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

Received: 08.11.2017
Received in revised form: 15.02.2018
Accepted: 07.03.2018

THE EFFECT OF TWO MODES OF INSTRUCTION: MODELING VS. PRESENTATIONAL

Mujgan Buyuktas Kara
Research Institute of Child Development and Education. University of Amsterdam
mujganbuyuktaskara@gmail.com

Elke Van Steendam
KULeuven, Belgium
elke.vansteendam@arts.kuleuven.be

Gert Rijlaarsdam
University of Antwerp, Belgium
g.c.w.rijlaarsdam@uva.nl

Huseyin Kuru
Sapienza University of Rome, Italy
kuru.1515904@studenti.uniroma1.it

Department of Child Development and Education, PO Box 15776, 1001 NG Amsterdam

Mujgan Buyuktas Kara holds a BA and MA in English Language and Literature. She is currently working on her PhD at University of Amsterdam. Her research interests include writing instruction and writing processes in English as a foreign language (EFL).

Elke Van Steendam is an assistant professor at KU Leuven and University of Antwerp, Belgium. Her research interests include writing instruction and assessment, writing processes in a first (L1) and second/foreign language (L2/FL), collaborative writing and revision.

Gert Rijlaarsdam is a professor and an investigator of the Research Lab of Language, Literature & Arts Education at the University of Amsterdam, a member of the AnhubLab at University of Antwerp and a visiting professor at Umea University, Sweden.

Huseyin Kuru holds a BA in Statistics from Istanbul Commerce University and MBA in Advanced Economics from Sapienza University of Rome. He worked as a marketing manager in Rome and is currently working as a digital marketing specialist in Barcelona.

THE EFFECT OF TWO MODES OF INSTRUCTION: MODELING VS. PRESENTATIONAL

Mujgan Buyuktas Kara
University of Amsterdam, The Netherlands
mujganbuyuktaskara@gmail.com

Elke Van Steendam
KULeuven, Belgium
elke.vansteendam@arts.kuleuven.be

Gert Rijlaarsdam
University of Antwerp, Belgium
g.c.w.rijlaarsdam@uva.nl

Huseyin Kuru
Sapienza University of Rome, Italy
kuru.1515904@studenti.uniroma1.it

Abstract

We implemented an intervention of four lessons and tested the effects of two instructional modes as compared to the regular curriculum practices for completing a synthesis task at the preparatory program of a Turkish university. Participants were 48 upper-intermediate English as a Foreign Language learners (mean age = 18) assigned to three conditions. The presentational condition received direct strategy instruction supported by mnemonics; the modeling condition observed a video of a peer doing the task using the same strategies mnemonic. In the control condition, there was no explicit reference to strategies; rather, students inferred the necessary information about writing an effective synthesis text from the instruction and the lesson materials. We hypothesized that both of the experimental conditions would have a positive effect on students' synthesis text quality and writing processes and that modeling of explicit strategy use would have an effect over and above the other conditions. Results showed that students in the modeling condition improved their source use skills significantly more than students in the presentational condition, which was maintained in the delayed posttest four weeks later. No statistically significant condition effect was observed for content and authenticity of students' texts. The modeling condition also showed and reported a more process-oriented approach to writing.

Keywords: L2 writing, higher education, strategy instruction, modeling

1. Introduction

Integrating content in teaching foreign/second language (L2) writing skills is rare (Hinkel, 2015) and is definitely a challenging experience for instructors and curriculum developers in time-constrained language-teaching programs. English preparatory schools of Turkish universities are no exception. Students are almost always asked to write persuasive or argumentative essays using prior knowledge and personal experience rather than synthesis texts. However, the primary means of receiving information in an academic context is

reading (Grabe & Stoller, 2001). Students often write to demonstrate competence across the curriculum in response to content material (Hinkel, 2015). Hence, segregated writing tasks fail to represent academic expectations and are not as effective in improving linguistic skills or contributing to students' intellectual growth as much as integrated writing tasks (Leki & Carson, 1997).

Content-integrated writing should ideally promote the processes described in the "knowledge-transforming" model of writing. In this model, writers actively generate and evaluate content and organize information in a more sophisticated manner than in the "knowledge-telling" model of writing (Bereiter & Scardamalia, 1987). On the other hand, integrating content into writing definitely adds up to the already overwhelming cognitive load that is commonly associated with the writing activity (Mateos & Solé, 2009). Writing a synthesis text entails critical evaluation of information in accordance with genre-specific features and the audience. This multifaceted structure of synthesis tasks necessitates decomposing the several different aspects of the complex reading and writing task and addressing each one separately and explicitly to foster performance.

Flower and Hayes (1981) described writing as "a goal directed thinking process" (p. 366) and the role of the writer as that of a problem solver. Solving problems entails employing heuristic strategies (i.e., optional techniques to approach the task at hand). Having conscious access to a repertoire of heuristics (i.e., procedures for writing) can make the writing process considerably easier for the writer; and, fortunately, these heuristic strategies can be translated into teachable techniques (Flower & Hayes, 1977). This is especially important in an EFL (English as a Foreign Language) context, in which the majority of learning takes place in instructional settings.

1.1. Current Instructional Strategies in EFL Instruction in Turkey

In 1997 and in 2008, the Turkish Ministry of National Education adopted several policy changes in an effort to reform Teaching English as a Foreign Language (TEFL) practices in Turkey. The primary objective for teaching English at secondary schools was defined as development of learners' communicative capacity (Kirkgoz, 2005). In this vein, the "communicative" approach to TEFL was introduced at a national level for the first time (Kirkgoz, 2007). Curricular issues, such as selection of teaching materials, curriculum design, and the role of the teacher in the classroom were also defined in line with the main goals of Communicative Language Teaching (CLT) (Ozsevik, 2010).

CLT can be defined broadly as an approach to TEFL that emphasizes the participation of the learner in meaningful L2 interaction with a focus on functional and communicative aspects of the language. It has been associated with *implicit learning*, since it provides a naturalistic view of learning the language, that is "acquiring skills and knowledge without conscious awareness", similar to learning the mother tongue. Explicit learning, on the other hand, refers to "learner's conscious and deliberate attempt to master some material or solve a problem" (Dörnyei, 2009, p.3), which is associated with more traditional approaches that dominated the TEFL domain until the beginning of the 1970s (Dörnyei, 2009). Thus, with the introduction of CLT approach, there has been a shift towards implicit instruction of all four skills of the language (i.e., reading, listening, speaking and writing) in TEFL settings in Turkey. However, this created a discrepancy between TEFL practices in Turkey and evidence-based writing practices, particularly in L1.

1.2. Current Instructional Strategies in L1 Writing Instruction

Empirical research studies both with learning disabled students and normally performing adolescent students in L1 proved that interventions with explicit teaching of strategies for planning, revising and/or editing text are much more effective in improving writing skills than non-explicit instruction conditions (i.e., teaching text structures, the process writing approach, traditional instruction, practise writing and literature study) (Graham & Perin, 2007). Among strategy instruction regimens, Self-Regulated Strategy Development (SRSD) designed by Harris and Graham (1996) is distinct with its explicit teaching of writing strategies. It is an all-encompassing training program tailored to the needs of the students in terms of its recursive nature and the time allocation of different stages during which strategies are presented, discussed, modeled, memorized, scaffolded and practiced collaboratively and individually. SRSD has proved effective in a myriad of research studies (cf. Brunstein & Glaser, 2011; De La Paz, 2005), including two meta-analyses (Graham & Harris, 2003; Graham, 2006), as well as in teaching reading and writing hybrid tasks (Mason, Hickey Snyder, Sukhram, & Kedem, 2006; Martínez, Mateos, Martin & Rijlaarsdam, 2015). SRSD is especially potent because three very effective components in writing instruction are intertwined in its instruction: promoting self-regulation, direct instruction and modeling of the strategies.

1.2.1. Self-regulation

Deliberately employing strategies involves self-regulation functions, which include self-monitoring, self-instruction, goal-setting, and self-reinforcement (Sawyer, Graham, & Harris, 1992). Once learnt, self-regulation skills can be internalized and maintained for use in similar future circumstances (Schunk & Zimmerman, 1997). This autonomy helps learners control their learning processes (1998). Self-regulation is not necessarily a stand-alone component, since it can also be triggered through other components of strategy-focused instruction.

1.2.2. Direct instruction

Direct writing instruction is a deductive approach to learning. It includes the explanation of rules followed by controlled practice and delivery of explicit feedback (Manchón, 2009). Instruction is teacher-led and conveyed in a presentational mode, especially in the initial stages of the training program. It also draws upon the use of strategies through mnemonics to help create a representational system (Reber, 1976). This representational system enables the individual use of the strategies by gradually releasing the control from the instructor to the student, also known as scaffolding. Memorization helps fasten this process and is enabled primarily with the use of mnemonics, but also through graphic organizers, think-sheets and/or prompt cards (Baker, Gersten, & Graham, 2003). Mnemonics reduce the task requirements to a single chunk in the required order (Worthen, & Hunt, 2011) and facilitate retrieval by locating information in memory with associations (Malhotra, 1991). Hence, they may alleviate the cognitive load of learning to write complex writing tasks such as synthesis tasks. Mnemonics might be especially helpful if students are "cramming" for an exam (McPherson, 2000), which is a typical circumstance in time constrained EFL programs, as is also the case in this study.

1.2.3. Modeling

A key component of SRSD programs that has been exclusively studied is *modeling*, or its mirrored learning activity, that is, *observational learning* (hereby used interchangeably to refer to the same concept). Several factors have contributed to the growing body of research in this domain. First, observing modeled experiences revives imitative functions; and, arguably, triggers self-monitoring, self-judgment, and self-reaction, which are the three

pillars of self-regulation (Bandura, 1986). Vicarious experiences improve learners' self-efficacy beliefs (Bandura, 1986), which promotes positive learning behavior (Schunk, 1996; Zimmerman, 1995). When learners observe peers completing tasks successfully, they may form outcome expectations, which in turn motivate behavior towards achieving the desirable outcome (Zimmerman, 1977).

The simultaneous orchestration of several cognitive strategies is especially difficult for novice writers, as they have not yet acquired skills needed to manage the process of writing that strong writers have (Breetvelt, Van den Bergh & Rijlaarsdam, 1994; De la Paz & Graham, 2002). Observing modeled experiences helps students direct their limited cognitive resources to learning-to-write instead of producing a text and, thus, counteracts the challenge of the "dual agenda" of having to zero in on both (Rijlaarsdam & Couzijn, 2000). Students who observe their peers on task also adopt a more recursive approach to writing and delay the executive writing activities to the later stages of the writing process and engage in more metacognitive activities, such as goal-orientation, and analyzing in the initial stages of writing. Students in a non-observation condition, on the other hand, adopt a more linear approach to writing with transcribing during the initial stages and with formulating spread throughout the whole writing process. Thus, observation encourages a purposeful temporal organization of cognitive activities and this has a positive effect on the quality of final written products (cf. Braaksma, Rijlaarsdam, Van den Bergh, & Van Hout-Wolters, 2004; Groenendijk, Janssen, Rijlaarsdam, & Van den Bergh, 2008).

We have reason to assume that peer modeling of explicit strategy use may be beneficial in teaching L2 synthesis tasks (the type of task in the present study) to Turkish university students. Writing is an academic requirement that both low- and high-achieving Turkish students find difficult and view as something to "persevere through in order to pass certain exams" (Yavuz & Genç, 1998, as cited in Erkan & Saban, 2011). This notion stems from negative student attitudes and writing apprehension, as well as low self-efficacy in writing (Erkan & Saban, 2011). As observation improves self-efficacy, it can have an activating power towards positive learning behavior and result in writing success.

