describe the strategy's content and means of instruction.

This strategy consisted of two phases, a semantic, and a text production phase. The semantic phase had three steps: 1) connecting arguments intertextually (semantic analysis, comprehension check), 2) generating and formulating an overarching concept that represents such a connection, which can probably serve as the topic sentence of a synthesis text paragraph, and 3) generating a critical note for each overarched argument pair. In Science Orientation such a critical stance to scientific information and reasoning is an essential learning aim (see also Meneses et al. (2023) in this special issue).

. Students learn to test the claims and arguments critically against their own knowledge base. In the case of contrasting information, the critical note is often a reasoned choice for one of the two representations, so that students can position themselves regarding the issue at stake. Students must be encouraged to add such critical notes to their summary scheme.

In the next stage, this semantic pattern is transformed into a paragraph. The strategy must support students' formulation process to create intertextual connections, which requires: (1) transforming the condensed information from the summary scheme into full sentences and adding cohesive ties; (2) monitoring the text's comprehensibility for uninformed readers, and (3) weaving in the self-generated critical notes.

Integration activities are best taught via demonstration and subsequent practice (Casado- Ledesma et al., 2021; Mateos et al., 2020). In other studies on source based and synthesis text writing, observation of models proved to be effective for syntheses' integration quality (Buyuktas Kara et al., 2018; Linderholm et al., 2014; Mateos et al., 2008; Mateos et al., 2018; Raedts et al., 2007). Explicit instruction in the form of modeling has the advantage that it is less cognitively demanding than performing the task itself: it shows the strategy in action, including the metacognitive self-talk.

Based on the findings from the literature, design Principle 3 is: If students are to acquire a typical cognitively demanding strategy to integrate notes and convert them into comprehensible text, they should

- observe a competent model demonstrating the semantic and formulation steps of the strategy, applied to one of the paragraphs of the task at hand,
- apply the strategy by reprocessing the information in the summary scheme by: (1) rearranging arguments by shifting post-it notes in the scheme, (2) connecting arguments with an overarching term, and (3) weaving their critical notes into the scheme, and
- apply the strategy of how to write a synthesis paragraph, and
- practice writing the other two synthesis paragraphs.
- After practicing with the three components of the reading and writing strategy described above, and in line with Merrill (2002)'s fourth principle ('Apply'), students should be given the chance to practice completing a whole synthesis text.

3. Research questions and hypotheses

The aim of this study was to design and test a course to teach students how to produce a written synthesis in Science Orientation classes. Two questions guided this study, one about the quality of the design itself, and another about the efficacy of the course:

> *RQ 1:* To what extent does the designed synthesis writing course seem feasible for teaching synthesis writing, to elicit the intended learning activities, to be appreciated by students, and to result in experienced learning effects that align with the learning objectives?
>
> *RQ 2:* To what extent does the synthesis writing course improve synthesis text quality?

Based on similar studies by Van Ockenburg et al. (2019, 2021a, b), Mateos et al. (2018), and Martinez et al. (2015) we expect positive findings for both questions.

4. Method

4.1 Experimental design and participants

The set up for the study was quasi-experimental: intact classes, nested in schools, were assigned to conditions, an experimental treatment condition and a no treatment control condition to control for maturation effects. We administered the same pretest and posttest measurements in both conditions: writing a synthesis text.

The intervention in the experimental condition consisted of four actions in seven lessons; the three components discussed before, and a composing task (see Figure 1 and Table 1). Participants were students in a pre-university track from all 10th Grade classes of two schools, on average 15.6 years old, and generally from well-to-do families due to the geographical location of the participating schools in the Netherlands. The experimental group lessons were carried out in seven classes at the first author's school (n=129) that offers Science Orientation as a subject in Grade 10, and the control group lessons in four classes at a partner school nearby (n = 86), which did not offer Science Orientation as a subject.

A science-orientation teacher, the first author of this paper and teacher-researcher in this study, with seven years of experience teaching Science Orientation classes and fifteen years' teaching Biology, taught four classes in the experimental condition, while another Science Orientation teacher, with one year of experience with Science Orientation classes, and two years as a Philosophy teacher, taught the other three classes. The control condition sessions were scheduled in the regular classes for Dutch language and literature, a compulsory subject in Grade 10. Three L1-language teachers, each with extensive teaching experience, taught the regular L1-curriculum in these classes.

All participating students and their parents were informed that the results of the pretest and posttest, and the content of the learning workbooks would be processed anonymously for scientific research.

The research activities were scheduled as part of the regular curriculum in Science Orientation and L1 classes. Therefore, the University Ethics Board considered asking for parents' passive consent regarding their children's participation in the study suitable in this case. No one objected to their child's participation in the study.

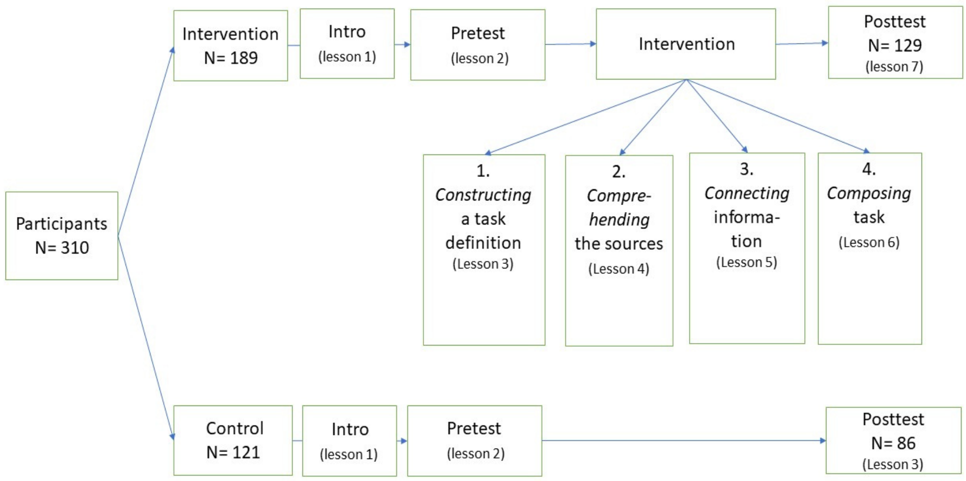


Figure 1. Design of the Synthesis Challenge writing course.

4.2 Procedures: Implementing the course & the Covid-19 challenge

The intervention was taught once a week for 75 minutes during the regular Science Orientation lessons. It started in the third week of January with an instruction lesson, followed by a week's holiday between lesson 2 and lesson 3 (constructing a task definition), and an unintended pause of 12 to 13 weeks caused by Covid-19 between lesson 4 (comprehending the sources) and lesson 5 (connecting information). Lesson 6 (the composing task) took place two weeks later via a *livestream*, in which teachers and assistants, at school, supervised the students who worked online from home. Lesson 7 (the posttest) took place a week later also via a *livestream.* Therefore, the whole course of seven lessons took a total of 23 weeks and lasted from January to June. For an overview of the lessons in the experimental and control group, see Table 1.

Table 1. Overview of course content in both conditions

Phase	Week	Lesson	Experimental Condition	Control Condition
0: Regular curriculum	Before the start of the writing course	.	Big History lessons on the origin of the universe and planet earth	Narrative analysis and literature history
1: Pretest	Week 4	1 (Introduction)	Introduction of the coming lesson series: aim of the research, giving a brief definition of a synthesis text and explanation of the continental drift issue, to prepare writing task 1 to equalize prior knowledge between conditions.	

	Week 6-7	2 (Pretest)	Pretest Synthesis task 1: Experience writing a first synthesis text. Issue: the continental drift.	
2: Intervention	Week 10	3	Strategy 1: Constructing a task definition by comparing and contrasting	Regular L1 Dutch curriculum, with writing, reading, literature
	Week 11	4	Strategy 2: Comprehending sources: by a retrieval assignment 'mind map evolution' followed by question-based summarizing	
	Covid-19 intermezzo Week 12-22		12/13 weeks of Big History lessons on the origin of life, animals, and humans.	
	Week 23	5	Strategy 3: Connecting source information: organizing content and adding a critical note	
	Week 25	6	Application: Composing task. The human evolution issue, based on preparations in lessons 4-5	
3: Posttest	Week 26	7 (Posttest)	Posttest Synthesis task 2: Writing a synthesis text. Issue: the definition of what language is.	

