Learning history by composing synthesis texts: Effects of an instructional programme on learning, reading and writing processes, and text quality

Isabel Martínez*, Mar Mateos*, Elena Martín* and Gert Rijlaarsdam°

*Autónoma University of Madrid | Spain
° University of Amsterdam | The Netherlands

Abstract: The combination of reading and writing can be a powerful tool for learning content. Previous research has shown the learning potential of the integrated use of these processes in what have come to be known as “hybrid tasks”. These tasks do not call for reading-then-writing, but for an interplay between being a reader and being a writer. The fact that students alternate between reader and writer roles explains the epistemic potential of such tasks.

The aim of the present study was to improve learning from texts via strategies that train students how to process synthesis texts. Processing such texts requires goal-oriented interaction between reading and writing activities. The participants were 62 sixth-grade students, 33 in the experimental and 29 in the control group. In a pretest-posttest design—with a control group—the effects of an experimental programme were tested on (a) the level of learning achieved, (b) the quality of the written texts produced, and (c) the synthesis text-processing activities (in a sub-sample of 32 participants). The experimental group was trained in the processes involved in writing a synthesis using two expository texts about history via a strategy-oriented programme, while the control group worked on the same content using the more conventional tasks in their regular textbook. Findings show that the experimental group outperformed the control group on a deep-learning content measure, wrote better texts, and exhibited more sophisticated text-processing activities.

Keywords: Synthesis text; text quality; writing processes; reading processes; strategy training.
1. Introduction

Today’s society offers students easy access to a vast amount of information provided from different sources, which often provide contradictory information. Students need to manage this load of incoming facts critically. Therefore, one of the biggest challenges for education is to teach students to initiate and direct their own learning, and to independently make decisions directed toward a self-determined learning goal (Goldman, 1997; Martín & Moreno, 2008; Mateos, 2001; Swartz, Costa, Beyer, Reagan, & Kallick, 2008).

Within this context, reading and writing processes are learning activities that can promote critical thinking. Given their epistemic function, these activities can potentially lead to transformation of knowledge (Bereiter & Scardamalia, 1987; Olson, 1994) guided by metacognitive control, or in other words, reflection on the activated process. However, in practice, research at different educational levels has shown that reading and writing are often primarily used to reproduce knowledge (Castelló, 1999; Goldman, 1997; Langer & Applebee, 1987; Mateos, Villalón, de Dios, & Martín, 2007; Solé et al., 2005).

Research on reading and writing has shown that integrated use can increase their epistemic potential (Moran & Billen, 2014; Tierney & Shanahan, 1996; Tynjälä, 2001). Spivey (1997) considers tasks that require both activities “hybrid” tasks, since they do not call for reading-then-writing, but for a complex interplay between being a reader and writer. The fact that students alternate between reader and writer roles explains the learning potential of such tasks: it requires an internal dialogue, and is a recursive process (Anmarkrud, Bråten & Strømsø, 2014; Langer & Flihan, 2000; O’Hara, Taylor, Newman & Sellen, 2002). Such an internal dialogue plays a role in good reading (Morrow, 1997) as well as good writing (Klein, 2014). When the task set requires reading for writing, the internal dialogue between the reader and the writer role usually starts in a natural way. However, not all hybrid tasks promote equally complex processes and the same levels of learning. Hence, for example, making a synthesis from multiple texts is cognitively more demanding than writing a summary of a single text. A summary of a single text could be created simply in accordance with a “knowledge-telling model”. However, a synthesis created from several source texts entails comparing, contrasting, and organising the ideas from the source texts around a central one, and integrating them with ones prior knowledge into a single new text (Mateos et al., 2014; Segev-Miller, 2004). Furthermore, according to Spivey (1997), in a synthesis task, three interrelated sub-processes – selection, organising and connecting - contribute to the epistemic potential. The learner must guide and regulate these sub-processes, which are goal-directed and occur recursively. In the selection process, the learners determine the degree of importance of the content of the information in the sources and include the ideas they regard as most important in their own products. During the organising process, learners look for keys in the sources that would enable
them to link the ideas. When they put the ideas in writing, they create their own groupings and sequences of ideas, including newly generated categories. During the connecting process, they link up the various pieces of information from the source texts with each other and with their prior knowledge, which can lead them to transform the contents. All in all, writing a synthesis text, based on multiple sources, requires three main high level processes: (1) the sources must be read integratively, that implies that readers must alternate between texts and build up semantic relations between these texts, they must compare and contrast, and abstract (Perfetti, Rouet & Britt, 1999); (2) during the production of the written text, the author must constantly alternate between the source texts and the developing synthesis text, executing activities such as underlining the source texts, taking notes, making a rough draft, and, (3) the author must constantly monitor and review the text already written. To a large extent the way students behave during the performance of such tasks is guided by their task representation. Flower et al. (1990) found that a simplified representation leads students to select certain information and include it in their text following the logic of the sources. They do not seek to integrate the ideas from the sources by means of a new structure. Such a simplified task representation is visible in the resulting texts. Mateos and Solé (2009) discuss two forms of insufficient products: a text which juxtaposed summaries of the sources or a single text that combines ideas from each source by alternating, but not integrating, them. Most of the existing research on these synthesis processes has been carried out with students at the higher levels of education (Flower et al., 1990; Gil, Bråten, Vidal-Abarca, & Strømso, 2010; Mateos & Solé, 2009; McGinley, 1992; Segev-Miller, 2004; Spivey, 1984). Those with younger children are less frequent (Lenski & Johns, 1997; Spivey & King, 1989). All these studies have found that synthesis tasks are difficult for many students at any level. Spivey and King (1989) found, for 6th, 8th and 10th grade students, that more advanced reader-writers make plans before writing, think about what and how they are going to write, and spend more time composing a synthesis. In contrast, students from the lower grades do not make plans and spend less time on the task. Lenski and Johns (1997) analysed the reading-to-write strategies of six 8th grade school students producing a written report based on source texts in a multiple case study. They observed three different patterns of searching, reading and writing – sequential, spiral and recursive – depending on the order in which the strategies appeared (Lenski & Johns, 1997, p. 25). Only one out of six participants in this study showed a recursive pattern, whereas the majority used a spiral pattern – defined as search-read-write –, repeating this pattern until the product was finished. The authors also found that the syntheses based on a spiral pattern were, by and large, copies or paraphrases of the source texts, while the participant who displayed a recursive pattern was the only one who produced an integrated synthesis. Mateos, Martín, Villalón and Luna (2008) report similar results with nine 15-year-old participants. All but one of the nine texts were unorganised, included few main ideas from the source texts, and were mere copies or paraphrases of these texts. Moreover, the participants rarely integrated information from within one source or from more than
one source. They made very few changes while rereading the text they had produced so far. Even when they did make changes, they did not reorganise what they had written. For the most part, the students did not take notes or make rough drafts. Mateos and her colleagues found clear relations between processes and texts: students who showed a relatively more recursive and flexible pattern of reading and writing activities produced the best syntheses. This implies that the activation of the sub-processes is not a random process. The distribution of activities across the process correlates with the quality of the text (Flower et al., 1990; Rijlaarsdam & van den Bergh, 1996; 2006). It seems that successful knowledge-construction through synthesis-writing based on multiple sources involves a recursive pattern of activities that is guided and controlled by metacognition.

