Heliyon 7 (2021) e08607

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

Integrating digital documents by means of concept maps: testing an intervention program with eye-movements modelling examples

Victoria García [a], Franck Amadieu [b], Ladislao Salmerón [a,*]

[a] University of Valencia, Spain
[b] University of Toulouse, France

ARTICLE INFO

Keywords:
Reading comprehension
Eye-movement modeling examples
Concept maps

ABSTRACT

When using the Internet to learn about a curricular topic students face the challenge of not only understanding each single document, but also of integrating the ideas in a combined representation. Several intervention studies have tested instructional methods, such as building concept maps, aimed at teaching integration of multiple documents to Secondary education and older students. However, building a concept map may be demanding for learners and requires competencies to build maps in an appropriate way. In the current study we explore the extent to which such integration processes relying a concept map mapping instruction can be efficiently taught to 6th grade students. Specifically, we tested a program which included eye-movement modeling examples (EMMEs) and self-explanation prompts on comprehension and concept mapping in a natural educational setting. An active control group received a similar instruction without EMMEs. Students were randomly assigned to the intervention ($n = 34$) or active control ($n = 32$) groups, and participated in 1-h sessions during four consecutive days. Results indicated that EMME group learnt better than the control group the phases to construct a concept map, as measured in a questionnaire. However, from pre to post test, the EMME group didn't improve comprehension as measured by open-ended reading comprehension questions or in a delayed summary. We conclude that EMMEs can be used to foster literacy strategy instruction of primary school students.

1. Introduction

A major challenge of studying multiple digital documents is to integrate the information from the different texts (Cho and Afflerbach, 2017; List and Alexander, 2018; Magliano et al., 2018; Salmerón, Strømsø, Kammerer, Stadtler, & van den Broek, 2018). An integrated understanding of multiple texts requires that students link content information across different texts by means of bridging inferences in order to achieve a complete understanding of the issue (Rouet and Britt, 2014). This aim can't be reached just by comprehending each text, as information across multiple texts tend to be part of a larger whole not specified in any single text (List and Alexander, 2018).

In a recent systematic review, Barzilai et al. (2018) analyzed interventions aimed at teaching integration of multiple documents. Based on their review of 21 empirical studies, they concluded that intertextual integration can be successfully taught across different grades and disciplines. Several instructional methods proved to be effective, such as by means of explicit instruction of integration, collaborative discussions, using graphic organizers or representations, modeling integration processes, and individual practice. The authors also pointed out some limitations on the scope and the methods used in the literature. Regarding the scope, research so far has addressed mostly secondary school or older students, and the scarce literature on primary school students (Martínez et al., 2015; Wissinger & De La Paz, 2016), needs to be complemented to examine if integration can be taught successfully to younger students. As for the methods used, few studies have included a pre-post design, or used delayed measures. Also, in most cases they combined several instructional practices. Those issues limit our ability to identify the causes of the changes, their persistence after intervention, or the effectiveness of specific interventions.

Building on a long line of research on the interventions to improve students' integration of multiple texts, and on the limitations identified by Barzilai et al. (2018), we investigated the effectiveness of a new instructional program relying on concept mapping from different texts. Specifically, we focused on primary school students, with a methodologically sound design, including a pre-post design in which only one instructional method was manipulated (i.e. Eye-movements modelling examples to build concept maps, EMMEs), and that included delayed measures of integration.

* Corresponding author.
E-mail address: ladislao.salmeron@valencia.edu (L. Salmerón).

https://doi.org/10.1016/j.heliyon.2021.e08607
Received 13 December 2020; Received in revised form 8 July 2021; Accepted 10 December 2021

1.1. Advanced organizers as facilitators of integration of digital texts

Advance organizers, such as concept maps or tables, are rhetoric devices that can represent graphically the organization of different units of information (Lorch, 1989). For example, in concept maps different idea units are represented as nodes which are linked together by means of links that represent potential relationships between the ideas. Students can benefit from having access to advanced organizers when reading digital texts (Britt et al., 1996; Fesel et al., 2015; Naumann et al., 2008; Potelle and Rouet, 2003; Salmerón et al., 2009; Salmerón et al., 2010; Stadtler and Bromme, 2008); for a review see Amadieu and Salmerón (2014). This way, students can focus on comprehending single text information, and can use the organizer to reflect on the potential relationships between the texts. For example, Salmerón et al. (2010) used a graphical organizer to display rhetorical relationships between web pages on the topic of climate change. The organizer displayed together the pages that focused on a similar subtopic (e.g. if pages dealt with solutions for climate change, rather than on the causes), and included a tag for each group that identified the overarching relationship of the group. Such graphical organizer, as compared to a list presentation of the pages, improved undergraduate's scores on a measure aimed to assess participant's ability to draw inferences across texts.