Prior to the intervention, the teacher-researcher discussed the lessons in detail with the colleague- teacher at the experimental school to ensure that they both strictly followed the planned scenario. When the regular lessons were converted into streamed lessons, due to Covid-19, the instruction was centralized: the teacher-researcher instructed all experimental classes, and lesson 7 (posttest) for the control group. All the lessons were discussed by the teachers and livestream supervisors involved, before and after they were carried out.

4.3 Instruction Materials

4.3.1 Lessons

The *Synthesis Challenge*, the experimental condition, consisted of seven lessons (see Figure 1). Table 1 provides information on the intervention and lessons and the influence of the Covid-19 pandemic. In phase 1, students were introduced to the Synthesis Challenge lessons and a brief definition of a synthesis text (lesson 1) and experienced writing a first synthesis text (lesson 2). In phase 2, students worked step by step towards composing the synthesis text in lesson 6. In phase 3, students wrote a synthesis text on a new and unprepared topic (posttest transfer task).

In the control group, students participated in an introductory lesson, in which they were taught about the topic that they had to write about, received information about the genre, and wrote their first synthesis text in lesson 2. Then they followed their regular language and literature curriculum for several weeks. Finally, they wrote a synthesis task on the same topic as the experimental group.

4.3.2 Teacher materials

Experimental condition. Teachers received a teacher workbook, with instructions per lesson and PowerPoint slideshows as instructional materials. Transcripts of the accompanying instructions were provided in teachers' workbooks as well. To facilitate observational learning, PowerPoint video clips were provided, see Appendix A, in OSF: Appendix A Taught strategies.docx

Control condition. Teachers received a teacher workbook, with instructions per lesson. For lesson 1, the control school's teachers received a PowerPoint video presentation recorded by the teacher-researcher, in which she explained the lessons' aims and provided students with background knowledge on the continental drift theory, the topic of the pretest task, their first synthesis writing assignment (lesson 2).

4.3.3. Student materials

Experimental condition. Students received paper workbooks, with instructions, assignments, and first recall questions about their experiences. The teacher collected all the workbooks at the end of each lesson, to ensure that students would have the relevant materials during each lesson and to avoid data loss. Source materials needed for writing synthesis texts were provided separately on paper, because earlier experiences had indicated that students focused better while reading from paper than online. Reading from paper enables readers to focus better (Baron et al., 2016; Clinton, 2019; Woody et al., 2010), especially if this is done under time pressure (Ackerman & Goldsmith, 2011; Delgado et al., 2018). Reading from paper also increases text comprehension (Gils et al., 2020; Mangen et al., 2013). In lesson 3 students received three anonymized student texts written in lesson 2 to work with during the task representation phase.

Control condition. The students in the control group received workbooks identical to the experimental condition, however, without the content of the four intervention lessons (lessons 3-6; see Figure 1).

4.4 Measurement Instruments

Various instruments were used to determine the quality of the design in terms of its feasibility, elicitation of intended learning activities, students' appreciation, experienced learning outcomes (RQ1), and text quality (RQ 2).

4.4.1 RQ 1: The design as implemented

Table 2 presents an overview of the instruments and methods used to answer RQ1.

Table 2. Instruments and methods for answering RQ 1*:* To what extent does the designed synthesis writing course meet the expectations?

Sub-question RQ 1	Instrument	Method
Feasibility for teaching synthesis writing	Time-on-task	Time-on-task measurements and procedures were based on Rietdijk et al. (2017). Eight students per class were observed for 1 minute each, one after the other. For each student, the observer noted whether the student had been on or off task during the previous two thirty second intervals. After observing the eighth student, the observer took a brief break and then carried out a new round of observations, six rounds in total Observations were carried out in six classes of the E-group and one class of the C-group in lessons 1 to 3, after which they stopped due to Covid-19 restrictions.
	Teachers' notes	Teachers in the E-condition had copies of detailed lesson plans, with a column to indicate to what extent the planned tasks had been carried out: completely, partly, or not at all. In addition, after each lesson they responded to five questions about classroom management, clarity of task instructions for teachers and students, the interest students showed, and practicality.
	Observations provided by teachers, observers and supervisors	Lesson 5 was carried out under different conditions due to the outbreak of Covid-19. The instruction was centralized, live streamed in multiple classrooms, while students were supervised by a supervisor in each room. Each supervisor completed an evaluation form afterwards to report whether students participated actively in class, whether they did what was asked of them, how they reacted to the instructions provided, and anything else they thought was worth reporting. During lessons 6 and 7, students worked online from home while writing the synthesis text they prepared in lessons 2-6, and the unprepared synthesis texts in Exam.net, a software tool especially developed for online exams. Students could read the sources and write their texts in a Word document, while each supervisor followed their progress and ensured texts were submitted and saved at the same time, in the same way, in groups of twenty students via Zoom.

Elicitation of the intended learning activities	Workbook task analysis	Per lesson we selected 3-7 key tasks to analyze. We coded the quality of students' responses to tasks. To judge to what extent the intended effect of the course could be linked to the three components, the assignments made in the workbooks were analyzed and coded for quality (see codebook Appendix B).
Appreciation by students	Questionnaire	Students responded after each lesson to questions about clarity (instruction, materials) and difficulty (tasks).
Result of experienced learning effects that align with the learning objectives	First-recall question.	At the end of each lesson, students completed the following sentence: *"I learned from this lesson that... "*. Students' learning experiences were processed and coded anonymously (see Appendix B in OSF: Appendix B codebook synthesistexts.docx Interrater reliability was found to be sufficient (Cohen's Kappa = .76).

4.4.2 RQ 2: Text quality: To what extent does the synthesis writing course improve synthesis text quality?

Writing Tasks

Two assignments were constructed on two different topics for the pre- and posttest. Both tasks required students to discuss the scientific issue-at-hand in a thoughtful way, by integrating the path of reasoning from the sources and their own topic and world knowledge into a text. Each task consisted of only two sources, because earlier experiences indicated that synthesis tasks with two sources were quite challenging for this age group (see also Barzilai et al., 2018).

Table 3. Features of the two writing tests

Measurement	Issue	Sources	Word count
Pretest	Continental drift question: How and why are theories generally accepted after a period of disbelief?	Text 1: Alfred Wegener and the continental drift hypothesis.	199
		Text 2: Criticism of the continental drift hypothesis.	305
		Total	504
Posttest	To what extent does language make us different from other animals?	Text 1: The case of Alex the parrot.	328
		Text 2: Human language differs from animal language.	258
		Total	586

The task was to write a synthesis text of 400 words maximum in 40 minutes. The test conditions in both conditions were kept as similar as possible and in alignment with regular classroom practices. Texts were written on digital devices which students were used to.

The assignments and sources were taken from the *Big History* website (https://www.oerproject.com/Big-History) and were translated into Dutch by the first author. Table 3 provides an overview of the assignments, topics and sources used in writing assignments in the present study. Each set of sources discussed a popular scientific issue and included a text with arguments for and a text with arguments against the issue in question. Opposing views in the issue facilitate comparison and integration (Barzilai et al., 2018; Wiley et al., 2018). The Composing Task (Lesson 6) in the experimental condition, had similar features.