In the studies on the process of creating a synthesis, content-learning was inferred from the processes employed and the quality of the texts written, rather than directly measured. However a recent study by Solé, Miras, Castells, Espino, and Minguela (2013) with 15-year-old students did include the level of learning as an outcome measure. Using a test to evaluate the degree of content reading comprehension, they measured the participants' ability (1) to retrieve information from the sources, which corresponds to low-level learning and, (2) to interpret and integrate ideas from the sources (one or more texts), which corresponds to high-level learning. Additionally, the written synthesis texts were evaluated in terms of their degree of organisation - coherence and cohesion -, selection, elaboration and integration. The processes were analysed on two dimensions: a linear/recursive and a direct/mediated dimension. The linear/recursive dimension was operationalized as either following a rigid sequence of steps or, on the contrary, using different sequences of activities depending on the task situation. The direct/mediated dimension was characterized by the absence/presence of specific activities, such as rereading or creating a rough draft, to elaborate the final synthesis. In this study relations were found between learning success, text quality and activity patterns created from individual audio-video registrations during the tasks. Therefore, the more recursive and the more mediated the process, the better the resulting text was and the higher the learning success score.

As has been shown, research suggests that students from different levels experience difficulties. They fail when: (1) deploying the complex processes involved in writing a synthesis of multiple texts to construct knowledge, and (2) activating these processes recursively. One of the approaches to tackle these difficulties is to provide instructional supports from the earliest educational level for knowledge integration via reading and writing. This may enhance their knowledge-transforming potential.

A small number of studies has examined the effectiveness of interventions aimed at the acquisition of strategies to create syntheses at higher levels of education. Boscolo, Arfé, and Quarisa (2007) designed and tested an intervention in which teachers revised the synthesis texts composed by undergraduate students. The authors claim that the participants improved their synthesising skill thanks to the intervention and their active
participation in this intervention by reflecting on their own experiences and discussing the different examples, although no control group was used in this study. In similar vein, Segev-Miller (2004) trained college students to synthesize through explicit instruction of relevant strategies (presentation and explanation, demonstration, and practice). She asked them to create a process log of the performance of two synthesis tasks, and to evaluate and discuss the differences between the processes and products of the two logs. From the content analysis of the process log and the products, Segev-Miller concluded that a significant improvement in the post-intervention synthesis occurred.

Few intervention studies have been conducted at beginner levels. Kirkpatrick and Klein (2009) sought to increase 7th and 8th grade students’ knowledge of the compare-contrast text structure and teach the students to plan the structure of their texts before writing. These authors found a positive intervention effect in terms of the holistic and structural quality of the texts produced. Wray and Lewis (1997) designed an intervention for teaching upper primary school students to read and write information texts combining a variety of teaching methods to achieve strategic learning of reading and writing as learning tools. The programme was tested in various studies and the results obtained were positive in all cases - the participants made strategic use of expository texts in order to learn the content.

In summary, some of the interventions concentrating on improving integrated use of reading and writing to construct knowledge indicate the effectiveness of promoting strategic learning in students. However, documented interventions have focused on improving reading and writing as learning tools, but did not directly measure content-learning results. Our view is that these processes ought to be taught in the context of domain-specific learning tasks to promote the learning in these domains and provide students with the opportunity to experience reading and writing as supporting specific content-learning (Bazerman, 1992; Bean, 2000; Vacca & Vacca, 1996). In the present study we embrace a comprehensive perspective on learning through reading and writing focusing on synthesis processes, products, and content-learning.

Our Strategies for Writing Syntheses to Learn (SWSL) programme has been designed to teach students of the final grade of primary education to use reading and writing as tools for learning in a particular domain. The program focuses on the integration of strategies involved when writing a synthesis text based on various source texts. The programme is based on strategy instruction in reading and writing (Graham & Harris, 2005; Graham, Harris & McKeown, 2013; Mateos, 2001; Sánchez, García & Rosales, 2010; Torrance, Fidalgo & García, 2007; Zimmerman, 2000), as well as the integrated use of both of them (Raphael & Englert, 1990; Wray & Lewis, 1997).

The main objective of the SWSL programme is that students learn to read and write by selecting, elaborating, connecting, and integrating the information coming from different texts with the goal of learning their content in a profound and constructive way. Therefore, for the construction of the intervention, we borrowed from writing and
reading strategy studies to set up a strategy oriented programme that provided instructions for the four main reading-writing activities mentioned. In this respect, the programme is a new branch in strategy-instruction research.

The main goal of our study was to design and test the effects of the Strategies for Writing Syntheses to Learn (SWSL) programme on (1) the degree of content learning, (2) the quality of the texts produced, and (3) the pattern of reading and writing activities.