Synthesis writing is cognitively more demanding than most other writing tasks (cf. Mateos & Solé, 2009), as it brings together task requirements such as organizing, selecting and connecting (Spivey, 1997). This complexity becomes daunting when the task is in L2, as in this case language proficiency level also comes into play (Plakans, 2009). Arguably, as observation may activate the "learner" capacity in a learning-to-write activity, students can use more of their cognitive resources for metacognitive and procedural knowledge instead of focusing predominantly on the production process, (Rijlaarsdam, Braaksma, Couzijn, Janssen, Kieft, & Broekkamp, 2005). This may be an effective strategy for them to complete synthesis tasks, where the "procedures" (i.e., organizing, selecting and connecting) are more burdensome than in the other writing tasks.

In Turkey students are discouraged to critically question text information (Turkkollu, 1994; Clachar, 2000). However, writing a synthesis requires engagement with the text at a critical level (Mateos & Solé, 2009). For learning to write a synthesis text, students need to familiarize themselves with this new discourse, as well as to produce a new text. In such a complex writing task as synthesizing, students may benefit from observation as it has the potential to alleviate the "dual-agenda" (cf. Rijlaarsdam & Couzijn, 2000) of learning-to-write and producing a text at the same time. In this way, it may also stimulate Turkish students' engagement with the task at a critical level.

Although the three components of strategy-focused instruction (i.e., self-regulation, direct instruction and modeling) have been tested separately and in combination with each other in

several studies, only one study has compared the effects of direct instruction and modeling (cf. Fidalgo, Torrance, Rijlaarsdam, Van den Bergh, & Álvarez, 2015). However, the differences in the participant profile, the tasks and especially the design issue certain caveats in the comparability of the two studies (see Conclusion section).

Another issue to consider is that although we expect effects of peer modeling on synthesis writing, individual differences between learners also affect the results of observation. In Zimmerman and Kitsantas's study (2002), college students benefited more from observation of a coping model than a mastery model. Braaksma, Rijlaarsdam and Van den Bergh (2002), in their study with students in secondary education, showed that when the task is novel, struggling writers benefit more from observing struggling models; and stronger writers from observing a stronger model. This shows that in studies looking into the effects of observation, individual differences should also be taken into consideration, which will be controlled for in this study.

2. The Present Study

We set out to improve the synthesis writing performance in the EFL program of a private Turkish university. Therefore, we tested the effects of a strategy-focused instructional design based on the principles of observational learning by comparing three treatment groups. In a modeling condition, students observed their peers modeling the use of strategies for completing a synthesis task. In a presentational condition, students received direct strategy instruction without modeling. Thus, the distinguishing feature of the two strategy instruction conditions is the mode in which the strategy instruction was conveyed (i.e., through modeling mode in the modeling condition compared to a direct, presentational format in the presentational condition). In the control condition, instruction was not strategy-focused, so there was no explicit presentation or modeling of a strategy. Teaching in the control condition took place more on the implicit rather than explicit end of the instructional scale, in line with the CLT approach to TEFL. In the control condition, students had to work out the task requirements (i.e., their own heuristic strategies) from the given materials in accordance with the guidance provided during the training session.

The hypotheses of the study are:

Hypothesis 1: writing performance. Modeling of strategies results in qualitatively better synthesis texts compared to presentation of strategies (1A), while presentation of strategies results in qualitatively better synthesis texts compared to a control condition (1B).

Hypothesis 2: writing processes. Modeling of strategies leads to improved synthesis writing processes compared to presentation of strategies (i.e., more meta-cognitive activities in the initial stages and more executional activities in the later stages of writing) (2A) and presentation of strategies leads to improved synthesis writing processes compared to a control condition (2B).

We also investigate, whether the students' motivational orientation is a confounding variable in the analysis of the results and, for generalization purposes, whether our hypotheses apply to students with different learner characteristics in terms of initial levels of motivation and writing performance.

3. Method

3.1. Participants

Participants were 48 (54% male; mean age: 18) pre-faculty course students in Module 1 of the 14-week combined program at a private Turkish university. They were a homogeneous

group of students in their reading, listening, writing and speaking skills, measured at the end of module tests, which were prepared in line with IELTS international exam specifications by the test office of the institution. Students' L1 was mainly Turkish except for six international participants (evenly distributed over the three conditions) admitted to study their entire degree at the university (two Syrian, one Afghan, one Iraqi, one Moldavian and one Macedonian). Students did not differ in their motivational orientation prior to the study (see Table 8), but did differ at pretest for text quality (see Table 9), which was corrected for using the pretest scores as covariate in the final analysis. The training and the tests were part of the curriculum apart from a summary task assigned as one of the pretests and a synthesis task assigned as one of the posttests (i.e., the delayed posttest). We informed the participants about the study before the delayed posttest, which would replace their previous grade should they get a higher score in the delayed posttest. Participants could ask for the removal of their data after that time until the end of the Module. All of the participants agreed to take part in the study.

Newly enrolled students were not eligible for studying in the pre-faculty course, so all of the students in this combined 14-week program were so-called "repeat students" who had failed at least once at any level in the previous academic year. The reasons for failure were mostly failing to meet academic standards, such as completing assignments, following ethical principles in writing and research and/or attending lessons regularly. Because of the distinct student profile, in Module 1, the management chose instructors with experience in the pre-faculty course with the particular (14-week) group of students.

3.2. Design

The number of lessons at the pre-faculty course is 20 lessons per week. There are four lessons of 50 minutes every day. Every week instructors have to allocate a total of four lessons to teaching writing skills with the materials prepared by a selected panel of instructors prior to the start of the academic year. Through weeks 1 – 5, the focus of the writing lessons is on critical thinking and argumentation skills, library skills, paraphrasing, summarizing, referencing and citing sources in APA style and writing an argumentative essay. Through weeks 6 – 7, the focus is on writing a synthesis text; and through weeks 7 – 14, students write an argumentative research paper through a process writing approach.

The study was conducted in nine sessions of 50 minutes each. Four sessions were reserved for training in Week 6 (Sessions 1-4) and five sessions for pre- and posttest administration, distributed over weeks 4 and 13: One session to collect baseline data about motivational orientation and text quality in the form of summary writing of a single source (Pre-session), three additional sessions for posttest 1 (Session 5), writing log training (Session 6) and delayed posttest administered with the process registration measure (Session 7). Motivation questionnaire and learner report data were collected in a Post-session. See Table 1 for the distribution of the sessions across weeks and pre- and posttests of the study.

Table 1. *Distribution of the sessions across weeks and pre and posttests of the study*

	Weeks		Activities		Text Genre
Pretests	Pre-session	4	MSLQ		ARG & SSS
	Sessions 1-4	6	Training		
	Session 5	7			SYN
Posttests	Session 6 Session 7	11 11	Log Training	Writing Logs	+SYN
	Post-session	13	MSLQ +	Learner Reports	

ARG: Argumentative essay; SSS: Summary of a single source; SYN: Synthesis task

We compared the effects of modeling and direct instruction of the use of strategies supported by mnemonics as opposed to a control condition on participants' synthesis writing performance and writing processes in an experimental pretest-posttest design. Participants were randomly assigned to three classes. These classes were assigned to three conditions, with the researcher's class appointed to the modeling condition because the majority of instruction in this condition is conveyed through the peer videos, so her instructional contact with the students would be minimal. The two other classes were randomly assigned to either the presentational or control conditions. All three instructors had around six years of experience in teaching EFL. Two of the instructors were female native speakers of the Turkish language, with English language and literature degrees and both PhD candidates. The other instructor was male, native speaker of English of Canadian origin, with an MA in ELT.

3.3. Materials

The focus of this study is the effect of different approaches to teaching a strategy. Two different strategy instruction conditions are compared to a control group. In both experimental conditions a multicomponent strategy to write a synthesis text was taught, but via a different instructional format. The materials were the same except for the ones immediately related to each instructional format, that is, peer videos, and the two materials for the instruction of the strategies: a slide on the introductory PowerPoint presentation (PPT) with the strategies mnemonic (i.e., TRAMPOLINE) and the accompanying handout.

3.3.1. TRAMPOLINE strategies

We adapted TRAP IDEAS reading and writing strategies for summarizing (Mason, Reid & Hagaman, 2012) into TRAMPOLINE strategies to write a synthesis text (see Table 2 for the TRAP IDEAS and Table 3 for the TRAMPOLINE strategies). The researcher and a colleague piloted the TRAP IDEAS strategies simultaneously in two pre-faculty level classes in the last module of the academic year preceding the actual experiment. The two trainers then liaised for the adaptations necessitated by the differences in the tasks (summary vs. synthesis), the contexts (L1 vs. L2) and the academic writing conventions for synthesis writing, such as using reporting verbs and APA principles.

Table 2. *TRAP IDEAS reading and writing strategies for summarizing*

Think before reading **R**ead the paragraph **A**sk: "What is the paragraph mostly about?" **P**araphrase the important information **I**dentify important details to support the main idea **D**elete trivial details **E**liminate redundant details **A**dd a term for a list of words or concepts **S**ummarize.

Table 3. *TRAMPOLINE strategies for writing a synthesis text*

Think – 3 steps a. Before reading the first extract: Think about the purpose why you are given different extracts on the same topic? After reading the second and the third extracts: What is the relationship of this extract to the previous one? b. What do you expect to learn from the extract? c. What do you already know about the general topic/the focus of the extracts? **R**ead the extracts **A**sk - What is the main idea? **M**ark the important details **P**araphrase the main idea and the important details (Repeat TRAMP for each extract and OLINE for the whole summary) **O**rganize the paraphrased ideas BY **L**inking the ideas with appropriate linkers **I**ncluding APA **N**esting reporting verbs **E**dit your summary

3.3.2. Videos

We shot three separate videos using the Camtasia Screencast Program (Techsmith, 2016) for the three substrategies: TRAM, P (Paraphrasing) and OLINE. The duration of the videos was 9, 17 and 12 minutes, respectively. All of the videos featured the same model: a female freshman student who had studied the preparatory program in the previous year. The videos were controlled think-aloud protocols written and performed in the English language. We prepared the framework of the script based on a more or less ideal student performance, that is, using the TRAMPOLINE strategies during task execution; but with occasional instances of the most common student mistakes in writing a synthesis text based on an error analysis we did with the instructors teaching in the pre-faculty course. The model refrained from adopting a prescriptive tone. She was asked to mimic an actual account of completing the writing task with the help of the strategies. To ensure authenticity, the model did not follow the script very strictly.

3.4. Conditions

Both conditions are based on the stages of the SRSD program, with memorization (through the use of mnemonics) present in both of the experimental conditions. The distinctive feature in the two strategy instruction conditions is the mode in which the instruction is conveyed: modeling versus a verbal presentational format manipulated in the experimental conditions. All other features appearing in the stages of SRSD are applied to the lesson plans at a micro level, rather than spread over a long period of time, as is the case in SRSD programs. Thus, the focus of the study is not SRSD as the time allocated for teaching synthesizing was predetermined by the administration to be four lessons, which is rather short for an SRSD program to be implemented. Table 4 provides an overview of the distinctive characteristics of the instructional conditions and Table 5 of the training session for the experimental conditions.