Rating instrument

We based the text quality evaluation form on earlier synthesis writing studies (Cuevas et al., 2016; Gonzáles-Lamas et al., 2016; Martinez et al., 2015; Mateos et al., 2008; Vandermeulen et al., 2020b), and improved it after feedback obtained during earlier iterations of the instructional design (see Alkema, 2022). The form covered five text quality criteria: 1) Information (argument, complete and correct), 2) Integration (relations between arguments), 3) Critical note, 4) Coherence and structure, and 5) Language and references (see Appendix C IN OSF: Appendix C Assesment rubric synthesis texts.docx. The form was accompanied by a text scale, made up of five annotated texts ranked from relatively poor (0 points) to high quality (100 points), to assist the rating process. Each text's annotations described the extent to which it met each of five criteria (see Appendix C). We selected these texts and the annotations based on information from a group of experts.

Text selection for assessment

Texts from students who missed more than one intervention lesson, or whose texts exceeded the 400-word limit by more than 25% were excluded from assessment. Posttest texts written by students who had not participated in the pretest were also excluded. Overall, after applying these criteria 129 experimental group texts (60 texts excluded) and 86 control group texts (35 texts excluded) were assessed.

Assessors

As the assignments focused both on writing to acquire and synthesize new knowledge as well as learning to produce synthesis texts, we chose to have text quality assessed by a group of 13 Dutch language teachers as well as 14 subject teachers (including Science Orientation, Geography, Philosophy, History, and other subjects). We created this assessor mix also to ensure that the results would not depend on one or a few assessors of one specialty but would be validated by subject teachers and language teachers. Based on a pilot study (Alkema, 2022) we concluded that both groups were equally strict in their assessment of synthesis texts, and thus found it justified to draw from a combined pool of both groups. Groups of three independently working raters were formed, each team consisted of a mix of language and subject teachers. We had 27 assessors at our disposal; therefore, we composed nine mixed juries with three assessors who rated overlapping portions of texts (Van den Bergh & Eiting, 1989).

The assessors received detailed instructions. They had to read the assignment, then study the rubric, the text scale and a sample text assessed by the researcher. Subsequently, they rated students' texts on a scale of 0-100, and provided a score for overall text quality. Raters were blind to conditions and measurement moments. Pre- and posttest texts were presented in a balanced order to assessors. The agreement between assessors was sufficient (Cronbach's alpha: pretest: .76, posttest: 71).

4.5 Data analysis

RQ 1: Design quality

Descriptive scores (means, percentages) were calculated for time-on-task observations (feasibility), workbook analysis (elicitation of learning activities), students' appreciation, and experienced learning effects.

RQ 2: Text quality

We checked whether the quality of students' texts in the experimental and control conditions did not differ at pretest. On the pretest, we observed no differences between the two conditions in a multivariate analysis with the five text quality criteria and global text quality (Pillai's trace = .046, F (6.21) = 1.67, p = .13). Upon closer inspection, none of the univariate analyses showed a significant difference. Thus, both conditions' Text Quality did not differ at pretest.

Because data in this study were nested within students, and students within classes, the residual variance in traditional analyses of variance was overestimated. Therefore, we used mixed model analysis, with students nested in classes as random factor, and compared the fit of a series of nested models, next to the general model (Intercept only): Model 0 (plus random factor subjects within class); Model 1 (plus random factor class); Model 2 (plus factor Measurement Occasion); Model 3 (plus Factor Condition); and Model 4 (plus the interaction between Measurement Occasion and Condition). We ran these analyses for each of the five text quality criteria and the global score.

5. RESULTS

5.1 RQ 1: Design quality

Tables with the results for various aspects are provided in Appendix D in this article (Workbook analysis (Table D2), questionnaire answers (Table D3) and first recall answers (Table D4).

Teachers signaled that the design was doable, even though component 1, Constructing a task definition, required some extra teacher guidance. But precisely because students first had to struggle with the rubric, they welcomed their teacher's help, and most students ended up with a complete definition of what a synthesis text is. Supervisors and observers present during the livestream and online lessons were positive about the lessons. The observed lessons' time-on-task was high (Mean: 90%; see Table D1).

Furthermore, the design elicited the intended learning activities, described in the design principles, except for the learning activity *topic knowledge retrieval*. This activity needs to be better scaffolded, for example by using a fill-in scheme. The relatively high percentage of misconceptions in component 3 (see Appendix D2) may be the result of this lack.

Students seemed to appreciate the design, defined as whether the students found the assignment useful, clear, or easy to carry out; most of their responses about the three components were neutral to very positive about the learning activities (see Appendix D3).

Finally, the design seems to generate learning outcomes in students that align with the learning objectives of each specific lesson (3-5) for the most part, as indicated by students' answers to the *'I learned from this lesson...'* question (see Appendix D4).

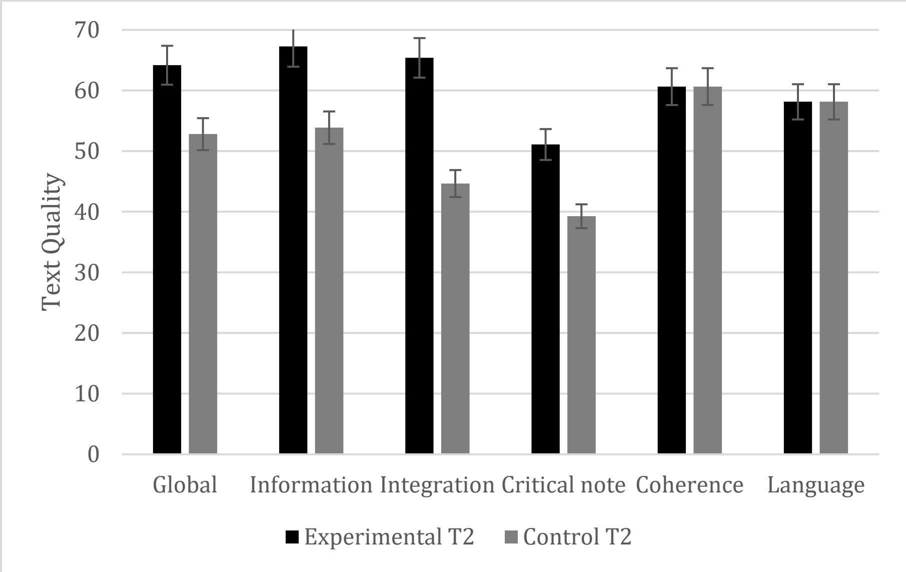
5.2 RQ 2: Text quality

Table 4 presents the estimated mean scores (and standard measurement errors) for both measurement moments and both conditions, based on the most appropriate model (for model comparisons, see Appendix E, this article). For four text quality variables, the model with the interaction between condition and measurement moment appeared the best fit. In each of these four cases, the effect of measurement moment in the experimental condition is larger: students' texts in the experimental group were better in terms of selected information (d = .52), integration (d = .73), and critical note (d = .40), and score higher on global quality (d = .44). For coherence and structure, an effect of measurement moment (model 2) was observed, while for language use quality no effects of measurement moment or condition were observed (model 1).

Table 4. Means (and standard measurements errors) for both conditions and measurement moments, estimated under the most appropriate model (see Appendix E for model comparisons).

		Experimental condition		Control condition	
		T1	T2	T1	T2
Variable	Best Model	M (Se)s	M (Se)	M (Se)	M (Se)
Global	4	53.48 (1.63)	64.19 (1.63)	50.62 (2.04)	52.81 (2.04)
Information	4	54.09 (1.64)	67.29 (1.64)	51.27 (2.04)	53.88 (2.04)
Integration	4	49.99 (1.80)	65.40 (1.80)	45.45 (2.24)	44.65 (2.24)
Critical note	4	33.35 (1.90)	51.10 (1.90)	35.44 (2.31)	39.27 (2.36)
Coherence and Structure	2	53.18 (1.56)	60.65 (1.56)	53.18 (1.56)	60.65 (1.56)
Language and References	0	58.15 (1.05)	58.15 (1.05)	58.15 (1.05)	58.15 (1.05)

Figure 2 shows the differences between text quality scores between conditions on both measurement occasions. For four variables the E-condition scored much better at posttest than the C-condition, while at pretest the differences between conditions were not statistically significant. The 'critical note' initially scored lowest at pre-test, but the E-condition made a big leap forward for this aspect, while the C-condition did not.