2. Method

2.1 Design
A quasi-experimental pretest-posttest design with a control group was set up to investigate our research questions. The independent variable - including two different levels: yes/no implementation - was the instructional programme, which was designed to assist students to develop more epistemic processes of reading and writing. The dependent variables were participants’ content learning outcomes, task processes patterns - the reading and writing activities - and the quality of the generated syntheses. We measured prior knowledge of history and reading comprehension, so that we could investigate initial differences between groups assigned to different conditions on relevant variables, and to be able to check for generalization of the effects across learner characteristics.

2.2 Participants
Sixty two 6th grade primary school students participated in the study. They came from four classes of two schools, two whole classes from each school, which were randomly assigned to the experimental and control groups: 33 in the experimental group and 29 in the control group. There were 34 girls and 28 boys, with a mean age of 10 years and 8 months.

In addition we collected process measures from a subsample of 32 of these students. 16 students with low scores on the two pretest variables prior knowledge and reading comprehension (8 for the control group and 8 for the experimental group) and 16 with high scores (8 for each group) were selected from the full sample (the same number of students from each school). All the students were part of the whole class during the interventions; the only difference was that the students in the subsample were tested individually during the pre- and post-sessions.

Participants in this study did not receive any compensation for their participation. The schools were chosen on the basis of convenience. None of the participants was diagnosed with special needs.
2.3.1 Intervention programme

The program consisted of 12 sessions of 60 minutes each, three per week, grouped in three blocks of four lessons, each block with one learning theme, two source texts and one synthesis task as the main ingredient.

The learning contents were three units from the regular history curriculum at these schools: Industrial Revolution, Capitalism, and Rural/Urban life in 19th century. See Table 1 for an overview of the intervention programme.

For each of the three learning tasks in the experimental programme we taught the same five strategies: (1) selecting important ideas from the source texts, (2) elaborating on the information, (3) organising the content, (4) integrating prior knowledge with new knowledge, and (5) integrating information from both source texts. The strategies were taught in an integrated and recursive manner – sometimes it was done simultaneously, we went back over things when necessary, etc. Based on other studies (Wray & Lewis, 1997), the strategies were taught using a variety of methodologies, such as: (1) teacher modelling, (2) collaborative activities, (3) guided activities, (4) individual student activities, and (5) the support of a written guide (for details see Table 1). The purpose was to gradually transfer control from the instructor to the students as this is advocated and tested in strategy-oriented research with a strong emphasis on self-regulation (Fidalgo, García, Torrance, & Robledo, 2009; Graham & Harris, 2005; Graham, Harris & McKeown, 2013; Zimmerman, 2000). A fundamental part of this programme, following previous work by Gárate and Melero (2004), was the construction of a written guide to support the process. The goal was that students benefit from this program now, as well as in the future when they are able to work on their own. To encourage students to self-regulate their learning process, we jointly develop a guide with them in which all aspects were posed in the form of a question. Some examples of phrases that were proposed in this material are: "What are our aims?", "What for?", "How can I make a map or scheme linking the thoughts in the texts?", etc… It is important to stress that the guide is not intended to be a fixed sequence of steps. Therefore, the intention was not to offer a finished version of this guide to the students from the start, but to gradually work on each of the steps in the sessions and recapitulate everything at the end and write it up into a guide. Students could then rely on this material when they perform the following synthesis tasks.

In the first block the instructional focus was on modelling and observational learning. In these sessions the instructor guided the students through the sequence to be followed. In the second block the focus moved to collaborative learning and writing in groups of four with instructor support. Finally, in the third block, students worked individually, with the support of the instructor and the written guide.
Table 1: Description of the aims, strategies, sessions, social organisation, and material of the SWSL instructional programme

<table>
<thead>
<tr>
<th>General aim</th>
<th>Specific aims</th>
<th>Strategies and techniques</th>
<th>Session, social organisation and material</th>
</tr>
</thead>
</table>
| To perform a first synthesis task -- composing a text based on two expository source texts -- through researcher modelling of the processes employed | Synthesis task representation  
Prior knowledge activation  
Global/local comprehension of texts  
Selection of main ideas  
Link between prior knowledge and text information  
Organisation of ideas from texts  
Elaboration of the information from texts linked to prior knowledge and aims  
Integration of prior knowledge with new knowledge  
Integration of the information within and between the two texts  
Writing of final text  
Elaboration of a student-based guideline to help students to control their own processes (to be used in all sessions) | Joint reflection  
Negotiation of aims  
Modelling  
Guided collaborative activity  
Guided questions | Sessions 1-4  
Whole class  
Pair of texts about the industrial revolution |
| To perform a second synthesis task in which the students have a little more autonomy than in the previous one through group | Prior knowledge activation  
Global/local comprehension of texts  
Selection of main ideas  
Link between prior knowledge and text information | Guided collaborative activity in small groups and with the whole class  
Joint reflection with whole | Sessions 5-8  
Whole class and small groups (4 students)  
Pair of texts about social |
To perform a third synthesis task in which the students do the activities individually, with aids:

<table>
<thead>
<tr>
<th>Work</th>
<th>Class</th>
<th>And political organisation under capitalism</th>
</tr>
</thead>
<tbody>
<tr>
<td>Organisation of ideas from texts</td>
<td>Individual activity</td>
<td></td>
</tr>
<tr>
<td>Elaboration of the information from texts linked to prior knowledge and aims</td>
<td>Global/local comprehension of texts</td>
<td></td>
</tr>
<tr>
<td>Integration of prior knowledge with new knowledge</td>
<td>Selection of main ideas</td>
<td></td>
</tr>
<tr>
<td>Integration of the information within and between the two texts</td>
<td>Link between prior knowledge and text information</td>
<td></td>
</tr>
<tr>
<td>Writing of final text</td>
<td>Organisation of ideas from texts</td>
<td></td>
</tr>
<tr>
<td>Elaboration of the information from texts linked to prior knowledge and aims</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Integration of prior knowledge with new knowledge</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Integration of the information within and between the two texts</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Writing of final text</td>
<td>Elaboration of a written guide of the process</td>
<td></td>
</tr>
</tbody>
</table>

Sessions 9-12
Whole class and individual
Pair of texts about rural and urban life in the 19th century
2.3.2 Assessment of reading comprehension

Reading scores were based on a 35 multiple choice-item reading comprehension test ($\alpha= .60$) produced and validated by IDEA ("Instituto de Evaluación y Asesoramiento Educativo": Educational Assessment and Counselling Institute; González Nieto, 2002). Students read different texts (story, instructive text, expositive text, and newspaper article) and answered questions focusing on text structure identification, global comprehension, linguistic aspects, and comprehension strategies.