Table 4. *Characteristics of instructional conditions*

Component	Modeling	Presentational	Control
Presentational Mode	-	+	-
Peer Modeling	+	-	-
Strategies (mnemonic) for synthesizing	+	+	-
Direct Instruction	+	+	-
Collaborative practice	+	+	+
Individual practice	+	+	+

3.4.1. Experimental condition 1: Modeling condition

In Session 1, students discussed a controversial topic aimed at creating a meaningful context for introducing the task and the strategies. Using a Power Point Presentation (PPT.) and a complementary worksheet, the instructor showed participants two extracts from different articles about the topic and then a weak and a strong sample of student syntheses of these two extracts. The comparison of the two syntheses enabled generating common knowledge for completing the task, which to some extent corresponded with some of the strategies in the TRAMPOLINE mnemonic. Next, the instructor introduced the TRAMPOLINE strategies mnemonic and gave each student a TRAMPOLINE handout detailing the steps in the strategy for self-reference. The instructor showed examples of each strategy step by referring to the sample summaries and checked comprehension through a one-item exercise for each strategy. Finally, the instructor checked memorization of the mnemonic with a whole class drill. In session 2, students received a synthesis task with extracts from three different articles, observed the model in a video completing the TRAM strategies for extract 1 while thinking-aloud. Subsequently, participants emulated the strategies individually (finding the main idea and the important details) for extracts 2 and 3. Then, the participants watched the model paraphrasing (P of TRAMPOLINE) the main idea and the important details in extract 1. In this video, the sub-skills of effective paraphrasing were shown. In session 3, in groups of four, participants paraphrased the main ideas and the important details of extracts 2 and 3 (subsequent collaborative emulation). The instructor supervised the activity, provided scaffolding and showed possible responses on the board. Finally, students observed the video model showing the OLINE strategies for organizing the ideas by using linkers, including APA, nesting reporting verbs and editing text. In Session 4, participants completed a new synthesis task individually with minimal support for practicing purposes. The videos were not available to the students after the screening, but they were free to refer to the TRAMPOLINE mnemonic handout and could ask for minor assistance from the trainer.

3.4.2. Experimental condition 2: Presentational condition

The main difference between the two experimental conditions is that the observation tasks in the modeling condition are replaced with the teacher presentation of the TRAMPOLINE strategies. The instructor taught the strategies in, what can be defined as, presentational mode (Hillocks, 1984), with occasional teacher–led whole class question-and-answer episodes. The content of the first and the last sessions and the sequence of learning content in the 2nd and the 3rd sessions, was the same in both conditions: TRAM and Paraphrasing strategies in the 2nd session and Paraphrasing practice and the OLINE strategies in the 3rd session. In Session 2, the instructor gave the synthesizing task to participants, and instead of showing the video she presented the TRAM strategies through extract 1 via a teacher-led question-and-answer session and subsequent individual practice of the strategies by the participants on extracts 2 and 3. This was followed by presentation of P by the instructor. In Session 3, the participants practiced the strategies collaboratively and the instructor supervised, provided scaffolding and showed sample paraphrased sentences to the students. Finally, the instructor presented the OLINE strategies. Session 4 was the same as in the modeling condition.

3.4.3. Control condition

This is the regular curriculum practice of the institution. In this condition we adapted and used a lesson plan previously prepared by an instructor and observed and approved by the administration as part of yearly course and instructor evaluations. This lesson plan also set the premise for the lesson plans used in the experimental conditions, by making concise adaptations; we differentiated the training sessions for the three conditions. The main

difference between the experimental conditions and the control condition was that there was no explicit and systematic strategy instruction in the latter. The learning content was the same as in the experimental conditions except for the materials related to the explicit presentation of strategies (i.e., the TRAMPOLINE strategy practice slides on the introductory PPT, the TRAMPOLINE handout and the peer videos. The content of the first session was the same as in the experimental conditions except for the brief introduction to TRAMPOLINE strategies. Giving participants more time for self-discovery of the task requirements filled this absence. The only explicit task requirements on the PPT were: "underline the key points" and "paraphrase," mentioned prescriptively, as well as some sentence-level paraphrasing practice. As participants in all conditions had already studied APA in-text referencing, reporting verbs and linking words in the previous weeks, there was a brief reference to that on the PPT, but no explicit instruction was provided with alternative structures as in the experimental conditions. In sessions 2 and 3, students worked on the same synthesis task as the students in the other conditions. Taking the weak and strong synthesis samples as reference tasks, the instructor asked guided questions to elicit task requirements that are similar to the strategies in the other conditions, that is, "finding the main idea, supporting ideas and the formalities of effective paraphrasing in the second session and more surface-level concerns such as reporting verbs, linkers, APA conventions in the third session. Some guided questions were: "What would you include in your synthesis?" "Why did you choose that sentence?" and so on. Although the instructor followed a plan for the overall session, the questions needed to be partially improvised according to the answer of the previous question. Each cluster of task requirements was followed by individual practice, whole class-check and collaborative practice. Students read the extracts and underlined the key points individually. After checking the answers, the participants paraphrased the underlined points in groups of four (collaborative practice). The instructor supervised the activity and showed sample paraphrased sentences, which were reported using various reporting verbs, APA conventions, combined with linkers and required some editing, which students were expected to discover and mention. Session 4 was the same as in the experimental conditions (see Table 5 for the training session for two experimental conditions).

Table 5. *Training session for two experimental conditions. (MC: Modeling Condition, PC: Presentational Condition)*

Training session for experimental conditions					
Condition	**Aim(s)**	**Contents**	**Instructor / Student Activities**	**Teaching Techniques**	**Materials**
SESSION 1					
MC = PC	Creating a meaningful context for introducing the task and the strategies Developing background knowledge about the task & task requirements	Discussing a controversial topic to create context. Reading the extracts. Comparing weak and strong sample student synthesis texts. Finding task requirements in students' sample synthesis texts. Doing exercises about the task requirements	Instructor facilitates and moderates discussion to set the context and introduces the task with a PPT. Instructor and students discuss the strong and weak points of student samples. Instructor elicits the strategies and gives the TRAMPOLINE handout to the students. Students complete exercises for each strategy. Instructor checks for memorization of the strategies.	Direct Instruction Brainstorming Guided questions (to generate common knowledge and retrieve joint experiences) Elicitation Awareness Raising Joint reflection Whole class drill	- PPT and complementary worksheet with the two extracts, strong and weak student sample syntheses. - TRAMPOLINE strategies on PPT -TRAMPOLINE handout.
SESSION 2					
MC	Modeling of the use of TRAM strategies Scaffolding the use of strategies	TRAM Video and subsequent individual practice Paraphrasing (P) Video	Students read the extracts of a synthesis task and look up the vocabulary using online and paperback dictionaries. Students watch video 1 for TRAM strategies. Individual student emulation. Instructor projects the answer key on the board and gives whole class feedback. Students watch video 2 for Paraphrasing strategies	Peer modeling via video	- A synthesis task: *Reasons for the increase in divorce rate* - TRAMPOLINE strategies handout - TRAM and P videos
PC	Presenting TRAM strategies Scaffolding the use of strategies	TRAM Strategies and subsequent individual practice Paraphrasing (P) Strategies	Students read the extracts of a synthesis task and look up the vocabulary. Instructor explains how to use the TRAM strategies. Individual student emulation. Instructor projects the answer key on the board and gives whole-class feedback. Instructor explains how to use Paraphrasing strategies.	Teacher presentation	- A synthesis task: *Reasons for the increase in divorce rate*- TRAMPOLINE strategies handout

Training session for all conditions					
Condition	**Aim**	**Contents**	**Instructor / Student Activities**	**Teaching Techniques**	**Materials**
SESSION 3					
MC	Practicing paraphrasing strategies	Paraphrasing	In groups of four, students paraphrase the main ideas and the details in all the extracts that they worked on previously. Instructor supervises the activity, provides scaffolding and shows sample answers. Students watch video 3 for OLINE strategies.	Collaborative practice Monitoring Scaffolding Feedback	- A synthesis task: *Reasons for the increase in divorce rate* -TRAMPOLINE strategies handout - OLINE video
PC			In groups of four, students paraphrase the main ideas and the details in all the extracts that they worked on previously. Instructor supervises the activity, provides scaffolding and shows sample answers. Instructor presents the OLINE strategies.		- A synthesis task: *Reasons for the increase in divorce rate* -TRAMPOLINE strategies handout
SESSION 4					
MC = PC	Enabling independent practice	Completing a synthesis task	Instructor gives the practice worksheet and walks around the students, monitors each student and provides individual help when students ask for it.	Individual Practice Monitoring Scaffolding Feedback	- A synthesis task: *Factors that play a role in academic success*

3.5. Training Delivery and Intervention Fidelity

We trained the instructors of presentational and control conditions in a one-hour training session; and provided detailed lesson plans and the materials organized in a folder in the order to be followed. As an implementation check, instructors rated their integrity in implementing each of the critical steps of the intervention on a 100-point scale, and we observed each of the two instructors in one session of the experiment. Both the treatment fidelity scores (mean score: 90) and our observations yielded satisfactory results.

3.6. Measures

Table 1 shows the list of pre- and posttest measures of the study and their distribution across weeks. We measured text quality with students' exam papers, writing processes with the use of writing logs and motivational orientation with the use of a questionnaire adapted from Motivated Strategies for Learning Questionnaire (MSLQ).

3.6.1. Product measures: Motivation questionnaire data and writing performance

We used an adapted version of the MSLQ questionnaire developed by Pintrich, Smith, Garcia, and McKeachie (1991) by selecting relevant items from the instrument in its original language, that is, English. We selected a total of 17 questions from three scales of the instrument (i.e., six items for task value, eight items for self-efficacy and three items for intrinsic goal-orientation) measured on a 7-point Likert scale, pre- and posttest. The reliability of the whole scale was .93 for the pretest and .98 for the posttest.

We evaluated text quality through three different genres and four different tasks: an argumentative essay and a summary as pretests and two synthesis tasks as posttests. Participants took the tests simultaneously in pen and paper written exam conditions on the dates preset by the directorate. Students wrote on A4 papers with the exam prompts written at the top. Except for the summary, all the tasks were compulsory assessment components of the course counting towards students' General Point Average.

In the argumentative task, students were asked to write an essay of about 350 words in response to a 50-word prompt prepared by the testing unit (See Appendix A for the list of writing prompts). The other sections of the exam (i.e., listening, reading and writing) were clustered around one theme selected from previously covered topics. Students received the writing exam paper with the writing prompt in the last 60 minutes of the exam after all other test materials were taken from them. The task was to write a well-organized argumentative essay for or against the given prompt.

In the summary task, students were asked to summarize a textbook article in 150 words. In both synthesis tasks, students were asked to write a synthesis of 150-200 words integrating extracts from three different articles (each one paragraph), hereby referred to as sources, in response to a writing prompt which was around 25 words including the instruction. The tasks were identical to the tasks used for teaching and practicing in all three conditions of the intervention, but the students saw the content of the exam materials including the sources and the prompt for the first time in the exams. To prevent a possible distracting effect of completing the writing logs in the delayed posttest (cf. Table 5) from putting students at a disadvantage, we extended the duration of the delayed posttest an extra 10 minutes (i.e., 60 minutes as opposed to 50) and gave students easier extracts to synthesize (i.e., 9.8 on Flesch Kincaid readability tests in the delayed posttest, as opposed to 12.2 in posttest 1).

3.6.1.1. Rating Procedure

For rating purposes, the handwritten student papers were typed (on word documents) to eliminate any possible negative effect of student handwriting on the raters (Klein & Taub,

2005). We trained an outside panel of seven raters for rating the papers. The papers were divided over the seven raters, that is, the raters rated the texts in panels of two or three. All raters were second- or third-year bachelor students of English Language and Culture at the University of Groningen, the Netherlands; 20 to 23 years of age, with Dutch as their mother tongue.

We assessed the argumentative texts on four traits: (1) structural organization, (2) strength of the argumentation, (3) lexical richness, and (4) range and accuracy of grammatical structures Therefore, we used benchmark essays, since earlier research has demonstrated the positive effects of this rating procedure on rater reliability (cf. Schoonen, 2005; Tillema, Van den Bergh, Rijlaarsdam, & Sanders, 2012). For each trait, we used as benchmark essay a text from the 48 student texts written for pretest 1 that was of average quality with respect to that feature. The benchmark texts got an arbitrary score of 100. A total of three raters scored the other argumentative essays in comparison with the benchmark texts. If a text was considered twice as good as the benchmark text, it was scored as 200, if it was half as good, it was scored as 50, and so on. Each benchmark text was enriched with a list of the weak and strong points of the text with regard to the feature that had to be assessed with it, in order to help the raters focus on the right aspects when scoring a text on a particular trait.