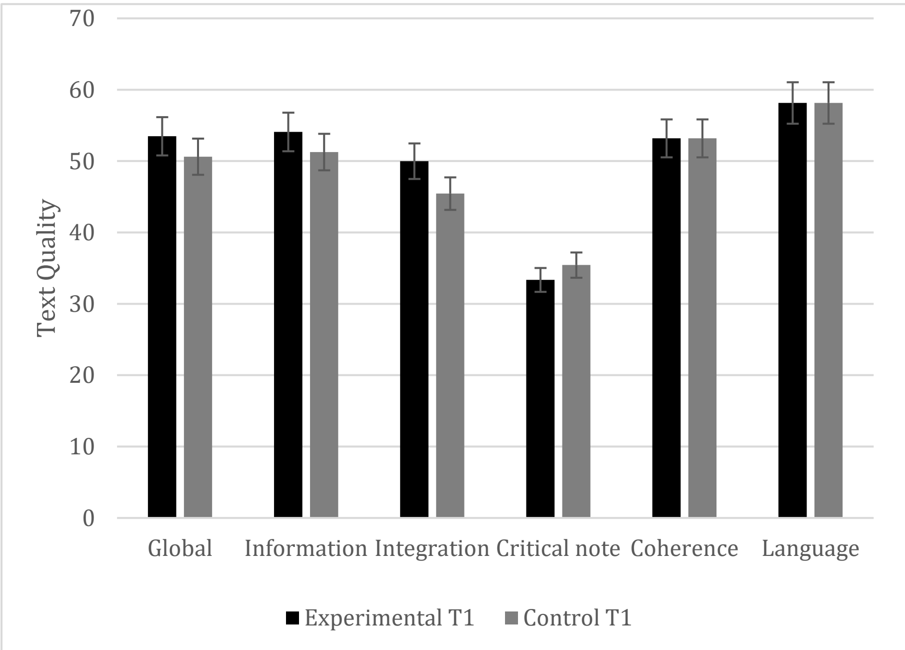


Figure 2. Quality scores for Global Text Quality and Five Sub-dimensions for Two Measurement Occasions (top panel: pretest; bottom panel: posttest) and Two Conditions (Means and errors bars, 95% confidence level).

Overall, the results indicate that texts in the E-condition scored significantly higher in terms of text quality at posttest after participating in the synthesis writing course than in the C-condition. Furthermore, participating in the course appears to have had an effect on the three text quality variables that were explicitly related to the three components (constructing the task definition, comprehending the sources, and connecting information) in the experimental course: information, integration, and critical note. For the variables coherence and structure and language and references, which only received some attention in the first component (constructing the task definition) of the course, no condition effect was found.

6. Discussion

This study tested the quality and effects of an instructional design to teach students to synthesize information on scientific issues, implemented in a Science Orientation 10th grade course. The course instructed students how to synthesize using several strategies, how to read sources and relate the information they contain for optimal intertextual comprehension, and how to produce a short, comprehensible synthesis text. The course had a positive effect on selecting information, integration of information and interweaving one's own substantive knowledge or critical note. The design of each component is based on available strategies, and is therefore a theoretically validated design, which proved feasible and effective. Students turned out to be positive about the lesson series.

The two conditions did not differ at posttest on two text quality criteria: coherence and structure, and language and references. This was expected, as these elements were only mentioned as part of the task representation component, and not further instructed. Improving synthesis texts' quality for these aspects should be a focus of the next round of instruction.

We can conclude that the design, based on the design principles, is theoretically valid. We chose to combine the components (Figure 1) into a complete learning unit, based on Graham and Perin (2007), who found that the most effective intervention consists of a combination of "all kinds of learning and instruction activities, including observation and modeling" (Graham & Perin, 2007, as cited by Rijlaarsdam et al., 2008, p.74) As far as we know, the combination of the four components for the successful writing of synthesis texts as is the case in our combined course (see Figure 1) has not previously been investigated.

The relevance of our combination approach lies mainly in the fact that the reading and integration task (components 1 to 3) is separated from the final writing task (component 4). The selection, connection and addition has taken place before writing the entire synthesis text, so that writing a synthesis text on a difficult issue becomes easier: everything is well thought out and noted in own words in a scheme, ready to be picked out to be used in the synthesis text. The design can be especially useful when students in higher grades are required to write synthesis texts for the first time. It is therefore recommended to practice synthesizing the selected main ideas from sources with the help of the instructional design from 10th grade onwards. Once it is clear what the intention is, component 1 could be omitted when writing subsequent synthesis texts. If the task involves difficult sources, the summary scheme can remain useful for implementing components 2 and 3. Especially if many sources must be read for students' final thesis, over a longer period of time, it is useful that students learn to summarize the sources per subtopic in a spreadsheet or scheme. The summary scheme then not only offers the possibility to organize and integrate but is also useful as a reference and documentation scheme, so that the consulted sources are not lost. In addition, the scheme gives teachers the opportunity to follow the developing literature overview, so that they can provide advice in time. It is then useful if students have practiced working with summary schemes in Grade 10 and have become confident using it.

If our instructional model now works for Grade 10 in SO-classes, and students must synthesize in other subjects, to what extent is our model more generally useful? Is it mainly a domain-specific instruction model, or can it also be considered a generic one? We tend to think that other school subjects could use our instructional design. The five evaluation criteria for text quality may guide the adaption process, as each component focuses on one to five quality criteria (see Table 5).

Table 5. Generic model for writing texts based on sources

Learning content	Approach	Text Quality criteria				
1. Learning strategy: Construct a task definition	Inductive: from experience to insight in possible outcomes and processes	1	2	3	4	5
2. Task strategy: Read-stop-think-note	Strategy instruction: observation and guided practice	1				
3. Task strategy: A: Critical semantic reprocessing of notes			2			
B: Integration of source information and own critical note in textual format				3		
4. Apply: Composing whole text	Practice	1	2	3	4	5

Note: Text quality criteria: 1. Information presented: complete and correct; 2. Integration: Information properly related; 3. Critical note: fact-check of the sources; 4. Coherence and structure; 5. Language and source reference.

We tested the model in Science Orientation, an interdisciplinary subject, which examines issues from the natural, social, and formal sciences. This school subject presents issues in the fields of language, history, nature, etc., and asks students to analyze sources, to discuss misconceptions and meta-concepts in these specific contexts, and to seek good argumentation for taking a position regarding the issue at stake. Such an interdisciplinary school subject is intended to connect different school subjects, and therefore, one may expect that an effective instructional model for synthesis writing might be adaptable to other school subjects, as most of them require reasoning and a way of critical thinking (Mayer, 1996). Generally, in all subjects, there are forms of cause-effect reasoning albeit with domain-specific concepts and type of relations which must be represented correctly, in an integrated text, with a critical stance, in a coherent and well-structured text. So, we assume that the structural elements of a synthesis text in different school subjects and in instructional models may largely overlap.

However, if we zoom in on what exactly happens in subject specific processes, then the processes for generating source-based texts seem to vary, as a result of different conceptions of the function of source-based writing. And when task representations about the product and the process differ, the synthesis strategy to be taught must be adapted to the school subject. Such differences in processes are demonstrated in Holdinga et al. (2021), in which students who took History and Philosophy classes wrote source-based texts in both subjects. Different reading-writing processes were observed, with different relations to text quality and quality of learned content. For History, these students separated the source analysis phase from the writing phase while when they wrote for Philosophy, the writing process was dominant and influenced their thought quality, in such a way that "…philosophical thinking and writing are intertwined" (Holdinga, et al., 2021, p. 577). For the instructional model this implies that the Task strategy component is one of the critical components to be included, with a different outcome for each school subject. Such observations may also lead to somewhat different strategies for analyzing source information (component 2 in our model) and connecting information (component 3); the content of these components will thus be guided by the specific subjects. On an abstract level, combining information from different sources by overarching the elements will be the same, but the specific semantics (concepts) and language might differ. In most subjects in upper secondary school, a critical approach to the information at hand will be required, but subjects will teach different schemata of cause-and-effect phenomena, and different types of relations and argumentation schemes.