2.3.3 Assessment of content-learning

The topics of the three syntheses tasks have not yet been taught by the teacher in the class group. A test was constructed to assess the pupils’ knowledge about the history contents taught during the study, before, and after implementation of the programme (Martínez, 2012). Therefore, this instrument served two different purposes. First, before the intervention, the test allowed us to assess the students’ prior knowledge about each of the topics taught in the intervention, to check differences between the conditions. Second, the prior knowledge scores were used to evaluate the degree of learning achieved by the participants and the effects of the intervention in this respect. The items covered the formal history contents programme which the three units were part of. The test comprised 17 true/false questions and five complete-this-sentence items in which the first part of a sentence was given and students had to choose the appropriate second part to complete it. All items had the same weight (1 point for a correct answer). The test-retest reliability index of this questionnaire was reasonable ($\alpha = .62$).

Following earlier models which focused on comprehension and learning from one text (Kintsch, 1998) and from multiple texts (Perfetti et al., 1999), this test intended to measure two levels of learning/comprehension. The students were asked to decide in each case whether the idea expressed could or could not be deduced from the information included in the texts they read. On the one hand, six statements were paraphrases of ideas contained in specific fragments of the texts and required minimum inferences -low-level learning- (e.g. the capitalism system was on the side of the working class –true/false-. This information was literally provided in one source text). On the other hand, 16 items were inferences extracted from the information given in the texts and/or required the integration of information from distant information -high-level learning- (e.g. "the steam engine had an impact on both the industry and transport sector” –true/false-. In this case the impact of the steam engine on the industry was described in one source text and its impact on transport in the other). The correct responses, therefore, are produced when identified student has gathered that the true statements can be deduced from the texts, and that the false statements cannot be deduced from the texts. This division of test items was validated by two experts, independently assessing the items in the two categories. The reliability indexes of these
parts of the questionnaire were .60 and .70 (pretest) and .58 and .62 (posttest) for the low- and high-levels respectively.

2.3.4 Texts for the synthesis tasks and tests
Five pairs of source texts were selected for the five synthesis tasks the students had to write in total: two for the tests and three for the intervention. The content of the five pairs of texts, selected in agreement with the students’ teachers, was related to various aspects of a particular teaching unit on contemporary history which had not yet been taught. The texts were selected from various textbooks by different publishers and care was taken to ensure that the students were familiar with their format and level. The participants’ regular teachers assessed the difficulty of these sources and considered the texts to be of an appropriate level for their students.

The texts in each pair included complementary information on different topics in history. They were of a similar length (with a mean of 256 words, ranging from 235 to 280). From these five topics, we randomly selected those for testing and those for the intervention.

2.4 Procedure
Pretest procedure. In session 1 (80 minutes), reading comprehension and topic knowledge questionnaires were administrated.

In session 2, all students were given 60 minutes to perform the first (pre-test) synthesis task: to write a synthesis based on two source texts individually without any help. The following instructions were given orally and in writing: Here are two social sciences texts. Read them carefully as many times as you need and write a synthesis of the information in them. You can take notes or make a rough draft. The selected subsample of 32 participants was individually video-recorded while performing the synthesis task, with the camera pointing at the material on the desk used by the students.

Instruction sessions. The instructional programme was implemented with the whole experimental group between session 3 and session 14. The researcher implemented the intervention programme in students’ regular classrooms, and with their teacher present.

The control group performed the same three synthesis tasks, but they were taught by their regular teachers, according to their conventional method, in the presence of the researcher. This implies that students studied the source texts individually via all kinds of learning activities included in the history textbook to build up their comprehension. Individual activities were, for instance, answering written questions about the texts, or completing sentences, etc. with the teacher’s support. Then, in the fourth lesson of each block, to equalize test conditions, the researcher presented and explained the synthesis task that rounded off the learning content acquired in the three former lessons.
In sum, the SWSL programme was implemented by a researcher in the experimental group with the regular teacher present, while the control group worked with their usual teacher and in presence of the programme instructor. In both conditions the same topics and synthesis tasks were instructed, but in a different way. The structure of the intervention was the same in both conditions. Both groups (experimental and control) had equivalent exposure to the content of the texts using for synthesis tasks and received instruction related to the learning test items. Moreover, variables such as a time on task, motivation, absenteeism, etc. were controlled to avoid differences between groups caused by different instructors.

Posttest. The students in both groups performed the post-test synthesis task in session 15 following the same procedure as in the pre-test. They also took the topic knowledge questionnaire again. A summary of the procedure followed by the two groups is presented in Table 2.

<table>
<thead>
<tr>
<th>Sessions</th>
<th>Experimental group</th>
<th>Control group</th>
<th>Topic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Session 1</td>
<td>Reading comprehension and Previous Knowledge tests</td>
<td>Reading comprehension and Previous Knowledge tests</td>
<td>Liberalism and Absolutism (Political systems)</td>
</tr>
<tr>
<td>Session 2*</td>
<td>Performance of the pre-test synthesis task individually without any help</td>
<td>Performance of the pre-test synthesis task individually without any help</td>
<td>Liberalism and Absolutism (Political systems)</td>
</tr>
<tr>
<td>Sessions 3-6</td>
<td>Performance of the first synthesis task through the strategies of modeling and observational learning</td>
<td>Work with texts and activities from the textbook (conventional method)</td>
<td>Industrial and Agricultural Revolution</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Performance of the first synthesis task individually and without help</td>
<td></td>
</tr>
<tr>
<td>Sessions 7-10</td>
<td>Performance of the second synthesis task through the strategy of collaborative writing with instructor support</td>
<td>Work with texts and activities from the textbook (conventional method)</td>
<td>Capitalism</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Performance of the second synthesis task individually and without help</td>
<td></td>
</tr>
</tbody>
</table>
Sessions 11-14 | Performance of the third synthesis task through the strategies of individual practice supported by the instructor and the written guide