Students can also benefit by constructing advance organizers while reading, as they can engage students on the search for relationships between the texts (Barzilai & Ka'adan, 2017; Hilbert and Renkl 2009; for a review see Barzilai et al., 2018). In their meta-analysis, Schroeder et al. (2018) concluded that concept mapping can be an effective learning activity in comparison to other teaching methods and learning strategies. Processing maps demand summarization which in turn leads to elaborative processing that could concern connection of information from different texts. However, studies have shown that building a map may be too cognitively demanding for learners (Colliot and Jamet, 2018a; 2018b) and, thus, such elaborating task should be designed in a way to support learners' need. To improve concept mapping effectiveness for learning, more investigations of effective mapping functions tools (Sanchiz et al., 2019) or instructional approach as strategy guidance for building maps (Amadieu et al., 2015) are necessary.

Hilbert and Renkl (2009) tested different instructional approaches to benefit from map construction. They tested a group of undergraduate students working with different texts about marketing. One group first practiced by constructing two concept maps, another studied two narrated solved examples on how to build concept maps, and third studied the examples together with self-explanation prompts, such as "“To which phase of the concept mapping process can you assign what Carolin just did? Why?” After instruction, they had to answer a series of integration questions that required combining marketing concepts described in different texts. Pre-post test analyses revealed that the first two groups didn't improve their integrated understanding (Exp. 1 & 2), while the group studying solve problems with prompts did so (Exp. 2).

As noted by Barzilai et al. (2018), the efficacy of constructing maps for integrated understanding has been only tested with Secondary school and older students, although previous studies suggest that students as young as 6th graders can be instructed to integrate information from texts (Martínez et al., 2015; Wissinger & De La Paz, 2016). In addition, recent research has shown that concept maps can be effectively use with 6th graders to teach other advance literacy strategies, such as navigating hypertexts (Fesel et al., 2016). The meta-analysis conducted by Schroeder et al. (2018) showed that the positive effect of concept mapping was observed for different age levels without significant differences between the intermediate level students in grade 4 to 8, secondary students and post-secondary students.

However, one could argue that for such grades a combined instruction based only on textual material, such as study solve examples together with self-explanation prompts used by Hilbert and Renkl (2009) may pose a challenge for some students, as some may still have not acquired full reading comprehension proficiency (van den Broek, 1997). As we describe next, a recent technique based on eye-movements modelling examples (EMME), which presents visual examples of students work on a task, may provide the correct scaffold for young students.

1.2. Eye – movement modelling examples

EMMEs are videos which combine student or expert actions while solving a task, with a moving dot representing their visual inspection patterns. This instructional technique, originally proposed by Tamara van Gog and colleagues, has been mostly applied to visual procedural tasks, such as medical diagnosis (Gegenfurtner et al., 2017; Jarodzka et al., 2012), identification of fish locomotion patterns (Jarodzka et al., 2013), problem-solving tasks (van Gog et al., 2009), or geometric problems (van Marlen et al., 2016, 2018) (for reviews see de Koning et al., 2017; Xie et al., 2021).

Recently EMMEs have also been tested as a way to instruct reading comprehension strategies to middle school students (Mason et al., 2015, 2016; Salmerón and Llorens, 2019) and undergraduate students (Krebs et al., 2019; Salmerón et al., 2020). The fact that EMMEs allow to convey visually complex strategic information, that otherwise must be presented on text, makes it particularly beneficial for middle school students, as in these grades not all students have acquired a proficient reading comprehension level. For example, Salmerón and Llorens (2019) used EMMEs to teach middle school students to self-regulate their navigation while solving comprehension tasks in a Wikipedia environment. Students worked in pairs to discuss a set of EMMEs, which presented examples of good and not so good students implementing a particular navigation strategy (e.g. clicking on a hyperlink before completely reading it). The control group also discussed examples, which were provided as textual narrated cases. Only participants from the EMME group improved their scores from pre to post test in a series of comprehension tasks in a Wikipedia scenario.

The effects of EMMEs on strategic knowledge and performance are potentially due to its efficacy in guiding students' attention during the instruction. Although visualizing the model's gaze can guide students' attention to relevant behaviors is a precondition for effective modelling, they also need to elaborate on the modeled information to improve the chances to learn from them (Bandura, 1986). Self-explanation prompts after watching an EMME, such as the ones used by Hilbert and Renkl (2009), could be used to induce students' elaboration on the model. Self-explanation prompts appear as good instructional factor because learners are brought to retrieve relevant previously learned information from memory in order to connect it with relevant aspects of new information to process.

1.3. The current study

The aim of the current study is to test the effectiveness of an instructional program to instruct 6th grade students to build concept maps from digital documents, to support their integrated understanding of the documents' set. Our study builds upon recent efforts to explore the possibilities to teach integration process to primary school students (Martínez et al., 2015; Wissinger & De La Paz, 2016). In addition, we aimed to test the efficacy of EMMEs in a natural educational setting, as most previous research has used them in laboratory settings (Xie et al., 2021), raising concerns about its ecological validity. Specifically, we adapted the program developed by Hilbert and Renkl (2009) which included narrated solved examples and self-explanation prompts, and which has been proved to be effective in undergraduate students. Our major adaptation consisted on transforming the textual examples into audio narrated videos, consisting of the screen capture of the movements done by the students (EMMEs). Following the recommendations by Barzilai et al. (2018), different methodological advances were considered to maximize the potential impact of the study. First, we included a pre-post test design, to allow testing changes specifically due to the instruction. Second, we included both immediate and delayed measures of comprehension, to detect potential learning effects over time. Third, we used an active control group, which also received an instruction, and that only varied from the experimental control in one instructional method (the presence or not of EMMEs).