Overall, we think that the instructional model we tested represents components and approaches (observation>practice) that can be applied in other school subjects, but with adapted strategies for components 2 and 3.

One important aspect of source-based writing activities, with which many models of multiple source literacy start (Rouet & Britt, 2011), is not included in our course. In regular Science Orientation classes, the reading and writing tasks start with a proper source analysis. We did not include this activity in the course on purpose, to enable students to focus on the cognitive and textual skills required to analyze, organize, and critically view reliable source information. Therefore, we presented only reliable sources in this course.

Our views cannot be separated from the limitations that any study has. First, one might wonder if the Covid-19-break was necessary at all. In retrospect, we could have started with the online lessons immediately after the lockdown started. However, as this was the first national lockdown, we were experimenting with new software such as Zoom etc. to enable us to continue our lessons online. This was new, for both students and teachers. So, it took some time to get used to the new restricted situation, develop new materials for streamed and online lessons, convince the school management of the importance of the study, and prepare students, parents, teachers, invigilators and others for the study's continuation.

Another factor to address is the difference in school subjects between the experimental (Science Orientation) and control groups (L1), which could also be an alternative explanation for the effects. However, both conditions did not perform differently at pretest. In addition, we controlled for topic knowledge at pre- and posttest. Additionally, writing a synthesis text was a new experience for students in both conditions at pretest. In both conditions, there was minimal instruction before the actual synthesis task, to avoid off task performances. In Science Orientation, at the first author's school, the curriculum is thematic: every week, a new topic is introduced and studied. These topics are largely independent. The pretest topic knowledge was provided in the E-condition via self-study, while in the C-condition, the researcher provided students with a PowerPoint presentation that included the necessary basic knowledge about the topic (the continental-drift-issue). For the posttest both groups received the explanation about the language issue in the same way, via a PowerPoint shown via YouTube. So, if the E-students score better on the instructed variables, it is not likely to be due to their prior knowledge or their Science Orientation classes, but more likely caused by the general strategies they practiced during the intervention. It is likely that the C-group would have achieved similar performance if they had also practiced our design strategies. Therefore, it seems likely that the generic model for writing integrated texts that we designed has ensured that domain-specific source content and students' own insights based on prior knowledge were well combined.

7. Conclusion

In conclusion, although there are certainly improvements that can be made on minor issues, such as ranking texts and knowledge retrieval, overall, this combined intervention, *The Synthesis Challenge*, was successful. The three key instructional components, instructional strategies for students to *construct a task definition* via a learning strategy based on comparing and contrasting texts and processes, *comprehend source information* via a read-stop-think-note strategy, and *connect source information critically* via a semantic-textual transformation strategy were all implemented as intended, despite the disruptions caused by the Covid-19 pandemic, and the quality of students' synthesis texts appeared to have improved.

Acknowledgements

This study was made possible thanks to a research grant awarded to Edith Alkema from the Dutch CVO-group (Christian Secondary Education Group) Driebergen-Zeist. Apart from the official corona measures, the school board in no way influenced the design and execution of this research. The study was also made possible thanks to the University of Amsterdam, for giving hospitality and facilities to carry out the research. Special thanks to Huub van den Bergh for advice on statistical analyses, to the colleagues from the first author school and the colleagues from the Universities of Amsterdam, Groningen and Antwerp for rating the texts, to all the invigilators who came to surveil during the livestream/Zoom lessons during the corona lockdown, and to the colleagues of the sister school who have made it possible for their students to act as a control group.

References

Alkema, E.G. (2022). *Schrijven met kennis van zaken. Een didactiek voor de zaakvakken/ Writing with knowledge. Didactics for subject classes*. (Doctoral dissertation, The University of Amsterdam, Amsterdam, The Netherlands). Retrieved from https://dare.uva.nl/search?identifier=d7627439-52b7-4914-b87f-d385f8748969

Ackerman, R., & Goldsmith, M. (2011). Metacognitive Regulation of Text Learning: On Screen Versus on Paper. *Journal of Experimental Psychology,17*(1), 18–32 https://doi.org/10.1037/a0022086

Barzilai, S., Zohar, A, R., & Mor-Hagani, S. (2018). Promoting Integration of Multiple Texts: A Review of Instructional Approaches and Practices. *Educational Psychology Review*, *30*(3), 973–999. https://doi.org/10.1007/s10648-018-9436-8

Bandura, A. (1986). Social foundations of thought and action: A social cognitive theory. Englewood Cliffs, NJ: Prentice Hall. https://doi.org/10.5465/amr.1987.4306538

Baron, N.S., Calixte, R.M., & Havewala, M. (2016). The persistence of print among university students: An exploratory study. *Telematics and Informatics, 34* (5) 590–604. https://doi.org/10.1016/j.tele.2016.11.008

Braaksma, M., Rijlaarsdam, G., & van den Bergh, H. (2002). Observational Learning and the Effects of Model–Observer Similarity, *Journal of Educational Psychology, 94*(2), 405–415. https://doi.org/10.1037/0022-0663.94.2.405

Braaksma, M. A. H., Rijlaarsdam, G., Van den Bergh, H., & Van Hout-Wolters, B. H. A. M. (2004). Observational learning and its effects on the orchestration of writing processes. *Cognition & Instruction*, *22*(1), 1–36. Http://dx.doi.org/10.1207/s1532690Xci2201_1

Braaksma, M., Rijlaarsdam, G., Van den Bergh, H., & Van Hout-Wolters, B. (2006). What observational learning in writing courses entails: A multiple case study. *L1-Educational Studies in Language and Literature*, *6*, 31-62. http://dx.doi.org/10.17239/L1ESLL-2006.06.01.05

Buyuktas Kara, M., Van Steendam, E., Rijlaarsdam, G., & Kuru, H. (2018). The effect of two modes of strategy instruction: Modeling vs. Presentational. *International Online Journal of Education and Teaching (IOJET), 5*(2),460-495. http://iojet.org/index.php/IOJET/article/view/313/247

Carr, E., & Ogle, D. (1987). K-W-L Plus: A Strategy for Comprehension and Summarization. *Journal of Reading*, *30(7)*, 626-631. https://www.jstor.org/stable/40031872

Casado-Ledesma, L., Cuevas, I., Van den Bergh, H., Rijlaarsdam, G., Mateos, M., Granado-Peinado, M., & Martín, E. (2021). Teaching argumentative synthesis writing through deliberative dialogues: instructional practices in secondary education. *Instructional Science*, *49*(4), 515-559. http://dx.doi.org/10.1007/s11251-021-09548-3

Clinton, V. (2019). Reading from paper compared to screens: A systematic review and meta- analysis. *Journal of Research in Reading*, *42*(2), 288–325. https://doi.org/10.1111/1467-9817.12269

Couzijn, M. & Rijlaarsdam, G. (2004). Learning to read and write argumentative text by observation of peer learners. Rijlaarsdam, G. (Series Ed.) & Rijlaarsdam, G., Van den Bergh, H., & Couzijn, M. (Vol. Eds.). Studies in writing. Vol. 14, *Effective learning and teaching of writing. A handbook of writing education* (2nd Edition, part 1, Studies in learning to write, 241-258). Boston: Kluwer Academic Publishers. http://dx.doi.org/10.1007/978-1-4020-2739-0_17

Cuevas, I., Mateos, M., Martin, E., Luna, M., Martin, A., Solari, M., González-Lamas, J., & Martinez, I. (2016). Collaborative writing of argumentative syntheses from multiple sources: The role of writing beliefs and strategies in addressing controversy. *Journal of Writing Research*, *8*(2), 205- 226. http://dx.doi.org/10.17239/jowr-2016.08.02.02