Work with texts and activities from the textbook (conventional method)

Urban and rural lifestyle in 19th century

Performance of the third synthesis task individually and without help

Session 15* | Performance of the post-test synthesis task individually without any help. Topic knowledge test

Performance of the post-test synthesis task individually without any help. Topic knowledge test

Liberalism and Absolutism (social life)

* The subsample was individually video-recorded in session 2 and 15

Therefore, it should be clear that during the study the two participant groups read the same texts, worked on the same history topics and the same synthesis tasks before, during, and after the intervention (traditional or SWSL). The only difference between the conditions was the instruction on the three synthesis tasks in the instructional units: the experimental group worked on them following the SWSL programme methodologies (modelling, joint activity, written guide support…) whereas the control group wrote them individually, and without help, after working on the topics through the traditional method.

2.4.1 Rating procedure

Content learning
Two scores were calculated for content learning (high- and low-level learning). These scores were derived from the sum of the correct answers in the test and the weighted subtraction of the incorrect answers (Abad, Olea, Ponsoda & García, 2011).

Products
The pre- and post-synthesis texts were rated, based on our previous studies, according to the degree of selection, elaboration, intra-textual integration, intertextual integration, and the title complexity (Martínez, Martín & Mateos, 2011; Mateos et al., 2008; Mateos & Solé, 2009). Two experts – blind to the student’s condition – scored the students on these variables according to four levels of complexity (see Table 3).
Table 3: The levels of quality for five text variables

<table>
<thead>
<tr>
<th>Selection</th>
<th>Does not select</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Includes important and unimportant ideas</td>
</tr>
<tr>
<td></td>
<td>Includes all or most important ideas, but also unimportant and repeated ideas</td>
</tr>
<tr>
<td></td>
<td>Includes all the important ideas and none, or only one, unimportant or repeated idea</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Elaboration</th>
<th>Copies</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Copies and paraphrases</td>
</tr>
<tr>
<td></td>
<td>Copies, paraphrases and some unimportant elaboration</td>
</tr>
<tr>
<td></td>
<td>Copies, paraphrases and some important elaboration</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Intratextual integration</th>
<th>List of unconnected ideas</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>List of ideas with connectors copied from the source texts</td>
</tr>
<tr>
<td></td>
<td>At least one attempt at connection</td>
</tr>
<tr>
<td></td>
<td>Connected ideas</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Intertextual integration</th>
<th>Two separate texts</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Two juxtaposed summaries</td>
</tr>
<tr>
<td></td>
<td>Integration of both texts with one idea</td>
</tr>
<tr>
<td></td>
<td>Integration of several ideas from both texts</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Title</th>
<th>No title or summary</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Copied title</td>
</tr>
<tr>
<td></td>
<td>Sum of titles</td>
</tr>
<tr>
<td></td>
<td>Integration of titles with new information added</td>
</tr>
</tbody>
</table>

To check whether the set of criteria referred to a common construct (the overall quality of the synthesis), we performed a reliability test (Cronbach, 1951). The alpha values were .60 for the pre-intervention task and .89 for the post-intervention task criteria scores. Since the correlations between all five criteria were positive and significant, we decided to calculate the overall mean of the scores on the five criteria.

The two raters scored all the 124 synthesis texts – from pre and post-test – on the overall quality criterion defined. The kappa index values were significant (p < .05) and high: .70 for the pre-intervention task and .71 for the post-intervention task.

**Cognitive activities carried out by the students**

To code the activities observed in the audio-video recordings (see procedure details in section 2.4), we created a graphic pattern recording of each activity, with the start and end times. We also examined the source texts and each student’s written products. The category system was adapted from previous studies (Mateos & Solé, 2009; Solé, Miras, & Gràcia, 2005; Solé et al., 2013) – see Table 4.-
Table 4: Category system of the activities performed by students

<table>
<thead>
<tr>
<th>Category</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reads source text (source text 1, 2 or both)</td>
</tr>
<tr>
<td>Reads and underlines source text (source text 1, 2 or both)</td>
</tr>
<tr>
<td>Reads and takes notes on source text (source text 1, 2 or both)</td>
</tr>
<tr>
<td>Refers to synthesis guidelines/material</td>
</tr>
<tr>
<td>Makes rough draft while referring to source text (source text 1, 2 or both)</td>
</tr>
<tr>
<td>Writes final text while referring to source texts (source text 1, 2 or both)</td>
</tr>
<tr>
<td>Writes final text while referring to rough draft</td>
</tr>
<tr>
<td>Revises rough draft and makes changes</td>
</tr>
<tr>
<td>Revises final text without making any changes</td>
</tr>
<tr>
<td>Revises final text and makes changes</td>
</tr>
<tr>
<td>Writes final text without referring to source texts or rough draft</td>
</tr>
<tr>
<td>Makes rough draft without referring to source texts</td>
</tr>
</tbody>
</table>

Two examples of the graphic patterns, in figures 1 and 2, illustrate the protocols we produced. Figure 1 includes the different colours and textures used to represent all the activities observed during the task.

<table>
<thead>
<tr>
<th>Reads</th>
<th>Reads and underlines source text</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Refers to synthesis guidelines/material</td>
</tr>
<tr>
<td>Reads and takes notes on source text</td>
<td>Writes final text while referring to source text</td>
</tr>
<tr>
<td>Makes rough draft while referring to source text</td>
<td>Revise final text without making any changes</td>
</tr>
<tr>
<td>Writes final text while referring to rough draft</td>
<td>Revises final text and makes changes</td>
</tr>
<tr>
<td>Revises rough draft and make changes</td>
<td></td>
</tr>
<tr>
<td>Makes rough draft without referring to source texts</td>
<td>Writes final text without referring to source texts or rough draft</td>
</tr>
</tbody>
</table>

*Figure 1.* Explanation of the representation of each reading and writing activity observed.
Figure 2 presents the protocol of the pre- and post-tasks of a control group participant. This participant followed the same pattern in the two tasks, spending five minutes less in the post-test than in the pre-test. Figure 3 shows a protocol of an experimental group participant. Here we see many differences between the pre- and post-tasks. Compared with the pre-test, the participant spent more time on the post-test, took notes on the source texts, looked at the strategy guide several times, reread the source texts at different moments, and revised during the process and at the end.