Unlike the argumentative texts where students used personal experience and knowledge, in summaries and synthesis texts students worked with sources. For rating of text comprehensibility of the summary and the synthesis tasks, we did not inform the raters about the nature of the task to ensure that they were able to evaluate it for readers not acquainted with the sources. Additionally, the summary and synthesis texts were examined with regard to the incorporation of main ideas, supporting ideas and examples of the sources to summarize/synthesize. Two raters received lists with the main ideas, supporting ideas and examples of the source(s), and had to determine independently of one another the percentage of (1) the total number of main ideas, (2) supporting ideas and (3) examples in each student texts for both the summary and the synthesis tasks.

The synthesis texts were also analyzed on authenticity, source comprehension, and source use. For the last two analyses, the same kind of holistic scoring procedure was used as for the grading of the argumentative texts. Authenticity was identified with the function ‘Compare and Merge documents’ in Microsoft Word. After comparing and merging the original texts with the student texts, the parts of the student texts that overlapped with the source(s) were highlighted. Two raters subsequently calculated how many words of each student text were highlighted on a scale of 0 to 100%. The overlap percentage was subtracted from 100 and the corresponding value constituted the authenticity score of the student. If the overlap was 60%, the student received a score of 40. A higher score from authenticity meant less plagiarism.

Before the raters individually rated the argumentative texts, summaries and syntheses on the different dimensions, they had practiced the rating method together in a short training session, during which they received the benchmark texts, lists with the main ideas, supporting ideas and examples of the original texts, and/or the rating scale for the authenticity assessment. They read them carefully and used them to individually score six argumentative texts, summaries, or syntheses on a particular trait. When the raters differed in their scoring, they discussed possible reasons and solutions for their disagreement.

Table 6 shows the reliabilities of the rating of the texts written for pretest and posttests with benchmark essays. Consequently, we opted to use (the same method and) the same benchmark essays to score the texts of the delayed posttest as the ones in posttest 1. Tillema (2012) and Bouwer, Béguin, Sanders, & Van den Berg (2015) suggest that different tasks in the same genre can reliably be assessed with the same benchmark essays. Therefore, we

hypothesized that the quality difference of the synthesis texts written for posttest 1 and the delayed posttest could reliably be determined with rating scales that were developed for the scoring of posttest 1. The same raters assessed posttest 1 and the delayed posttest. Their scoring of the texts of the delayed posttest appeared to be reliable (cf. Table 6).

Finally, we calculated the mean of the scores the raters had given to the student texts on the particular traits. We determined the effect of our intervention with these mean scores.

Table 6. *Reliability in Cronbach's Alpha of the text scoring on the different traits (2 to 3 raters per text and 2 to 3 items for Cronbach's alpha)*

	Pretest 1: Argumentative essay	Pretest 2: Summary	Posttest 1: Synthesis text	Posttest 2: Synthesis text
Structural organization	.88			
Argumentation strength	.83			
Lexical richness	.84			
Grammar and punctuation	.90			
Incorporation of main ideas		.84	.70	.77
Incorporation of supporting ideas		.82	.86	.67
Incorporation of examples		.77	.87	.72
Comprehensibility		.94	.87	.82
Correct paraphrases			.79	.77
Source presentation			.93	.91
Authenticity			.97	.96

3.6.2. Process measures: Writing logs and learner reports

We tested the effects of the training on the writing processes of the students through the time sampled self-report method (i.e., writing logs) that Fidalgo, Torrance and Garcia (2008) implemented in their studies with six graders. In this method, students hear a bleep sound at regular intervals of 1-2 minutes during the writing and they are supposed to tick a box on the writing logs indicating the activity they are engaged in at that moment. We administered the writing log measure where participants responded to 45 bleep sounds concurrent with the delayed posttest. We showed the exact numbers of the bleeps projected on the board in case they lost track of the order of the bleep sounds. The eight activity categories in the writing logs were adapted from the original (Fidalgo et al., 2008) taking into consideration the possible activities that students would do when completing a synthesis task. We also included a simple graphic representation next to each activity category to help students locate the activity on paper easily. The activities in writing logs were categorized and defined as follows:

1. I am reading the sources: I am trying to understand the sources
2. I am paraphrasing: I am writing the sentences in my own words
3. I am working on the sources: I am trying to find the main idea important details, writer, year of publication, etc.
4. I am editing: I am making changes to the writing: correcting spelling mistakes, changing/-adding words
5. I am writing my text: I am writing my synthesis text
6. I am reading my text: I am reading through part or all of my text
7. Other: I am doing something unrelated: looking for a pen, looking out of the window
8. Finished writing

We adopted Torrance, Fidalgo and Garcia's (2007) strategy to train the students on how to complete a writing log prior to the actual practice (50 minutes in total). First, the participants went through the activity names, their explanations and the graphic representations. Then, they watched a video of a student model doing the synthesis task interrupted with occasional bleeps at different moments. Students were asked on a demo version of the writing log to tick the box, which showed the writing activity category that the model was engaged in at the moment they heard the bleep sound. We made sure that the students were able to distinguish between the different writing categories. We checked and discussed the answers with the students. Subsequently, we simulated the exam conditions and gave the participants a synthesis task similar to the one they would do in the exam and the writing logs. Finally, in the delayed posttest, after a quick reminder of the different categories of writing activities on the writing logs, students wrote their synthesis texts and filled out the writing logs simultaneously.

To increase validity of the results, we used multiple process measuring methods (Schellings & van Hout-Wolters, 2011), so combined online self-reported data of the writing logs with the offline learner report method (De Groot, 1980). The latter was used to provide insight into learners' experiences and identify the extent of conscious knowledge that students were able to retain after the training prompted by open-ended questions (See Appendix B for the questions). After an instruction and a standardization session, two coders (i.e., the researcher and another instructor) independently coded the responses. The two coders found and underlined the predetermined key words in student reports (i.e., main idea, details, paraphrasing, summarizing, organization, linkers, APA, think, edit, steps/stages, time-management). As a second step, they clustered the key words and relevant semantic units under four categories: main ideas and details under *content management*; paraphrasing and summarizing under *synthesizing skills*; organization, linkers, APA under *source use skills*; think, edit, steps/stages and time management under *process knowledge*. The inter rater reliability was 0.70 based on a sample of 10 cases.

4. Analyses

To test Hypothesis 1, that is, whether peer modeling results in qualitatively better synthesis texts compared to presentation of strategies (1A) and presentation results in qualitatively better synthesis texts compared to a control condition (1B), we used students' summaries and argumentative writings as two pre-tests and two synthesis texts as post-tests. There was a positive correlation between the four subscores of pretest 1 Summary of a single source, that is, main ideas, supporting ideas, examples and text comprehensibility, ranging from (r =.436, p =.003) between main ideas and text comprehensibility to (r =.550, p <.001) for examples and text comprehensibility, and Cronbach's alpha reliability coefficient yielding highly satisfactory results (.77). Thus, we merged these four subscores and created a composite score, hereby referred to as *pretest score 1 for Summary of a single source.* In the same way, a statistically significant positive correlation was observed between the four subscores of pretest 2 Argumentative writing, that is, organization, strength of argumentation, lexis and grammar ranging from (r=.387, p =.007) between lexis and organization to (r=.641, p <.001) for strength of argumentation and organization. Reliability analysis for the different measures of pretest 2 yielded a highly satisfactorily alpha coefficient (α = .823). In subsequent analyses we used a composite score for pretest 2 Argumentative Writing, hereby referred to as *pretest score 2*. These two composite pretest scores provide us with a strong and valid (i.e., as more generalizable) measure of students' initial level of writing skill in two different genres. Controlling for students' initial writing skills across genres by including the two pretest measures as covariates in the analyses will increase the validity of our results (i.e., effect after having balanced out pretest differences in writing skill).

In the two (synthesis) post-tests, that is, posttest 1 and the delayed posttest, we had seven subscores, that is, *main ideas*, *supporting ideas*, *examples*, *text comprehensibility*, *source comprehension*, *source use and plagiarism.* There was a strong correlation between the subscores of *main ideas*, *supporting ideas*, *text comprehensibility* and *source comprehension*, with Cronbach's alpha reliability coefficient .88 for posttest 1 and .75 for the delayed posttest. Hence, we merged these subscores and calculated a composite score, subsequently referred to as *content*. Although there was a low correlation between supporting ideas and text comprehensibility in the delayed posttest (r =.264, p = . 070), there was high correlation between the rest of the values ranging between .408 and .680, which indicated a good internal consistency for merging of the sub-scores, so we followed the same procedure for standardization purposes. The examples subscale was expected to belong to the content composite score, but there was no correlation with any of the subscores for posttest 1. For the delayed posttest, there was low correlation with supporting ideas (r =.322, p <.05), but not with other subscores, so we eliminated the score from the final analysis. Therefore, in the final analysis we had three aspects for the quality of the two posttests (i.e., synthesis texts), that is, *content*, *authenticity* and *source use* (cf. rating procedure).

The triadic subset of TRAMPOLINE strategies corresponds with the three aspects for rating the quality of student texts and the composite scores (See Table 7). Thus, we expected to observe mastery in the related set of strategies in the corresponding quality of paper, and hence, the corresponding (composite) scores.

To explore whether one of the learning conditions resulted in better scores for a particular group of participants than another condition, we analyzed interactions between the three conditions and two learner variables: motivation and writing skills, both based on pretest scores, on three aspects of posttest text quality: quality of content, source use and authenticity. We applied Hayes moderator regression analyses (Hayes, 2013), as add-in in SPSS, which allowed us to estimate the regions within the moderator variable in which differences between the learning condition were statistically significant, using the Johnson-Neyman procedure. We present the explorations per posttest variable.

Table 7. *A cross match of the TRAMPOLINE strategies, text quality aspects and the (composite) scores*

Strategies	Text Quality Aspects	(Composite) Scores
Think Read Ask – What is the main idea Mark the important details	Main ideas Supporting ideas Text comprehensibility Source comprehension	Content
Paraphrase the main idea and the details	Plagiarism	Plagiarism
Organize the ideas BY Linking the sentences Including APA Nesting reporting verbs	Source Use Skills	Source Use Skills
Edit your text	(Not rated in text quality Measured with writing logs)	

To test Hypothesis 2, that is, whether peer modeling leads to improved synthesis writing processes compared to presentation of the strategies (i.e., more meta-cognitive activities in the initial stages and more executional activities in the later stages of writing) (2A) and presentation of the strategies leads to improved synthesis writing processes compared to a control condition (2B), we used writing logs and students' learner reports. To analyze the writing logs, we divided the process time of each individual student (calculated by the total number of responses in the log) into three phases (i.e., Phase 1, Phase 2 and Phase 3) of equal length, based on the number of responses to the bleeps (max. 45, each bleep occurring on average every 90 seconds, at random intervals of between 60 and 120, a practice in line with previous studies for generalizability purposes). It is customary to divide the writing process into stages to interpret the temporal distribution of the cognitive effort over the writing process (Rijlaarsdam, Van den Bergh & Van Steendam, 2015). When the result of the division of the number of responses into three phases was not a whole number, we transferred the surplus value in the decimals to Phase 3 (e.g., for 13 responses for student x, four were allocated to phase 1, four to phase 2, and then five to phase 3). We calculated the frequency of writing activities for each Phase calculated by the number of activity items reported in the logs.

To analyze the learner reports, we counted the number of categorical statements of all conditions and calculated proportions for each condition (by dividing the total number of statements by the number of students in that condition). We applied analysis of variance to analyze the learner reports with the writing pretest as a covariate.

As an implementation check, we tested the improvements in students' motivational orientation over time through MSLQ, by applying analysis of variance, with the measurements as within factor and conditions as between factors.

5. Results

We hypothesized that strategy instruction would be more effective in improving students' text quality and writing processes compared to regular curriculum instruction (i.e., the Control Condition) and that modeling would have an effect over and above the two conditions.