Delgado, P., Vargas, C., Ackerman, R., & Salmerón, L. (2018). Don't throw away your printed books: A meta-analysis on the effects of reading media on reading comprehension. *Educational Research Review, 25*, 23-38. https://doi.org.10.1016/j.edurev.2018.09.003 https://doi.org/10.1016/j.edurev.2018.09.003

Dewitz, P., Jones, J., & Leahy, S. (2009). Comprehension strategy instruction in core reading programs. *Reading research quarterly*, *44*(2), 102-126. http://dx.doi.org/10.1598/RRQ.44.2.1

Escorcia, D., Passerault, J., Ros, C., & Pylouster, J. (2017). Profiling writers: analysis of writing dynamics among college students. *Metacognition Learning, 12*(2), 233-273. https://doi.org/10.1007/s11409-016-9166-6

Galbraith, D., & Baaijen, V. M. (2018) The Work of Writing: Raiding the Inarticulate, *Educational Psychologist, 53*(4), 238-257. https://doi.org/10.1080/00461520.2018.1505515

Gils, J. van, Bakker, N., & Evers-Vermeul, J. (2020). De tablet, het geheime wapen voor leesbevordering? De invloed van tekstdrager op leesmotivatie, immersie en tekstbegrip [The tablet, the secret weapon for reading promotion? The influence of text carrier on reading motivation, immersion and text comprehension]. *Levende Talen Tijdschrift Jaargang, 21*(2), 3-11. https://lt-tijdschriften.nl/ojs/index.php/ltt/article/view/2056/1657 https://doi.org/10.5117/tvgesch2019.1.003.groot

González-Lamas, J., Cuevas, I., & Mateos, M. (2016) Arguing from sources: design and evaluation of a programme to improve written argumentation and its impact according to students' writing beliefs. *Journal for the Study of Education and Development*. *39(1)*, 49–83. https://doi.org/10.1080/02103702.2015.1111606

Granado-Peinado, M., Mateos, M., Martín, E., & Cuevas, I. (2019). Teaching to write collaborative argumentative syntheses in higher education. *Reading and Writing*, *32*(8), 2037-2058. https://doi.org/10.1007/s11145-019-09939-6

Hammann, L. A., & Stevens, R. J. (2003). Instructional approaches to improving students' writing of compare-contrast essays: An experimental study. *Journal of Literacy Research, 35*(2), 731- 756. http://dx.doi.org/10.1207/s15548430jlr3502_3

Hattie, J. (2009). Visible learning – A synthesis of over 800 meta-analyses relating to achievement. New York: Routledge. https://doi.org/10.4324/9780203887332

Hillocks, G. (1986), *Research on written composition: new directions for teaching*. Urbana, Illinois: ERIC Clearinghouse on Reading and Communication Skills.

Holdinga, l., Janssen, T., & Rijlaarsdam, G. (2021). The Relationship Between Students' Writing Process, Text Quality, and Thought Process Quality in 11th-

Grade History and Philosophy Assignments. *Written Communication*. *38*(4), 544-586. https://doi.org/10.1177/07410883211028853

Kiewra, K. A. (1989). A review of notetaking: The encoding- storage paradigm and beyond. *Educational Psychology Review*, *1*(2), 147–172. https://doi.org/10.1007/BF01326640

Kirkpatrick, L.C., & Klein, P.D. (2016). High-achieving high school students' strategies for writing from Internet-based sources of information. *Journal of Writing Research*, *8*(1), 1-46. https://doi.org/10.17239/jowr-2016.08.01.01

Linderholm, T., Therriault, D. J., and Kwon, H. (2014). Multiple science text processing: building comprehension skills for college student readers. *Reading Psychology, 35,* 332–356. https://doi.org/10.1080/02702711.2012.726696

Luna, M., Villalón, R., Mateos, M., & Martín, E. (2020). Improving university argumentative writing through online training. *Journal of Writing Research*, *12(1)*, 233-262. https://doi.org/10.17239/jowr-2020.12.01.08

Luo, L., & Kiewra, K. A. (2019). Soaring to successful synthesis writing: An investigation of SOAR strategies for college students writing from multiple sources. *Journal of Writing Research*, *11*(1)*,* 163-209. https://doi.org/10.17239/jowr-2019.11.01.06

Mangen, A., Walgermo, B.R., & Brønnick, K. (2013). Reading linear texts on paper versus computer screen: Effects on reading comprehension. *International Journal of Educational Research, 58*, 61-68. https://doi.org/10.1016/j.ijer.2012.12.002

Martinez, I., Mateos, M., Martin, 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

Mason, L.H., Reid, R., & Hagaman, J. (2012). *Building comprehension in adolescents: Powerful strategies for improving reading and writing in content areas.* Baltimore: Brooks Publishing.

Mateos, M., Martin, E., Villalon, 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

Mateos, M., Martín, E., Cuevas, I., Villalón, R., Martínez, I., & González-Lamas, J. (2018). Improving written argumentative synthesis by teaching the integration of conflicting information from multiple sources. *Cognition and Instruction*, *36*(2), 119–138. https://doi.org/10.1080/07370008.2018.1425300

Mateos, M., Rijlaarsdam, G., Martin, E., Cuevas, I., Van den Bergh, H., Solari, M. (2020). Learning paths in synthesis writing: Which learning path contributes most to which learning outcome? *Instructional Science* (2020). https://doi.org/10.1007/s11251-020-09508-3

Mayer, R. E. (1996). Learning strategies for making sense out of expository text: The SOI system of guiding three cognitive processes in knowledge construction. *Educational Psychology Review*, 8, 357–371. https://doi.org/10.1007/bf01463939

Meneses, A., Montenegro, M., Acevedo, D., Figueroa, J., & Hugo, E. (2023). Cross-disciplinary language changes in 4th graders as a predictor of the quality of written scientific explanation. *Journal of Writing Research, 15*(1), 575-602. https://doi.org/10.17239/jowr-2023.15.01.05

Merrill, M.D. (2002). First Principles of Instruction. *Educational Technology Research and Development, 50*(3), 43–59. https://doi.org/10.1007/BF02505024

Plakans, L. and Liao, J.-T. (2018). Task Representation. In The TESOL Encyclopedia of English Language Teaching (eds J.I. Liontas, T. International Association and M. DelliCarpini). https://doi.org/10.1002/9781118784235.eelt0532

Raedts, M., Rijlaarsdam, G., Van Waes, L., & Daems, F. (2007). Observational learning through video- based models: Impact on students' accuracy of self-efficacy beliefs, task knowledge and writing performances. In G. Rijlaarsdam, P. Boscolo, & S. Hidi (Vol. Eds.), *Studies in Writing, vol. 19, Writing and Motivation* (pp. 219–238). Oxford, United Kingdom: Elsevier. https://doi.org/10.1108/s1572-6304(2006)0000019014

Reynolds, G. A., & Perin, D. (2009). A comparison of text structure and self-regulated writing strategies for composing from sources by middle school students. *Reading Psychology*, *30*(3), 265–300. Https://doi.org/10.1080/02702710802411547

Rietdijk, S., Janssen, T., van Weijen, D., Van den Bergh, H., & Rijlaarsdam, G. (2017). Improving Writing in Primary Schools through a Comprehensive Writing Program. *Journal of Writing Research*, *9*(2), 173–225. https://doi.org/10.17239/jowr-2017.09.02.04

Rijlaarsdam, G.C.W., Janssen, T.M., Rietdijk, S., & Van Weijen, D. (2017). Reporting Design Principles for Effective Instruction of Writing: Interventions as Constructs. In Fidalgo, R., Harris, K.R. & Braaksma, M.A.H. (Eds), *Design Principles in Writing Instruction. Studies in Writing*, Volume 34. Brill Publishers: Leiden, the Netherlands. https://doi.org/10.1163/9789004270480_013

Rouet, J.F., & Britt, M. A. (2011). Relevance processes in multiple document comprehension. In M. T. McCrudden, J. P. Magliano, & G. Schraw (Eds.), *Text relevance and learning from text* (pp. 19– 52). Greenwich, CT: IAP. https://psycnet.apa.org/record/2011-20070-002

Salmerón, L., Strømsø, H. I., Kammerer, Y. A. K., Stadtler, M., & van den Broek, P. (2018). Comprehension processes in digital reading. In M. Barzill, J. Thomson, S. Schroeder, & P. van den Broek (Eds.), *Learning to read in a digital world* (pp. 91-120). John Benjamins Publishing Company https://doi.org/10.1075/swll.17

Solé, I., Miras, M., Castells, N., Sandra 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, N. (1991). Transforming texts: Constructive processes in reading and writing. *Written Communication,* 7(2)*,* 256–287. https://doi.org/10.1177/0741088390007002004

Stein, V., (1989). Elaboration: using what you know. *Reading-to-Write Report,* No. 6. The Center for the Study of Writing, Pittsburgh, Pennsylvania/Berkeley, California. https://eric.ed.gov/?id=ED306596

Surma, T., Vanhoyweghen, K., Sluijsmans, D., Camp, G., Muijs, D., & Kirschner, P. (2019). Wijze lessen, 12 bouwstenen voor effectieve didactiek. Ten Brink Uitgevers, Meppel.