From these graphic patterns we created two indicators that allowed us to analyse the different process characteristics the students exploited when performing the task.
1. The time each student spent on the task as a whole;
2. The number of different reading, writing and revision activities employed.

In addition, based on previous studies (Martínez et al., 2012; Solé et al., 2013) we defined three other variables (see Table 5).

Two experts – blind to the student’s condition – scored 20% of the protocols for these three variables. The Cohen’s kappa index values were significant (p < .05) and high: .76 for the interactions between source texts and student’s text, .80 for the review variable, and .84 for the first reading of the source texts. Then one of the experts, the first author of this paper scored the rest of the students for these variables.
Table 5: The different levels description of procedure variables

<table>
<thead>
<tr>
<th>Interactions</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Interactions</strong></td>
<td>Moves backwards and forwards between the source texts and the student’s own text in short periods (less than 30 seconds)</td>
</tr>
<tr>
<td></td>
<td>Moves backwards and forwards between the source texts and the student’s own text in long periods (more than 30 seconds)</td>
</tr>
<tr>
<td></td>
<td>In addition to level 2, adds a longer rereading that breaks the backwards and forwards sequence and/or occurs during revision</td>
</tr>
<tr>
<td><strong>Revision</strong></td>
<td>Does not revise or rereads final text only for a few seconds</td>
</tr>
<tr>
<td></td>
<td>Revises and makes small changes</td>
</tr>
<tr>
<td></td>
<td>Revises and makes substantial changes</td>
</tr>
<tr>
<td><strong>First reading of source texts</strong></td>
<td>Parallel reading</td>
</tr>
<tr>
<td></td>
<td>Serial reading</td>
</tr>
</tbody>
</table>

3. **Results**

First, ANOVAs and Chi-squared tests were used to check that there were no differences between the control group and the experimental group on any of the variables in the pre-test. Second, repeated-measures ANOVAs with two factors – a between-subjects (group) factor and a within-subjects (time) factor – were run for all variables, except the first reading of the source texts criterion – including two different levels –, which was analysed using the Chi-square test.

3.1 **Initial equivalence between the two groups**

We found no initial differences between the control group and the experimental group and between girls and boys inside the groups for the seven variables tested (Table 6).

Table 6: Descriptive statistics for each of the groups on the initially measured variables

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Control M</th>
<th>Control SD</th>
<th>Experimental M</th>
<th>Experimental SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reading comprehension</td>
<td>62</td>
<td>17.17</td>
<td>3.22</td>
<td>17.67</td>
<td>4.83</td>
</tr>
<tr>
<td>Low-level prior knowledge</td>
<td>62</td>
<td>1.57</td>
<td>1.96</td>
<td>1.52</td>
<td>2.57</td>
</tr>
<tr>
<td>High-level prior knowledge</td>
<td>62</td>
<td>2.55</td>
<td>3.48</td>
<td>1.76</td>
<td>2.95</td>
</tr>
<tr>
<td>Time (mins.)</td>
<td>32</td>
<td>36.13</td>
<td>5.57</td>
<td>37.81</td>
<td>10.18</td>
</tr>
<tr>
<td>Number of activities</td>
<td>32</td>
<td>3.38</td>
<td>1.26</td>
<td>3.31</td>
<td>1.25</td>
</tr>
<tr>
<td>Interactions (1-3)</td>
<td>32</td>
<td>1.19</td>
<td>0.54</td>
<td>1.25</td>
<td>0.48</td>
</tr>
<tr>
<td>Revision (1-3)</td>
<td>32</td>
<td>1.25</td>
<td>0.58</td>
<td>1.50</td>
<td>0.82</td>
</tr>
<tr>
<td>Overall quality (1-4)</td>
<td>62</td>
<td>1.70</td>
<td>0.43</td>
<td>1.68</td>
<td>0.50</td>
</tr>
</tbody>
</table>
3.2 Effects of the intervention on content-learning

Table 7 shows the means and standard deviations for each of the groups at different moments of the study on the content-learning criterion. As mentioned before, the scores were the result of the sum of the right answers minus the wrong ones.

Table 7: Means and standard deviations for each of the groups in the pre and post task on the content learning criterion

<table>
<thead>
<tr>
<th>N</th>
<th>Control</th>
<th>Experimental</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PRE</td>
<td>M</td>
</tr>
<tr>
<td>Low-level learning (max. 6)</td>
<td>62</td>
<td>1.57</td>
</tr>
<tr>
<td>High-level learning (max. 16)</td>
<td>62</td>
<td>2.55</td>
</tr>
</tbody>
</table>

3.2.1 Low-level learning

The interaction between time (pre-test and post-test) and group (control and experimental) was not significant ($F(1, 60) = .15, p > .05, \eta^2_p = .002$). However, a significant time factor was found ($F(1, 60) = 30.29, p < .001, \eta^2_p = .34$); both groups' scores improved between the pre-test and post-test tasks.

3.2.2 High-level learning

Interaction between the within-time and between-subject factors was significant ($F(1, 60) = 9.48, p < .05, \eta^2_p = .14$); the two group's scores varied differently over time. With regard to simple effects, the experimental group obtained higher scores on the post-test task ($F(1, 60) = 8.98, p < .01, \eta^2_p = .13$). There were no significant changes in the control group's scores between the two tasks, whereas those of the experimental group did change over time ($F(1, 32) = 57.83, p < .001, \eta^2_p = .64$) - see Figure 4-.
3.3 Effects of intervention on product quality.

Table 8 shows the means and standard deviations for each of the groups at different moments of the study on the overall product quality criterion.