5.1. Preliminary Analyses

We checked students' motivational orientation as an implementation check through MSLQ and did not find initial differences between conditions ($F(2,42) = 1.596$, $p = 0.214$, $\eta^2 = .066$). There was an effect of time ($F(1,42) = 11.724$, $p = 0.001$, $\eta^2 = .207$) indicating a progress over time, but no interaction between time and condition ($F(2,42) = 0.697$, $p = 0.503$, $\eta^2 = .030$). See Table 8 for the mean pre- and post-test scores for motivational orientation.

Table 8. *Pretest and posttest means and standard deviations for motivational orientation (7-Point scale 1=Strongly Disagree, 7=Strongly Agree)*

Variable	Condition	Pretest		Posttest	
		M	*S.D.*	*M*	*S.D.*
Overall Score	Modeling	4.80	0.76	5.40	0.92
	Presentational	5.09	0.69	5.35	0.74
	Control Condition	4.51	1.08	5.21	1.11

5.2. Text Quality

Table 9 shows mean text quality scores for the pretests for the three conditions. For pretest score 2, no statistically significant differences were observed between the three conditions ($F(2,44) = 0.460$, $p = 0.634$, $\eta^2 = 0.020$). For pretest score 1, however, an analysis of variance (ANOVA) showed a significant difference between the conditions ($F(2,44) = 3.964$, $p = 0.027$, $\eta^2 = 0.162$), so we used Tukey post hoc tests to identify sample means that are significantly different from each other. This analysis showed that there were initial differences between control and the presentational conditions for pretest 1 (*MD*: -21.14, $p = 0.047$) with the control condition scoring lowest. No differences were found between the presentational and modeling conditions (MD: 3,028, $p = 1,00$) or between modeling and the control conditions (MD: 18,109, $p = 0,062$). Consequently, we will include both pretest scores as covariates to adjust for initial level of writing skill.

Table 10 A shows text quality scores for writing performance in posttest 1 for all three conditions. A multivariate analysis of covariance including the two pretests as covariates with the three text quality measures for posttest 1 (content, authenticity and source use) showed a statistically significant condition effect: $\Lambda = 0.682$, $F(6, 74) = 2.566$, $p = 0.026$, $\eta^2 = 0.176$). Separate follow-up ANOVAs on the outcome variables showed no effect of condition for content and plagiarism was observed ($F(2,38) = 2.301$, $p = 0.114$, $\eta^2 = .108$, and ($F(2,38) = 0.537$, $p = 0.589$, $\eta^2 = .027$ respectively). For source use, a condition effect was observed in favor of Modeling ($F(2,38) = 3.905$, $p = 0.029$, $\eta^2 = .170$).

Table 10 B shows text quality scores for writing performance in the delayed posttest for all three conditions. A multivariate analysis including the two pretest composite scores and also partialling out the effect of the content, authenticity and source use scores of posttest 1 (including 5 covariates), using Wilks's Lambda showed a statistically significant main effect for condition $\Lambda = 0.612$, $F(6, 66) = 3.057$, $p = 0.011$, $\eta^2 = 0.217$. For content and plagiarism, no effect of condition was observed ($F(2,35) = 0.175$, $p = 0.840$, $\eta^2 = .010$, and $F(2,35) = 0.844$, $p = 0.439$, $\eta^2 = .046$ respectively). For source use, a condition effect was observed in favor of Modeling ($F(2,35) = 9.426$, $p = .001$, $\eta^2 = .350$).

Table 9. *Mean text quality scores at pretests for three conditions*

	Modeling			Presentational			Control Condition		
	N	Mean	SD	N	Mean	SD	N	Mean	SD
SSS	19	52.01	20.59	12	55.04	22.14	13	33.90	20.27
ARG	18	81.37	15.84	13	86.20	17.75	16	81.20	13.72

ARG: Argumentative essay, SSS: Summary of a single source

Table 10. *Mean text quality scores for synthesis texts for three conditions*

A. *Posttest 1*

	Modeling			Presentational			Control Condition		
	N	*Mean*	*SD*	*N*	*Mean*	*SD*	*N*	*Mean*	*SD*
Content	18	67.33	27.59	12	57.06	20.56	13	67.42	21.01
Authenticity	18	85.11	22.58	12	91.25	10.48	13	91.61	12.39
Source Use	18	123.44	51.46	12	76.25	32.20	13	93.77	54.70

B. *The delayed posttest*

	Modeling			Presentational			Control Condition		
	N	Mean	SD	N	Mean	SD	N	Mean	SD
Content	18	84.21	18.28	12	87.02	23.34	13	76.44	18.95
Authenticity	18	88.83	16.20	12	87.50	17.52	13	97.00	6.22
Source Use	18	155.17	30.91	12	73.50	41.48	13	120.15	40.71

5.3. Explorations: Interactions between Learner Characteristics and Learning Condition

5.3.1. Effect of initial levels of motivation

When testing for moderator effects for students' initial level of motivation (pretest based and with Hayes moderator regression analyses (Hayes, 2013), a statistically significant interaction effect was shown on source use and authenticity of posttest 1. For the content aspect of the text quality of posttest 1 no statistically significant interaction effect was observed. For source use, the regression between pretest motivation scores and the modeling condition was strongest. Next to main effects of pretest motivation scores (t = 2.36, p = .02) an interaction between motivation and learning condition was found (t = -2.18, p = .03). The effect was statistically significant for the 35% most motivated participants. In the modeling condition the most highly motivated students produced the best texts whereas the most motivated students in the control condition wrote the poorest texts. The condition did not affect the source use scores of the other 65% of the participants.

For authenticity, we observed a main effect of pretest motivation (t = 3.09, p = .001), and condition (t = 3.06, p = .001) as well as an interaction between motivation and authenticity scores (t = -2.75, p =.01) (See Figure 1). Here the interaction holds for the lower scoring group (37%) on pretest motivation: participants in this group scored significantly lower on authenticity in the modeling condition than in the control condition. The general pattern observed is that the modeling condition is more sensitive to motivation than the control condition (See Figure 1).

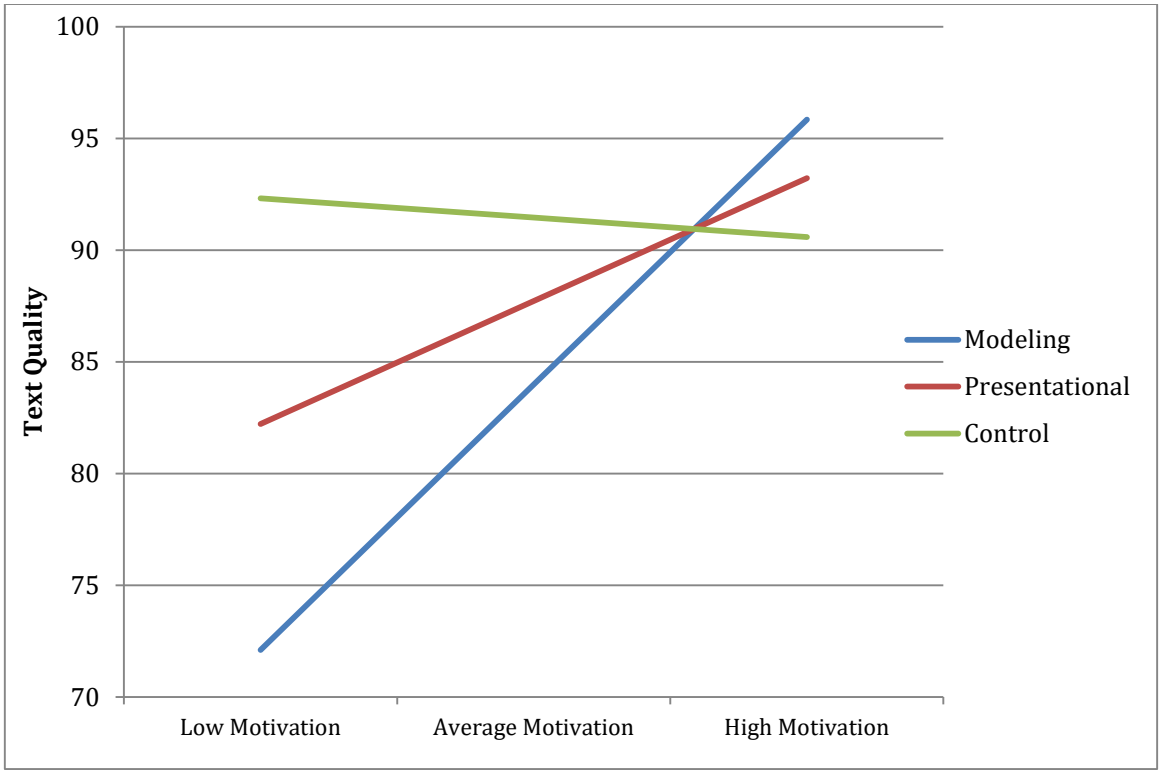

Figure 1. Interaction between learners' initial levels of motivation and learning condition on authenticity of synthesis texts

5.3.2. Effect of initial levels of writing performance

For the pretest Writing Performance we had two pretest scores that did not correlate to a statistically significant degree (r = .26, p = .086): pretest score 1 and pretest score 2.

For *source use* no interaction effects were observed. For c*ontent* we found an interaction between pretest score 1 and learning condition (t = -3.30, p = <.01) next to a main effect of the pretest scores (t = 2.44, p = .02). The interaction effect is significant for the 25% lowest scores on the pretest, as well as for the 25% highest scoring group of participants. The effect is strongest in the modeling condition, and non-significant in the control condition. Participants who have a relatively high score in Pretest score 1, score best in the modeling condition; while the participants scoring lowest in the pretest score 1, score lowest in this condition.

For the scores on *authenticity* at the posttest, we found an interaction between the pretest score 2 and learning condition (t = 2.29, p = .03). The effect is significant for the 26% highest scoring group of participants in the pretest. Again, the strongest effect of pretest scores is in the modeling condition: participants with relatively high pretest scores scored significantly lower on posttest *authenticity* in this condition than the participants that started with lower pretest scores. This effect is not observed in the control condition (See Figure 2).

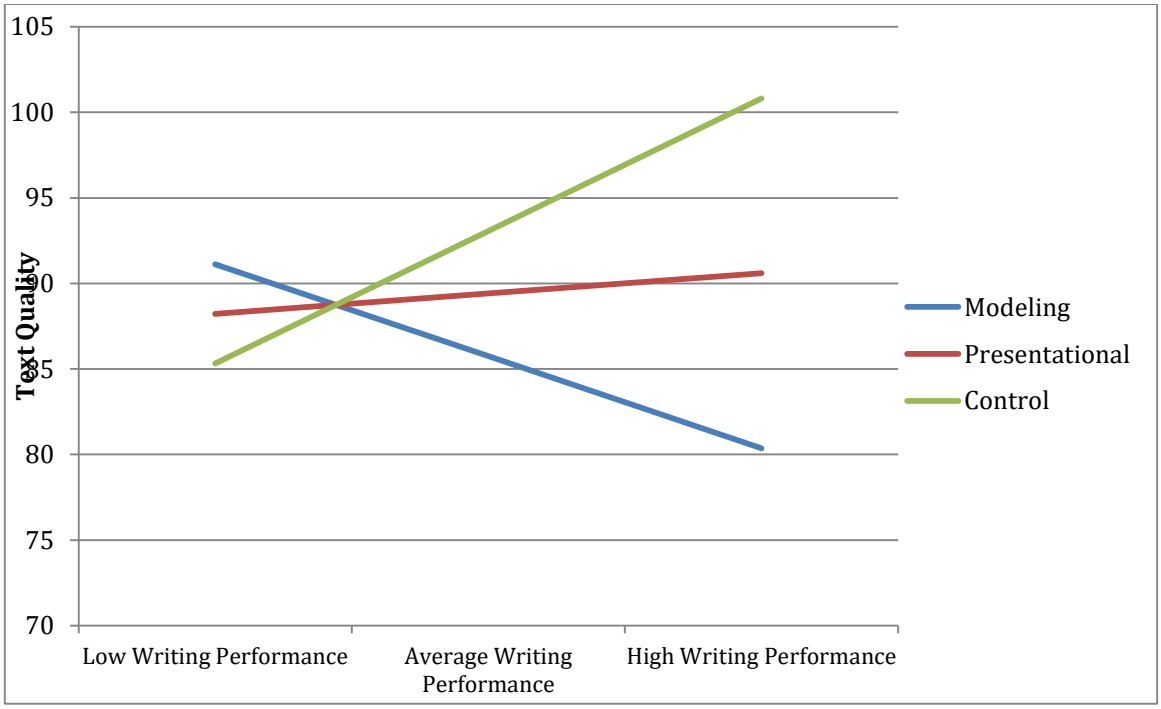

Figure 2. Interaction between learners' initial levels of writing performance and learning condition on authenticity of synthesis texts

5.4. Writing Processes

Table 11 shows the process data for the three conditions. The numbers in the table stand for the average number of the reported activities, that is, the frequency of each activity per phase for each condition. To have an overall picture of the most frequently reported activity in each condition, we also calculated percentages for each phase by adding up the averages for each one of the eight activity items and dividing it by the average for the relevant activity item.