Ter Beek, M., Spijkerboer, A. W., Brummer, L., & Opdenakker, M-C. (2018). Gemotiveerd, actief en zelfstandig lezen: Hoe een digitale leeromgeving zowel de leerling als de docent kan ondersteunen bij het begrijpend lezen van informatieve zaakvakteksten in het voortgezet onderwijs. Rijksuniversiteit Groningen. https://www.nro.nl/sites/nro/files/migrate/rapport-gazelle-nro-ppo-project-405-15-551.pdf

Törmälä, V., & Kulju, P. (2023). Work descriptions written by third-graders: An aspect of disciplinary literacy in primary craft education education. *Journal of Writing Research, 15*(1), 485-510. https://doi.org/10.17239/jowr-2023.15.01.02

Van den Bergh, H., & Eiting, M. H. (1989). A method of estimating rater reliability. *Journal of Educational Measurement*, *26*(1), 29–40. https://doi.org/10.1111/j.1745-3984.1989.tb00316.x

Van den Broek, P., & Helder, A. (2017). Cognitive processes in discourse comprehension: Passive processes, reader-initiated processes, and evolving mental representations. *Discourse Processes*, *54*(5-6), 360-372. https://doi.org/10.1080/0163853X.2017.1306677

Vandermeulen, N., Leijten, M., & Van Waes, L. (2020a). Reporting Writing Process Feedback in the Classroom: Using Keystroke Logging Data to Reflect on Writing Processes. *Journal of Writing Research*, *12*(1), 109-140. Https://doi.org/10.17239/jowr-2020.12.01.05

Vandermeulen, N., De Maeyer, S., Van Steendam, E., Lesterhuis, M., van den Bergh, H., & Rijlaarsdam, G. (2020b). 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), 187-236. https://doi.org/10.17239/jowr-2023.14.03.05

Vandermeulen, N., Van Steendam, E., van den Broek, B., & Rijlaarsdam, G. (2020c). In search of an effective source use pattern for writing argumentative and informative synthesis texts. *Reading and Writing*, *33*(2), 239–266. https://doi.org/10.1007/s11145-019-09958-3

Van Driel, J., Van Drie, J. & Van Boxtel, C. (2022). Writing about the significance of historical agents: the effects of reading and writing instruction. *Reading and Writing*, 1-27. https://doi.org/10.1007/s11145-022-10404-0

Van Gelder, T. (2005). Teaching Critical Thinking: Some Lessons From Cognitive Science. College Teaching, Winter 2005; 53,1, pp. 41 - 46. https://doi.org/10.3200/CTCH.53.1.41-48

Van Ockenburg, L., Van Weijen, D., & Rijlaarsdam, G. (2018). Syntheseteksten leren schrijven in het voortgezet onderwijs. Het verband tussen schrijfaanpak en voorkeur voor leeractiviteiten. [Learning to write synthesis texts in secondary education. The relationship between writing approach and preference for learning activities]. *Levende Talen Tijdschrift 19*(2), 3-9. https://lt-tijdschriften.nl/ojs/index.php/ltt/article/view/1795

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), 402-4028. https://doi.org/10.17239/jowr-2019.10.03.01

Van Ockenburg, L., Van Weijen, D., & Rijlaarsdam, G. (2021a). Choosing how to plan informative synthesis texts: effects of strategy-based interventions on overall text quality. *Reading and Writing*, 1-27. https://doi.org/10.1007/s11145-021-10226-6

Van Ockenburg, L., Van Weijen, D., & Rijlaarsdam, G. (2021b). 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, 1-33. https://doi.org/10.17239/l1esll-2021.21.01.06

Venneker, M. (2017). De kracht van de aantekening. [The power of notes]. *NVOX*, 5, 232-233. https://www.nvon.nl/nvox/de-kracht-van-de-aantekening

Weston-Sementelli, J. L., Allen, L. K., & McNamara, D. S. (2016). Comprehension and writing strategy training improves performance on content-specific source-based writing tasks. *International Journal of Artificial Intelligence in Education*, *28*(1), 106–137. https://doi.org/10.1007/s40593-016-0127-7

Wiley, J., Jaeger, A. J., & Griffin, T. D. (2018). Effects of task instructions on comprehension from multiple sources in history and science. In J. L. G. Braasch, I. Bråten, & M. T. McCrudden (Eds.) *Handbook of multiple source use* (pp. 341-361). New York, NY: Routledge. https://doi.org/10.4324/9781315627496-20

Woody, W., Daniel, D., & Baker, C. (2010). E-books or textbooks: Students prefer textbooks. *Computers & Education, 55*(3), 945–948. https://doi.org/10.1016/j.compedu.2010.04.005

Yeh C.-C. (2009). Student perceptions of an EFL undergraduate research writing project. *RELC Journal, 40*(3), 314-332. https://doi.org/10.1177/0033688209343863

Appendix A: OSF - Taught strategies.docx

Appendix B: OSF - Codebook synthesistexts.docx

Appendix C: OSF - Appendix C Assesment rubric synthesis texts.docx

Appendix D: Additional results for Time-on-task (D1), workbook analysis (Table D2), students' appreciation (Table D3) and responses to open learning questions (Table D4)

Table D1. Time-on-task

Date	Lesson	% On task
5-2-20	2	90
5-2-20	2	89
5-2-20	2	87,5
12-2-20	2	87,5
10-2-20	2	100
11-2-20	2	94,7
4-3-20	3	85

Table D2. Content workbook lessons 3 (component 1), 4 (component 2) and 5 (component 3): Percentage of assignments made.
The number of responses and students are provided in brackets in the Activities column.

Variable	Activities	Desired answer	Percentage of assignments created
1: Constructing a task definition	Reproduce the thinking steps of the thinking writer (146/189)	1) Selecting arguments from sources, 2) Relating the arguments, and 3) explaining with your own knowledge.	97.9% three correct answers
	Arrangement of three text examples, A, B and C (154/189)	Text A = a well-ordered summary; Text B= a perfect synthesis text, and text C = a poorly written synthesis text	56.5% think B is the strongest and C the weakest text, while C has integration and own knowledge, and text A does not.
	Reason text ordering A (152/189)	Text A has no integrations, clarification and critical notes.	It is most often mentioned that text A has selected the arguments correctly (10.5%). It is just as often

		mentioned that one misses the conclusion (10.5%). Text A is seen as the average text (59.4%). The integrations, clarifications and critical notes are >50% well estimated (58.6%, 66.1% and 75.7% respectively).
Reason text ordering B (151/189)	Text B has integrations, clarifications and critical notes	Students appreciate the selection of arguments, the order and the language aspect. Although the text has many integrations, that aspect is hardly mentioned. Text B is seen as the strongest (69.4%). The integrations, clarifications and critical notes are >50% well estimated (78.2%, 73.6% and 64.0% respectively).
Reason text ordering C (148/189)	Text C has also some integrations, clarifications, and critical notes (but because the text is poorly written as regards language, students have difficulties to recognize these qualities).	Although text C contains integrations and own notes, this is hardly experienced as positive. Students most often mention that the arguments and structure are unclearly represented. Text C is the weakest (64.4%). The integrations, clarifications and critical notes are >50% underestimated (75.3%, 68.8% and 64.9% respectively)
List of qualities a good synthesis text must cope with (152/189)	1.All arguments selected 2.All arguments integrated 3.There is a critical note 4.Cohesion and structure (Introduction, main part, end)	The selection of arguments, coherence & structure and citation are most often mentioned (17.1 %)