<table>
<thead>
<tr>
<th>N</th>
<th>Control</th>
<th>Experimental</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PRE</td>
<td>POST</td>
</tr>
<tr>
<td>Overall quality</td>
<td>62</td>
<td>1.59</td>
</tr>
</tbody>
</table>

The interaction between time (pre- vs. post-intervention task) and group (experimental-control) was significant ($F (1, 60) = 13.15, p < .001, \eta^2_p = .63$). Looking at each group separately, we found no significant change between the two tasks in the control group, but did find differences in the experimental group ($F (1, 32) = 181.01, p < .001, \eta^2_p = .85$) – see Figure 5.

*Figure 4.* High-level learning score for each group on the pre and post tests.
3.4 Effects of the intervention on the processes

Table 9 shows the means and standard deviations of each group’s process variables scores on the pre- and post-intervention synthesis tasks they performed.

Table 9: Means and standard deviations for each of the groups in the pre and post task on procedure variables

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Control</th>
<th></th>
<th>Experimental</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PRE</td>
<td>POST</td>
<td></td>
<td>PRE</td>
<td>POST</td>
</tr>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>M</td>
<td>SD</td>
<td>M</td>
</tr>
</tbody>
</table>
| Time spent on task     | 32     | 36.13           | 5.57  | 28.13          | 7.75  | 37.81| 10.18
|                        |        |                 |       |                |       |      | 51.44| 8.99 |
| Number of different activities | 32    | 3.38            | 1.26  | 2.56           | 0.63  | 3.31 | 1.25 | 5.06 | 1.61 |
| Interactions between source texts - student text | 32    | 1.19            | 0.54  | 1.00           | 0.00  | 1.25 | 0.45 | 1.88 | 0.81 |
| Presence and nature of revision | 32    | 1.25            | 0.58  | 1.13           | 0.50  | 1.50 | 0.82 | 1.88 | 1.02 |
3.4.1 Processing Time
The interaction between time and group was significant ($F(1, 30) = 47.44, p < .001, \eta^2_p = .61$). The mean scores of the two groups on the post-intervention task differed ($F(1, 30) = 61.71, p < .001, \eta^2_p = .67$). Both groups changed over time but in different directions: the mean time spent on the task by the control group decreased ($F(1, 15) = 18.73, p < .01, \eta^2_p = .56$), whereas the mean time spent by the experimental group increased ($F(1, 15) = 28.82, p < .001, \eta^2_p = .66$).

3.4.2 Number of activities
Table 4 shows the type of activities carried out by the students during the synthesis task. The interaction between the number of activities performed and the learning condition was significant ($F(1, 30) = 33.22, p < .001, \eta^2_p = .53$). The control group performed fewer activities in the post-intervention task than in the pre-intervention task ($F(1, 15) = 9.64, p < .01, \eta^2_p = .39$), whereas the experimental group performed more activities in the post-intervention task than in the pre-intervention task ($F(1, 15) = 23.71, p < .001, \eta^2_p = .61$).

3.4.3 Interactions between the source texts and the text produced
With regard to the interactions between the source texts and the students’ texts, we observed a significant interaction between time and group ($F(1, 30) = 13.00, p < .01, \eta^2_p = .30$). Looking at the simple effect of each group separately, the control group’s performance showed no significant change between the two tasks, whereas an increase in the experimental group scores ($F(1, 15) = 12.10, p < .01, \eta^2_p = .45$) was observed.

3.4.4 Revision activities
The interaction between time and group with regard to revision was significant ($F(1, 30) = 6.32, p < .05, \eta^2_p = .17$) –see Table 9 for descriptive statistics-. When this interaction effect was broken down into simple effects, we found that the two groups’ scores on the post-intervention task differed significantly ($F(1, 30) = 6.92, p = .01, \eta^2_p = .19$), whereas their pre-test scores did not.

3.4.5 First reading of the source texts
Table 10 shows the frequencies for each type of change in the scores on this criterion. As mentioned in the rating procedure, this variable was composed by two alternatives: parallel/serial reading. To analyse it, the following were applied: (1) “no change” signified that the student read the texts in the same way in both tasks; parallel or serial reading; (2) “go up” described those cases in which the student moved from serial reading in the pre-task to parallel reading in the post-task; and (3) “go down” referred to the opposite change, from parallel to serial reading.
According to these results ($\chi^2 (2) = 9.55, p < .01$), and to the data in Table 10, the scores of the students in the experimental group on this variable went up to a greater extent than would have been expected, while those of the control group went down.

### 4. Discussion

The overall aim of this study was to assess the effect of a SWSL programme designed to improve learning via strategy training in processing texts through synthesis generation, which requires goal-oriented interaction between reading and writing activities. After the instruction, pupils participating in the experimental group attained a greater degree of high-level content-learning, that is they were better able to interpret and integrate distant information. This indicates that teaching the processes involved in producing syntheses (selection, elaboration, organisation, and integration) helped students to achieve learning that goes beyond mere knowledge reproduction and fosters knowledge transformation. Previous studies had shown that integrating information from more than one text into a single, new, and different product promotes learning (Miras et al., 2008). Therefore, the present study contributes to earlier research with the direct and integrated evaluation of the three essential aspects involved in synthesis performance: (1) the degree of content learning, (2) the quality of the products written, and (3) the pattern of reading and writing activities.

This study confirms that previous findings on how instruction helps develop older students' cognitive processes as they engage in reading and writing through synthesis (e.g., Martínez et al., 2011) can be adapted and transferred to younger students when appropriate scaffolding is in place.

This study also focused on the cognitive activities students perform in carrying out synthesis tasks to arrive at a final product. In line with other research (Lenski & Johns, 1997; Mateos et al., 2008; Martínez et al., 2011; Solé et al., 2013), the students in this study tended to follow rigid, linear patterns in the beginning. After the intervention, it was found that, generally speaking, the experimental group’s score changed significantly between the pre- and post-intervention task.
The participants in this group replicated the results obtained by Spivey and King (1989): they increased the time they spent on performing the task and also carried out a larger number of different activities. Even so, this increase in the time taken and the number of activities might not be sufficient for the successful performance of the task. We also found that the nature of these activities changed qualitatively: the duration of the interactions between the source texts and the students’ own texts increased –enabling probably the elaboration of the ideas included in the final text. Some students even returned to the source texts while revising their own products and students who started off by reading the texts separately later came to read them together. These results are in line with those obtained by Solé and her colleagues (2013) when they evaluated the function of the re-readings of the sources carried out by the participants in the process of creating a synthesis.