Results showed that students reported spending most of their exam time, that is, 45 bleeps/60 minutes, *writing* their summaries (i.e., for 31% of the time). In Phase 1 of the writing process, the modeling and the presentational conditions wrote their syntheses 26% and 38% of the time, respectively, whereas the control condition reported reading the sources 32% of the time. There is a significant difference between conditions in the activity *reading the sources*: the control condition reported spending more time on the activity than the other two conditions ($F(2,42)$ 4.743, p = 0.014, η^2 =.184). In Phase 2, the modeling condition and the presentational condition reported writing their syntheses for 31% and 41% of the time respectively, whereas the control condition reported reading their own synthesis texts 35% of the time. A condition effect was observed for *working on the sources* with the modeling condition spending more time on this activity than the presentational condition ($F(2,42)$ 2.358, p = 0.048, η^2 = .101). In Phase 3, the proportion of total writing time spent in the activity *writing my synthesis* was 33%, 32% and 55%, respectively. The control condition reported spending more time on this activity in this phase than the experimental conditions ($F(2,42)$ 6.678, p = 0.003, η^2 = .241).

Table 11. *Process results: frequency of activities per condition and phase*

	Modeling		Presentational		Control	
	M	*S.D.*	*M*	*S.D.*	*M*	*S.D*
Phase 1						
Reading the sources	2.68	1.77	2.00	1.15	3.81	1.68
Paraphrasing	2.63	2.39	2.23	1.88	3.12	2.03
Working on the sources	3.05	2.20	1.85	1.52	1.87	1.41
Editing	.32	.58	.38	.51	.19	.54
Writing my text	3.16	2.59	4.15	3.18	3.00	2.83
Reading my text	.32	.67	.46	.88	.13	.50
Other	.11	.32	.08	.28	.56	1.26
Phase 2						
Reading the sources	1.68	1.42	1.85	1.91	1.00	1.41
Paraphrasing	2.95	2.70	2.08	2.25	2.19	2.17
Working on the sources	2.89	2.83	.85	1.07	2.19	2.32
Editing	.42	.61	.92	.86	1.31	2.75
Writing my text	3.58	2.69	4.38	2.63	4.75	2.35
Reading my text	.21	.42	.77	1.36	.81	2.17
Other	.53	1.02	.31	.63	.44	.63
Phase 3						
Reading the sources	.32	.75	.38	.77	.94	1.88
Paraphrasing	1.11	1.63	1.46	1.98	1.00	1.21
Working on the sources	1.47	1.81	.77	.93	.56	1.15
Editing	2.11	1.59	1.92	2.43	1.13	2.13
Writing my text	3.68	2.08	3.15	1.86	6.50	3.16
Reading my text	2.47	1.65	2.46	2.11	1.87	1.86
Other	.63	.90	.69	.95	.88	1.31

5.5. Learning Experiences

Table 12 shows the results from the learner reports for the three conditions (see Appendix B for a sample learner report). The numbers in the table show the percentage of the average number of categorical terms mentioned in each condition. The majority of the participants in the modeling condition reported learning experiences mostly about process knowledge, in the presentational condition about synthesizing skills and in the control condition about content management. No effect of condition was observed for content management ($F(2,42) = .171$, $p = .843$, $\eta^2 = .008$), synthesizing ($F(2,42) = .757$, $p = .476$, $\eta^2 = .035$) and source use skills ($F(2,42) = 2.348$, $p = 0.108$, $\eta^2 = .101$). The only category that seemed to be sensitive to conditions was process knowledge with an effect of ($F(2,42) = 3.418$, $p = .042$, $\eta^2 = .140$), in favor of modeling condition.

Table 12. *Results from learner reports: mean percentage and standard deviation of statements per condition and category*

	Modeling				Presentational				Control			
	Weak		Strong		Weak		Strong		Weak		Strong	
	Mean%	*SD*	*Mean%*	*SD*	*Mean%*	*SD*	*Mean%*	*SD*	*Mean%*	*SD*	*Mean%*	*SD*
Content management	3.4	.036	1.2	.026	1.9	.022	1.7	.020	1.7	.020	3.1	.032
Synthesizing skills	1.7	.015	2.4	.015	3.4	.010	2.0	.017	1.6	.017	2.4	.016
Source use skills	2.7	.029	3.1	.025	0.0	.000	1.9	.025	1.6	.024	2.1	.019
Process knowledge	3.6	.018	3.2	.039	2.7	.034	.08	.016	1.0	.017	1.4	.020

Mean percentages are calculated according to the occurrence of key phrases per student by multiplying the total number of responses by the number of participants in each condition divided by total number of participants in the study.

Students in the experimental conditions expressed improved process knowledge, especially about managing the stages and/or steps and using their time more efficiently in the timed writing synthesizing tasks of the Posttests with sentences such as:

Students no 8: *'I learnt if I want to synthesize I should paraphrase first.'*

Students no 10: *'Synthesizing is not as hard as I thought. I noticed the ways to write a synthesis in an orderly way.'*

Student no 12: *'When I Iooked at the exam paper (before the training) everything interfered. But now I know what I am doing. There are some processes for writing something.'*

Student no 18: *'Before the training I read all the extracts at once, now I learnt that I have to work on one extract at a time.'*

Student no 32: *'It works like processing (any) work. You can do everything step by step.'*

Hence, the learner reports pointed to positive learning experiences and outcomes for students in PO.

6. Discussion

We examined the effects of modeling and presentational modes for strategy instruction as compared to a control condition, which was the regular curriculum instruction of the institution on students' synthesis text quality, synthesis writing processes and learning experiences. This study is unique in that it tests the effectiveness of the separate components in strategy-focused instruction and in an L2 context. We partially confirmed hypothesis (1A). Modeling condition resulted in better synthesis texts in the posttest and the delayed posttest than the presentational condition for *source use* (i.e., knowledge of the available sources, using a variety of citation techniques, correct use of APA conventions, including a variety of reporting verbs and linkers), one of the three aspects of text quality, but not for *content* and *plagiarism*. However, we did not confirm hypothesis (1B). Presentational condition did not differ in effects from the control condition for synthesis text quality.

The success of the modeling condition in two consecutive posttests in the source use aspect of text quality and not in content and authenticity was an intriguing finding. It has been shown in previous studies (Braaksma et. al, 2002), that observation works best for novel tasks. Unlike discursive essays with similar organizational and thematic requirements that the students are mostly familiarized with throughout their preparatory year, writing from sources is a novel task both within the L1 and L2 writing curricula. Source use skills are predominantly related with the specifics of this novel task rather than the common skills across different writing tasks such as paragraphing, argumentation, etc., so this unfamiliarity with the novel features of the task might have contributed to students' success, all of which have been rated under the source use skills aspect of text quality. Another reason for the relatively lower gains of the modeling condition in content and authenticity might be the underlying linguistic skills needed to improve these aspects of text quality, which tap into students' reading comprehension, content selection, paraphrasing, elaboration, and writing formulation skills. These skills are cognitively demanding "higher order" skills (Mateos & Sole, 2009), and they are, to a great extent, dependent on L2 proficiency. Thus, a strategy-focused instruction session conveyed through modeling or presentation may not have sufficed to improve formulation skills in L2 that are required to independently apply the strategies for improving content and authenticity. Improvement in these skills may be possible in long-term strategy training programs, or in an L1 context where it has been

shown that speed of access to linguistic features, (i.e., fluency) poses less of a hindrance to writing performance than is the case in an L2 context (Schoonen, et al., 2003).

Preliminary analyses on the motivation pretest showed (cf. Results section) no differential effects of condition on motivation and contrary to our expectations, we did not find an effect of modeling on students' motivational orientation, including their self-efficacy beliefs (Bandura, 1986). Thus, the higher performance of the modeling condition in the source use aspect of text quality cannot be attributed to differences in motivation. However, it was interesting to observe in the analyses for initially more and less motivated students (with initial level of motivation as a moderator variable) that 35% of the most motivated students fared best in the modeling condition for the source use aspect of their text quality; whereas, the most motivated students in the control condition wrote the poorest texts for the same aspect. Hence, for success in mastering the novel aspects of the task (source use in this task), it seems that a higher level of initial motivation combined with modeling yields better results than the regular curriculum instruction in this institution. Another interesting finding was that 37% of the least motivated participants scored significantly lower in the modeling condition in the authenticity aspect of their text quality. This provides further evidence that modeling is the most sensitive condition to motivation and that in future studies assigning students on the basis of motivation to an instructional condition can be an interesting exploration.

For generalization purposes, we also looked into the effects of conditions on students with differing levels of initial writing performance (with initial level of writing performance as a moderator variable). Analyses showed that 25% of the highest scoring participants in pretest score 1 (summary of a single source) scored highest in posttest *content* in the modeling condition; and accordingly, 25% of the lowest scoring participants wrote the poorest texts in this condition. However, a reverse effect is observed for the 26% of the highest scoring participants in pretest score 2 (argumentative essay), who scored significantly lower on posttest *authenticity* in the modeling condition than the participants that started with lower pretest scores. The differing condition effects on the success of the students in the posttests could be the result of the different nature of the tasks, that is, summary of a single source in the former and argumentative essay in the latter. As being able to summarize a single source is a prerequisite in a synthesis task, it is presented as one of the strategy steps in the teaching of the synthesis task in the presentational and the modeling conditions. It may explain students' higher scores on the content aspect of the posttests in the modeling condition, which was not observed for the students who scored highest on the pretest score 2. The fact that the results were only observed in the modeling condition and not in the presentational condition is further evidence to the added benefit of observation in teaching of the strategies. These results are in line with the results of previous studies, which have demonstrated different condition effects on learners with different aptitude levels in composing written texts (Braaksma et al., 2002) and on learners with different creativity levels in creating design products (Groenendijk, Janssen, Rijlaarsdam & Van den Bergh, 2013).

We used a multiple methods approach (i.e., writing logs and learner reports) in measuring the writing process to answer the question whether strategy-focused conditions lead to more effective temporal organization of cognitive activities (Hypotheses 2A and 2B) than the regular curriculum instruction. We evaluated the writing log data according to the parameters we set on the basis of previous studies (cf. Modeling section). Accordingly, instead of executive writing activities spread over the whole writing process in a linear manner, we expected more metacognitive, or planning activities (i.e., reading the sources, working on the sources) in the initial stages and more executive activities (paraphrasing,

reading the synthesis text, writing the synthesis text) in the later stages of the writing task. We also expected two different ways of dealing with the sources: reading to understand the sources and strategic reading of the sources (i.e., finding, underlining and/or highlighting the main ideas, supporting ideas, etc.), as is labeled in the writing logs as *working on the sources*. Thus, we expected writers with effective temporal organization of cognitive activities to spare a minimum amount of time on *reading the sources* and more time on *working on the sources* especially in Phases 1 and 2 of the writing process. In Phase 3, more executive activities should ideally follow.