		5.Language and reference	
	List of qualities a good synthesis text must cope with, after a class conversation (150/189)	Ibidem	After a class discussion, integration and own note were most often mentioned.
2: Question-based' summarizing	Mind map evolution (177/189)	A body cell with a nucleus containing DNA and proteins, noted in the right place. DNA that has been partially copied, reproductive cells with a piece of single-stranded mutated DNA, a mutated gene in a fertilized egg cell etc.	87.6% not filled in correctly.
	The two theses filled in scheme (182/189)	1.men no longer evolve 2.men are still evolving	75.3% done correctly, 17.6% not done, 7.1% not done correctly
	Scheme with arguments on post-its (182/189)	Three arguments from source A (1.no natural selection anymore, 2. less mutations in sex cells, 3.no isolation anymore/decrease of genetic variation), and three arguments from source B (1. increase of genetic variation, 2. still natural selection, 3.no decrease in mutations)	91.2% done with >50% good.
3: Connecting and integrating source information and own knowledge	Post-its ordered (178/189)	A versus B: 1 versus 2 2 versus 3 3 versus 1	20.2% unordered, 6.2% half correct, 73.6% correctly ordered
	Critical note on post-it (178/189)	1 versus 2: I agree with B, there is still natural selection, because I learned in geography classes, that there are	Three, with 1x correct: 33.1%; Three, with 2x correct: 28.7%; Three with 3x correct: 15.2%. Misconceptions are for example for example,

	enough people who starve from hunger etc. 2 versus 3: I agree with B because there are still enough mutations in sex cells caused by radiations, chemicals etc. 3 versus 1: e.g. I don't understand B well, his arguments remind me of those from Lamarck etc.	that people become more and more similar through sexual reproduction (sexual reproduction provides more variation).
Synthesis text paragraphs written (176/189)	A paragraph in which argument 2 from source A is compared with argument 3 from source B, and A second paragraph in which argument 3 from source A is compared with argument 1 from source B.	Two done with integration and own note: 58.5% Two done with own note: 13.6% Two done with integration: 9.7% Two done without integration and own knowledge: 14.2%

Table D3. Students' appreciation of key activities of each instructional component. --/-= negative to very negative; +/- = neutral, and +/++ = positive to very positive response.

Students' response in percentage

	-/-	+/-	+/++
Component 1: Constructing a task definition			
Instructiveness assessing sample texts: Clear?	5	43,9	51,1
Rubric: Clear?	16.8	34,3	48,9
Video definition synthesis text: Clear?	15.2	44,2	40,6
Component 2: Comprehending sources by question-based summarizing			
Working with post-its: useful?	3.7	23.2	73,1
Video question-based summary: Clear?	9.4	56.2	34.4
Thinking about school knowledge while reading: Done?	13.1	38.8	48,1
Asking questions after each paragraph: Done?	7,5	26,3	66,2
Component 3: Connecting information: Observational learning			
Reordering with post-its: useful?	4,6	21,9	73,5
Videos reordering with post-its: useful?	5,7	27,7	66,7
Video about writing a synthesis paragraph: Clear?	4,6	24,9	70,5
Writing a first synthesis paragraph: Easy?	27,7	31,3	41,1

Table D4. Percentages of responses given to the open learning questions of three instructional components. For a detailed explanation of the code numbers, see Appendix B.
Numbers in bold refer to content that matched the instructional goal for that specific lesson.

Component		1 Constructing Task definition	2 Comparing sources: Question-based summary	3 Connecting information and critical note
Code	Categories of learning experiences?			
numbers 1		55.2	1.3	2.9
	Analyze sources			
3	Writing synthesis text	25.4	1.9	20.1
4	Selection of main ideas	5.2	19.4	1.1
5	Integration	3.7	1.3	42.5
7	Critical note	2.2	0.6	19.0
10	Question-based summary	0.7	62.5	1.7
	Total answers given.	134	160	174
	Total students	189	189	189
	Direction of the response:			
	Negative = -	7.5	4.4	5.7
	Neutral = ±	76.9	74.2	77.6
	Positive = +	15.7	21.4	16.7

Appendix E. Model comparisons from Mixed Model Analysis

Variable	Models	Comparison					
Global Quality		X^2	df	Models	X^2	df	P
Model 0	Intercept plus random factor (ID)	3542.02	3				
Model 1	Plus: Random factor Class	3524.18	4	0 vs 1	17.84	1	<.001
Model 2	Plus: Measurement Occasion (MO)	3486.501	5	1 vs 2	37.679	1	<.001
Model 3	Plus: Condition	3479.885	6	2 vs 3	6.616	1	.010
Model 4	Plus: Interaction MO*Condition	3466.031	7	3 vs 4	13.854	1	<.001
Information		X^2	df	Models	X^2	df	P
Model 0	Intercept plus random factor (ID)	3674.129	3				
Model 1	Plus: Random factor Class	3662.081	4	0 vs 1	12.048	1	.001
Model 2	Plus: Measurement Occasion (MO)	3622.486	5	1 vs 2	39.595	1	<.001
Model 3	Plus: Condition	3613.914	6	2 vs 3	8.572	1	.003
Model 4	Plus: Interaction MO*Condition	3598.816	7	3 vs 4	15.098	1	<.001
Integration		X^2	df	Models	X^2	df	P
Model 0	Intercept plus random factor (ID)	3756.069	3				
Model 1	Plus: Random factor Class	3725.102	4	0 vs 1	30.967	1	<.001
Model 2	Plus: Measurement Occasion (MO)	3694.597	5	1 vs 2	30.505	1	<.001
Model 3	Plus: Condition	3681.198	6	2 vs 3	13.399	1	<.001
Model 4	Plus: Interaction MO*Condition	3653.535	7	3 vs 4	27.663	1	<.001
Critical note		X^2	df	Models	X^2	df	P
Model 0	Intercept plus random factor (ID)	3731.269	3				
Model 1	Plus: Random factor Class	3722.549	4	0 vs 1	8.72	1	.003
Model 2	Plus: Measurement Occasion (MO)	3672.904	5	1 vs 2	49.645	1	<.001
Model 3	Plus: Condition	3669.824	6	2 vs 3	3.08	1	.079
Model 4	Plus: Interaction MO*Condition	3651.703	7	3 vs 4	18.121	1	<.001
Coherence & Structure		X^2	df	Models	X^2	df	P
Model 0	Intercept plus random factor (ID)	3595.935	3				
Model 1	Plus: Random factor Class	3585.63	4	0 vs 1	10.305	1	0.001
Model 2	Plus: Measurement Occasion (MO)	3555.485	5	1 vs 2	30.145	1	<.001
Model 3	Plus: Condition	3551.888	6	2 vs 3	3.597	1	.058
Model 4	Plus: Interaction MO*Condition	3551.133	7	3 vs 4	0.755	1	.385
Language & References	Y = C + [variances]	X^2	df	Models	X^2	df	P
Model 0	Intercept plus random factor (ID)	3461.206	3				
Model 1	Plus: Random factor Class	3459.024	4	0 vs 1	2.182	1	.140
Model 2	Plus: Measurement Occasion (MO)	3459.024	5	1 vs 2	0	1	1.000
Model 3	Plus: Condition	3457.466	6	2 vs 3	1.558	1	.212
Model 4	Plus: Interaction MO*Condition	3457.418	7	3 vs 4	0.048	1	.827