Nevertheless, there was an exception with regard to revision behaviour. Although there was a pronounced trend towards improvement in this activity, in line with results obtained by Torrance, Fidalgo, and García (2007), the result was not significant. This finding might be explained by the difficulty students at beginner levels have in revising their texts, particularly in picking up not only spelling and grammar mistakes, but also substantial errors of interpretation, organisation, etc. (Graham & Harris, 1996; 2000; Martínez et al., 2011; Mateos et al., 2008). It appears that, despite the fact that the students who received the programme went from not revising their texts to doing so, the intervention was not effective enough to enable them to revise with the aim of doing anything more than checking spelling and grammar. Another explanation, following Torrance et al. (2007), might be that students may have spent considerable time in planning what to say and then produced their best possible text in one go. In such cases, the revision process may not be as necessary to improve the quality of the text produced; the revision has already been done mentally during the process of reading and writing. In any case, it seems that students did not experience the need to revise their text when they had finished the draft. This might be due to the instructional setting. In fact, Mateos and Solé (2009) found that students tend to review their texts to a greater extent when the context demands they improve the text, for example because an audience will read it. However, the present study failed on this point. Students were instructed on the revision process, but they may not have felt the need to improve the text for a specific aim. Nevertheless, given the very large effect of the intervention on text quality (about 3 standard deviations) it is doubtful whether more revisions could have contributed much to the effect.

In view of these results, it can be said that, after the intervention, the students tackled the task following certain general patterns. The participants in the experimental group followed patterns displaying more recursive and flexible use of reading and writing: they spent more time on the task, performed more activities, revised more and substantially modified their texts, and returned repeatedly to the texts during long periods of time that would enable them to elaborate the information they read in order to include it in their own texts from the start in an integrated fashion.
The implementation of the intervention had a positive effect on the degree of learning achieved by the students and a positive influence on the end products and the activity pattern that followed. It seems, therefore, that the methodology employed in the sessions, based on the programme produced by Wray and Lewis (1997) –completed with other methodological support like the written guide-, was effective. Hence we conclude that this type of intervention has important implications for education, as it enables pupils to improve the use they make of reading and writing, leading them to be more recursive and flexible in their processes and, therefore, to write better texts (including relevant ideas, organised and integrated) and achieve deep learning of the content.

Previous studies found that students wrote syntheses of poor quality (Lenski & Johns, 1997; Mateos et al., 2008). Their texts failed to include important relevant information, were copies or paraphrases of information from the sources, had an incoherent structure, and failed to integrate information from the two source texts. As expected, as in the study by Martínez and her colleagues (2011), it was found that the two groups scored differently on the post-intervention task: whereas the experimental group improved their synthesis texts significantly, there was no significant improvement in the control group’s products. The experimental group improved with regard to selection, elaboration, coherence, and the integration of ideas from the two sources.

To conclude, we point to future research lines derived from this study’s indicators or limitations. First, regarding programme design, the intervention was effective when the methodologies applied were considered as a whole. In future, a new research design might be tested, in which it might also be useful to break down the programme into the instructional methods used. This would make it possible find out more about the contribution of each of the teaching methods to the improvements we found in this study. In addition, the intervention was carried out over four weeks, with a single teaching unit of a particular subject. Under these conditions, positive effects were found, so a new research question arises: can these findings be generalised to other units, subjects, kinds of texts (non-complementary), etc? Another issue for future research would be to focus on the teaching of the revision process in order to enable students to review their texts in a deeper way. Finally, the decision was taken to implement the programme by the researcher to control for teachers’ personal variables. As a result, the treatment effect might have been confounded with the instructor, as the researcher in the experimental groups was new to the students and perhaps less experienced than the regular teachers in the control groups. In future studies, an option could be to train the regular teachers to implement the programme so that, on the one hand, its usefulness can be examined in a more natural context and, on the other hand, to avoid confounding variables. The ultimate purpose is for this programme to become a useful and confident instrument in educational settings and for teachers to be able to implement it efficiently in their classes.
Second, regarding the measurement of knowledge acquisition, two questions arise. The measure of content learning was based upon questions drawn from all the six texts studied during the intervention. This offered us relevant information because, from our point of view, students learned in a different way due to the whole intervention. But a test based on items from six texts might have caused the relative low reliabilities while the learning is topic dependent. In studies like the one we presented here, it would also be favorable to establish a measure directly related to the content learning from the pairs of texts used in the pre- and post-synthesis tasks.

There is still a long way to go. However, it could be said that this is the first study to look at the learning outcome of composing synthesis texts in primary education. In this respect, this study has taken knowledge of the teaching of reading and writing as learning instruments a step forward in the right direction.

Acknowledgements
This Research was supported by the Spanish Ministry of Economy and Competitiveness (EDU2013-46606-C2-1-R) and a grant from the Autonomous University of Madrid, Spain. We are also very grateful to the students and teachers of “el Carmelo Teresiano” and “Mirasierra” schools who took part in this study, and Huub van den Bergh for his assistance in data analysis.

References
Anmarkrud, Ø., Bråten, I., & Størmo, H.I. (2014). Multiple-documents literacy: Strategic processing, source awareness, and argumentation when reading multiple conflicting documents. Learning and Individual Differences, 30, 64–76. Doi: http://dx.doi.org/10.1016/j.lindif.2013.01.007


Martínez, I. (2012). Enseñar a leer y escribir para aprender en la Educación Primaria: Diseño y evaluación de un programa de intervención de escritura de síntesis a partir de varios textos...


Acting in changing worlds: Learning, communication and minds in intercultural activities. Sevilla (Spain).