The analysis of writing log data could only partially be explained with the aforementioned parameters. In Phase 1, the control condition reported being engaged in the activity *reading the sources* (trying to understand the sources). Their reading time may have been extended due to a failure to strategically work on the sources. However, modeling and presentational conditions did not prove to have better strategies as they were predominantly engaged in writing their texts, an activity ideally expected as the dominant activity in Phase 3. In Phase 2, the modeling condition reported being engaged in working on the sources, a very important indication of effective temporal organization of cognitive activities, in a way following the TRAMPOLINE strategy steps that were shown to them in the peer videos, that is, planning before writing. In Phase 3, all three conditions reported writing their texts, as would ideally be expected.

In none of the conditions did students report editing their texts. One reason might be that editing was only minimally modeled in the videos with mostly sentence level corrections; thus, it may not have taken up writing process time at a statistically significant level. Previous studies also showed that even in higher education only a minority of students improved revision skills with instruction in L1 (Torrance, Thomas & Robinson, 1999); and in L2, less experienced writers only detected surface level corrections instead of making global revisions (see Van Steendam, Rijlaarsdam, Sercu, & Van den Bergh, 2010). It is also possible that the effects of the training did not emerge within the time span of the training program, because editing emerges later than planning in developing writers in L1 (Berninger & Swanson, 1994), which may also be the case in L2 learners. Additionally, an improved planning phase may have cut down on the time devoted to editing, so there might have been a trade-off in favor of planning (Torrance et al., 2007; Martinez, et al., 2015).

The results of the learner reports showed that the modeling condition reported to have more process knowledge required for writing a synthesis text. They reported learning how to deal with the steps of writing a synthesis text and/or, attributed their success to this process knowledge, a finding in line with previous research (cf. Groenendijk, Janssen, Rijlaarsdam & Van den Bergh, 2013). Why this is not entirely reflected in the process registration technique (i.e., writing logs) we used is an issue to consider. The discrepancy might have been a result of low concurrent validity of online versus offline process registration techniques (cf. Veenman, 2011). As we distributed the learner reports two weeks after the actual writing task, the writing process might have been reconstructed in the writer's memory with possible memory failure and distortions (Schellings & van Hout-Wolters, 2011). Also, at preparatory schools in Turkey, it is customary to write in pen and paper conditions in exams. Thus, for ecological validity reasons, we were not able to combine self-reported data with online process registration techniques such as keystroke logging to record the writing activity in real time and provide data from different perspectives, which give more leeway for interpretation. In future studies, it might be worthwhile to use multiple online registration techniques to get more insight into the writing processes of students.

In evaluating the results, some issues of validity should be discussed. Several studies showed that within a certain participant profile, initially stronger writers (in terms of aptitude) benefit more from observation than initially weaker writers (Groenendijk et al., 2013). Also, strong writers benefit more from observing mastery models and weak writers from coping models, also known as model-observer similarity (Braaksma et al., 2002). Although we looked into the effects of the training on initially weaker and stronger writers within this specific group, and reported the results, we may expect the effects to be different on a group of academically high achieving students. The model in this study was designed to demonstrate a standard student performance; however, considering the low-achieving student profile, we may not have been able to create the best environment for model-observer similarity. In a future study, special attention should be paid to student profile and heterogeneity of the participants in terms of competence and motivation that could affect the results of the study.

Previous research (Braaksma, Van den Bergh, Rijlaarsdam & Couzijn, 2001; Sonnenschein & Whitehurst, 1984) has shown that observation results in larger learning gains when it is set with evaluation and elaboration tasks. In our study, such a task was absent; instead, we opted for presenting the target behavior in digestible chunks of knowledge (Bandura, 1986), as in the presentation of the three subsets of strategies (i.e., TRAM, P, OLINE) which is, arguably, another way of channeling learner's attention to the modeled activity. Nevertheless, in future studies, it is worthwhile to include evaluation and elaboration tasks in the observation activity to ensure students' utmost engagement with the modeled behavior.

Another issue to consider is the assignment of the classes to conditions. Participants were randomly assigned to three classes by the administration prior to the training, but the assignment of intact classes to conditions had to be arranged in line with institutional priorities. The main instructor of each class had to give the training in their own classes because having a totally different instructor teach in each class would mean interfering with the natural course of the module, which would not be favored by the students or supported by the administration. We did not have the researcher give the training in each class because she was the main instructor in one of the classes that was assigned to a condition and in that case her class would have had the privilege to be instructed by their own instructor and the others would not. To control for a possible trainer effect on the results, the researcher's class was appointed to the modeling condition where the interaction between the instructor and the students was minimal compared to the other conditions, because the key element of the intervention, that is, the strategy instruction, was administered through peer videos. Thus, the assignment of the modeling condition was not random, but the presentational and the control conditions were randomly assigned to intact classes. We believe that in future studies, the same instructor should train the students in all conditions or a design with random assignment of students to conditions should be implemented.

Another limitation with regard to the trainers might be their different profiles. Although the trainers in the experimental conditions held a similar profile in terms of nationality, educational background and gender (i.e., Turkish, with an MA in literature, PhD candidates and female), the trainer in the control condition was Canadian, with an MA in ELT and male. This difference may seem to be in favor of the control condition since having a native speaker as the instructor might have improved students' motivation and possibly overall language proficiency. However, it is a strict requirement at Turkish preparatory schools that the lessons are conducted in English (British Council, 2015, p. 93), especially at this university for standardization purposes since half of the instructors were native speakers of

English at the time. Therefore, any possible effect of nationality might have been counteracted by the rules of the institution.

We should also point out the fact that students are placed at pre-faculty level classes according to certain criteria. In the previous academic year, they either failed the pre-faculty level exit exam; or passed the upper-intermediate level exit exam, but failed another level previously. Thus, the participants are a homogenous group of learners with an Upper-intermediate initial level of language proficiency, tested through a four-hour exam prepared by the testing specialists of the university in four skills (i.e., reading, listening, speaking and writing) according to carefully designed testing specifications of the level. Students' initial motivation and level of writing skill in L2 was tested as well, to control for pretest differences. Initial differences in writing were taken into account by running analyses of covariance.

It should also be mentioned that, although we staged our sessions similar to an SRSD program especially focusing on two of its main stages, that is, *modeling* and *memorization* (through mnemonics), this intervention is not a strategy training program, but an attempt to improve the regular curriculum instruction of a specific task through strategy-focused instruction. Although full strategy training programs are likely to yield better results, we found it worthwhile to decompose the training program and address each component separately to see the main acting agents in the training. This study is also an example of how training programs can be tailored to fit in a realistic time slot allocated for the actual teaching of a specific task. Although this is one reason why the results are not generalizable to studies that operationalize a form of SRSD, it is also a strong aspect of the study towards ecological validity as trainings in hectic EFL programs need to be carefully fit in the time slot allocated for specific tasks for reasons of practicality.

It has been suggested that to be able to reach generalizable conclusions in L2 research with secondary level students, there should be a total of 3 to 4 assignments per student each rated by two raters (Schoonen, 2005). In this study, we measured students' text quality in three different genres (i.e., summarizing, synthesizing and argumentative writing) with four different tasks, two pretests and two posttests. Thus, the results of text quality are generalizable within the strategy-focused writing intervention programs and provide grounds for future strategy-focused interventions in L2 writing.

7. Conclusion

Our results are complementary to previous studies, which proved the effectiveness of the modeling component of strategy-focused instruction on certain aspects of writing performance. There are two unique aspects of this study that, to the best of our knowledge, are not present in other studies. First, it separately investigates the effectiveness of two distinct instructional modes both of which are typically present in strategy-focused instruction studies, that is, *modelling* and *direct instruction.* The results of one study that singles out the separate components of a strategy-focused programme by Fidalgo et al., (2015), differs from our study in that in the latter (1) participants are 6 grade Spanish students instead of L2 students in higher education and (2) L1 tasks are in a different genre (i.e., compare-contrast and opinion essays instead of synthesis writing). Additionally, (3) the differences in the designs of the two study seriously affect the comparability of the two studies. In Fidalgo et al. (2015) direct instruction follows modeling, and the type of modeling is implicit; whereas, in our study direct instruction precedes modeling, and the type of modeling is explicitly supported by mnemonics.

Secondly, we conducted this study in ecologically valid circumstances and the effectiveness of the intervention was tested compared to regular curriculum practice, which tends to be standard across universities in Turkey. Thus, it has the potential to be insightful for similar teaching environments to revisit their practices and encourage further research. Although this specific university is one of the few institutions, which includes synthesis writing in their preparatory school curriculum, in the departments of other universities where the medium of instruction is English, synthesis writing is a common task in compulsory academic writing courses starting in the freshman year. Synthesis writing is also a task that is commonly encountered in faculty departments, both in L1 and in L2 contexts. This expands the scope of the results of this study to educational practices in other universities in Turkey and other comparable institutions around the world.

References

Baker, S., Gersten, R., & Graham, S. (2003). Teaching expressive writing to students with learning disabilities: Research-based applications and examples. *Journal of Learning Disabilities, 36,* 109 - 123. http://dx.doi.org/10.1177/002221940303600204

Bandura, A. (1986). *Social privates of thought and action: A social cognitive theory.* Englewood Cliffs, NJ: Prentice-Hall.

Bereiter, C., & Scardamalia, M. (1987). *The psychology of written composition.* Hillsdale, NJ: Lawrence Erlbaum.

Berninger, V. W., & Swanson, H. L. (1994). Modification of the Hayes and Flower model to explain beginning and developing writing. In E. Butterfield (Ed.), *Advances in cognition and educational practice. Children's writing: Toward a process theory of development of skilled writing* (Vol. 2, pp. 57–82). Greenwich, CT: JAI Press.

Bouwer, R., Béguin, A., Sanders, T., & Van den Bergh, H. (2015). Effect of genre on the generalizability of writing scores. Language Testing, 32(1), 83-100. doi: 10.1177/0265532214542994

Braaksma, M. A. H., Rijlaarsdam, G., & Van den Bergh, H. (2002). Observational learning and the effects of model-observer similarity. *Journal of Educational Psychology*, *94*, 405-415. http://dx.doi.org/10.1037//0022-0663.94

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*, 1–36. http://dx.doi.org/10.1207/s1532690Xci2201_1

Braaksma, M. A. H., Van den Bergh, H., Rijlaarsdam, G., & Couzijn, M. (2001). Effective learning activities in observation tasks when learning to write and read argumentative texts. *European Journal of Psychology of Education, 1,* 33-48. http://dx.doi.org/10.1007/BF03172993.

Breetvelt, I., Van den Bergh, H., & Rijlaarsdam, G. (1994). Relations between writing processes and text quality: When and how? *Cognition and Instruction, 12*(2), 103-123.

British Council (2015). *The State of English in Higher Education in Turkey. A Baseline Study*. Retrieved August, 2016, from

http://www.britishcouncil.org.tr/sites/default/files/he_baseline_study_book_web_-_son.pdf

Brunstein, J. C., & Glaser, C. (2011). Testing a path-analytic mediation model of how self-regulated writing strategies improve fourth graders' composition skills: A randomized controlled trial. *Journal of Educational Psychology*, *103*(4), 922–938. http://dx.doi.org/10.1037/a0024622

Clachar, A. (2000). Opposition and Accommodation: "An Examination of Turkish Teachers' Attitudes toward Western Approaches to the Teaching of Writing" *Research in the Teaching of English, 35*(1), 66-100. Retrieved from http://www.jstor.org/stable/40171507

De La Paz, S. & Graham (2002). Explicitly teaching strategies, skills, and knowledge: Writing instruction in middle school classrooms. *Journal of Educational Psychology, 94* (4), 687 – 698. http://dx.doi.org/10.1037//0022-0663.94.4.687

De La Paz, S. (2005). Teaching historical reasoning and argumentative writing in culturally and academically diverse middle school classrooms. *Journal of Educational Psychology*, *97*, 139–158. http://dx.doi.org/10.1037/0022-0663.97.2.139

De Groot, A. D. (1980). Over leerervaringen en leerdoelen [About learning experiences and teaching goals] Handboek voor de onderwijspraktijk, 10 (november), B.1-B. 18. Deventer, The Netherlands: Van Loghum Slaterus.

Dörnyei, Z. (2009). The 2010s. *Communicative Language Teaching in the 21st Century: The "Principled Communicative Approach". Perspectives*, *36*(2), 33-43.

Erkan, Y. D., & Saban, A. I. (2011). Writing performance relative to writing apprehension, self-efficacy in writing, and attitudes towards writing: A correlational study in Turkish tertiary-level EFL. *The Asian EFL Journal Quarterly*, *13*(1), 163-191.

Fidalgo, R., Torrance, M., & Garcia, J. N. (2008). The long-term effects of strategy focused writing instruction for grade six students. *Contemporary Educational Psychology*, *33*(4), 672–693. http://dx.doi.org/10.1016/j.cedpsych.2007.09.001

Fidalgo, R., Torrance, M., Rijlaarsdam, G., van den Bergh, H., & Álvarez, M. L. (2015). Strategy-focused writing instruction: Just observing and reflecting on a model benefits 6th grade students. *Contemporary Educational Psychology*, *41*, 37-50. http://dx.doi.org/10.1016/j.cedpsych.2014.11.004

Flower, L. S., & Hayes, J. R. (1977). Problem-solving strategies and the writing process. *College English*, *39*(4), 449-461.

Flower, L.,& Hayes, J. (1981). A cognitive process theory of writing. *College Composition and Communication*, *32*, 365 - 387. http://dx.doi.org/10.2307/356600

Grabe, W., & Stoller, F. L. (2001). *Teaching and researching reading.* London: Pearson Education. http://dx.doi.org/10.1017/S0267190504000030

Graham, S. (2006). Strategy instruction and the teaching of writing: A meta-analysis. In C. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (pp. 187 207). New York: Guilford.

Graham, S., & Harris, K. R. (2003). Students with learning disabilities and the process of writing: A meta-analysis of SRSD studies. In L. Swanson, K. R. Harris, & S. Graham (Eds.), *Handbook of research on learning disabilities* (pp. 383–402). New York: Guilford.

Graham, S., & Perin, D. (2007). A meta-analysis of writing instruction for adolescent students. *Journal of Educational Psychology*, *99*(3), 445-476. http://dx.doi.org/10.1037/0022-0663.99.3.445

Groenendijk, T., Janssen, T., Rijlaarsdam, G., & Van den Bergh, H. (2008). How do secondary school students write poetry? How creative writing processes relate to final products. *L1 Educational Studies in Language and Literature*, *8*(3), 57-80.

Groenendijk, T., Janssen, T., Rijlaarsdam, G., & Van den Bergh, H. (2013). Learning to be creative. The effects of observational learning on students' design products and processes. *Learning and Instruction*, *28*, 35-47.

http://dx.doi.org/10.1016/j.learninstruc.2013.05.001

Harris, K. R., & Graham, S. (1996). *Making the writing process work: Strategies for composition and self-regulation*. Cambridge, MA: Brookline.

Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. Guilford Press.

Hinkel, E. (2015). *Effective curriculum for teaching second language writing: Principles and techniques*. Routledge: New York.

Kirkgoz, Y. (2005). 'English language teaching in Turkey: Challenges for the 21st century', in G. Braine (Ed.), *Teaching English to the world: History, curriculum, and practice* (pp. 159-175). Mahwah: Lawrence Erlbaum Associates.

Kirkgoz, Y. (2007). English language teaching in Turkey: Policy changes and their implementations, *Regional Language Centre Journal, 38*(2), 216-228. http://dx.doi.org/10.1177/0033688207079696

Klein, J., & Taub, D. (2005). The effect of variations in handwriting and print on evaluation of student essays. *Assessing Writing, 10*(2), 134-148. http://dx.doi.org/10.1016/j.asw.2005.05.002

Leki, I., & Carson, J. (1997). Completely different worlds: EAP and the writing experiences of ESL students in university courses. *TESOL quarterly*, *31*(1), 39-69. http://dx.doi.org/10.2307/3587974

Malhotra, N. K. (1991). Mnemonics in marketing: A pedagogical tool. *Journal of the Academy of Marketing Science*, *19*(2), 141-149. http://dx.doi.org/10.1007/BF02726006

Manchón, R. (Ed.). (2009). *Writing in foreign language contexts: Learning, teaching, and research* (Vol. 43). Multilingual Matters.

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). http://dx.doi.org/10.17239/jowr-2015.07.02.03

Mason, L.H., Hickey Snyder, K., Sukhram, D.P., & Kedem, Y. (2006). Self-regulated strategy development for expository reading comprehension and informative writing: Effects for nine 4th grade students who struggle with learning. *Exceptional Children, 73,* 69–89. http://dx.doi.org/10.1177/001440290607300104

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 Co. Inc.

Mateos, M., & Solé, I. (2009). Synthesising information from various texts: A study of procedures and products at different educational levels. *European Journal of Psychology of Education*, *24*(4), 435-451. http://dx.doi.org/10.1007/BF03178760

McPherson, F. (2000). *The memory key*. New York, NY: Barnes & Noble Books.

Ozsevik, Z. (2010). The Use of Communicative Language Teaching (CLT): *Turkish EFL Teachers' Perceived Difficulties in Implementing CLT in Turkey.* (Doctoral dissertation). Retrieved from: http://hdl.handle.net/2142/16211

Pintrich, P. R., Smith, D. A. F., Garcia, T., & McKeachie, W. J. (1991). *A manual for the use of the motivated strategies for learning questionnaire (MSLQ).* Ann Arbor: University of Michigan, National Center for Research to Improve Postsecondary Teaching and Learning.

Plakans, L. (2009). The role of reading strategies in integrated L2 writing tasks. *Journal of English for Academic Purposes, 8*, 252–266.

Reber, A. S. (1976). Implicit learning of synthetic languages: The role of instructional set. *Journal of Experimental Psychology: Human Learning and Memory, 2*, 88-94.

Rijlaarsdam, G., Braaksma, M., Couzijn, M., Janssen, T., Kieft, M., & Broekkamp, H. (2005). Psychology and the teaching of writing in 60 minutes. In P. Tomlinson, J. Dockrell, & P. Winne (Eds.*), Pedagogy and the Teaching for learning. BJEP monograph, Series II, No. 3* (pp. 127e153). Leicester, UK: The British Psychological Society. http://dx.doi.org/10.1348/000709905X62156

Rijlaarsdam, G., & Couzijn, M. (2000) Writing and learning to write: a double challenge. In R. J. Simons, J. van der Linden, & T. Duffy. (Eds.), *New Learning*. (pp. 157-189). Dordrecht: Kluwer Academic Publishers.

Rijlaarsdam, G., Van den Bergh, H., & Van Steendam, E. (2015). Writing process theory: A functional Dynamic Approach. In: MacArthur C., Graham S., Fitzgerald J. (Eds.), *Handbook of Writing Research*, Chapter 4. New York: The Guilford Press, 57-71.

Sawyer, R. J., Graham, S., & Harris, K. R. (1992). Direct teaching, strategy instruction, and strategy instruction with explicit self-regulation—Effects on the composition skills and self-efficacy of students with learning-disabilities. *Journal of Educational Psychology, 84*(3), 340–352. http://dx.doi.org/10.1037//0022-0663.84.3.340

Schellings, G., & Van Hout-Wolters, B. (2011). Measuring strategy use with self-report instruments: theoretical and empirical considerations. *Metacognition and Learning*, *6*(2), 83-90. http://dx.doi.org/10.1007/s11409-011-9081-9

Schoonen, R. (2005). Generalizability of writing scores: An application of structural equation modeling. *Language Testing, 22(1)*, 1–30.

Schoonen, R., van Gelderen, A., de Glopper, K., Hulstijn, J., Simis, A., Snellings, P., & Stevenson, M. (2003). First language and second language writing: The role of linguistic knowledge, speed of processing, and metacognitive knowledge. *Language learning*, *53*(1), 165-202. http://dx.doi.org/10.1191/0265532205lt295oa

Schunk, D. H. (1996). Goal and self-evaluative influences during children's cognitive skill learning. *American Educational Research Journal, 33*, 359-382. http://dx.doi.org/10.3102/00028312033002359

Schunk, D. H., & Zimmerman, B. J. (1997). Social origins of self-regulatory competence. *Educational psychologist, 32*(4), 195-208.

http://dx.doi.org/10.1207/s15326985ep3204_1

Schunk, D. H, & Zimmerman, B. J. (Eds.). (1998). Self-regulated learning: *From teaching to self-reflective practice.* New York: The Guilford Press.

Spivey, N.N. (1997). *The constructivist metaphor.* New York: Academic Press.

Sonnenschein, S., & Whitehurst, G. J. (1984). Developing referential communication: a hierarchy of skills. *Child Development, 55*, 1936e1945. http://dx.doi.org/10.2307/1129940

TechSmith. (2016). [Camtasia ComputerSoftware]. http://www.techsmith.com/camtasia.html

Tillema, M. (2012). *Writing in first and second language: Empirical studies on text quality and writing processes*. Utrecht: LOT Dissertation Series.

Tillema, M., Van den Bergh, H., Rijlaarsdam, G., & Sanders, T. (2012). Quantifying the quality difference between L1 and L2 essays: A rating procedure with bilingual raters and L1 and L2 benchmark essays. *Language Testing, 30*(1), 71-97. http://dx.doi.org/10.1177/0265532212442647

Torrance, M., Fidalgo, R., & García, J. N. (2007). The teachability and effectiveness of cognitive self-regulation in sixth-grade writers. *Learning and instruction*, *17*(3), 265-285.http://dx.doi.org/10.1016/j.learninstruc.2007.02.003

Torrance, M., Thomas, G. V., & Robinson, E. J. (1999). Individual differences in the writing behaviour of undergraduate students. *British Journal of Educational Psychology*, *69*(2), 189-199. http://dx.doi.org/10.1348/000709999157662

Turkkollu, N. (1994). *The error hierarchy of Turkish EFL teachers in their assessment of compositions* (Master's Thesis, Bilkent University). Retrieved from: http://hdl.handle.net/11693/17571

Van Steendam, E., Rijlaarsdam, G., Sercu, L., & Van den Bergh, H. (2010). The effect of instruction type and dyadic or individual emulation on the quality of higher-order peer feedback in EFL. *Learning and Instruction*, *20*(4), 316-327. http://dx.doi.org/10.1016/j.learninstruc.2009.08.009

Veenman, M. V. J. (2011). Learning to self-monitor and self-regulate. In R. Mayer & P. Alexander (Eds.), *Handbook of research on learning and instruction* (pp. 197–218). New York: Routledge.

Worthen, J. B., & Hunt, R. R. (2011). *Mnemonology: Mnemonics for the 21st century*. Psychology Press.

Yavuz, D. & Genc, A.B. (1998). *Flexibility of setting up a writing unit at YADIM.* Unpublished Action Research Study. Çukurova University, Adana, Turkey.

Zimmerman, B. J. (1977). Modeling. In H. Hom and P. Robinson (Eds.), *Psychological processes in early education* (pp. 37-70). New York: Academic.

Zimmerman, B. J. (1995). Self-efficacy and educational development. In A. Bandura (Ed.), *Self-efficacy in changing societies* (pp. 202-231). New York: Cambridge University Press. http://dx.doi.org/10.1017/CBO9780511527692.009

Zimmerman, B. J., & Kitsantas, A. (2002). Acquiring writing revision and self-regulatory skill through observation and emulation. *Journal of Educational Psychology, 94*(4), 660–668. http://dx.doi.org/doi:10.1037//0022-0663.94.4.660