We considered different hypotheses regarding the effects of instruction. For the group participating in the instructional program with EMME, from pre to post test we expected an improvement on students' knowledge of the map construction phases and on immediate and delayed integrated comprehension, after controlling for potential confounding effects of students' reading comprehension skills. For the group participating in the control instructional program, we expected lower improvements on knowledge about how to construct a concept map and comprehension than those observed at the instructional group.

2. Method

2.1. Participants

Sixty - six 6th grade students (11–12 years old) from a public school in a city near (removed for blind review) participated in the study (51.5% female). Data from six students who didn't attend the four sessions, and that of a student diagnosed with learning disabilities, were not included in the analyses. None of the participants from the final sample have been diagnosed with hearing or attention deficits. They were all native Spanish speakers. The school followed a conservative approach towards technology. Although the group had a computer room, they only used it once or twice a month for academic purposes. Participation was approved by the School board, who planned the sessions as part of the instructional activities in the school aimed at improving students' digital skills. The study was approved by the Research Ethics committee of the University of (removed for blind review), and followed the guidelines of the Declaration of Helsinki.

The sample size was determined based on a power analysis (GPower 3.1: Faul et al., 2009) and considerations about feasibility and cost. The expected effect of EMMEs on performance was estimated to be of medium size, based on a metanalytical study (Xie et al., 2021). With the present sample size, a medium effect (f = .25) in the within-between interaction (that of Time -pre vs. post test- and Instruction Group -control vs. EMME-) can be detected with a power of $1\text{-}\beta = .80$ and $\alpha = .05$.

2.2. Materials

2.2.1. Documents

We used two sets of short documents related to the topics Energy and Arthropods, obtained from printed and digital textbooks dedicated to 6th grade science classes (Cáliz, 2009; Martínez, 2005) (see examples in Figure 1). Both topics were part of the official curriculum for 6th grade in the school, but students had not received any instruction on them yet. Each set included 12 short documents (Energy: 559 words in total, Arthropods: 493). While most document types included plain text pages resembling textbooks excerpts, some of the documents consisted of bar graphs, infographics, and visually rich web pages. We selected the documents trying to keep a similar underlying structural relation between them. First, a document described the core idea of the set. For example, the document "Energy" included a technical definition of the term, and a bar graph representing the consumption of different types of energies in Spain and in the Valencian region. Then, two documents further developed the main topic in two subtopics (e.g. 'Renewable energy' and 'Non-renewable energy'). Finally, nine additional documents elaborated on particular aspects of one of the two subtopics (e.g. "Hydroelectric power plant", included an infographic to explain how a hydroelectric central works, as was expected to be linked to the "Renewable energy" branch of the map), or to both of the subtopics (e.g. "Electricity").

2.2.2. Instructional program

Our instructional program aimed at teaching how to build concept maps was based on the program developed by Hilbert and Renkl (2009) described in the introduction, which included narrated solved examples and self-explanation prompts. We adapted the original program in two ways. First, we allowed students to work in pairs, as this has proved to be an effective way to engage young students in more elaborated discussions (Sizmur and Osborne, 1997). Second, we transformed the textual examples into audio narrated videos, consisting of the screen capture of the movements done by the students and a narrated voice explaining the reason behind the movements. In addition, only in the experimental group, those videos included a moving dot representing students' eye-movements while solving the task. A detailed description of one of the videos is provided in Table 1. The program had two phases. In step 1, all students watched a whiteboard animation video that provided direct instruction about: a) the parts of a conceptual map; b) the phases to access the documents and to build the map: planning, monitoring and evaluating; and c) strategies that have proved useful in each phase (Amadieu et al., 2015): initial access to the documents based on thematic coherence, planning the construction, organizing the structure, and revising the map. In step 2, students listened and watched seven solved examples videos (average length 59 s), which included eye-movements modeling or not, depending on the condition. The videos illustrated specific phases of the map construction process (see example in Table 1 and https://go.uv.es/KCXYWa5). After each video, students worked in pairs to identify the phase that corresponded to the actions depicted in the video. They were required to reach an agreement, and to write down the phase they identified.

2.3. Questionnaires

2.3.1. Knowledge test about how to construct a concept map

During step 2 of the instructional program (see above), students worked in pairs to identify the phase of the map construction process depicted in each of the seven solved examples videos: planning, monitoring and evaluating. For each solved example video participants were given a score of 1 if they correctly identified the phase described in the video, or a 0 if the failed to do so.

2.3.2. Comprehension questions

For each set of multiple documents ('Energy' and 'Artrophods') we developed a set of short open-ended comprehension questions targeting literal and inferential comprehension. For each topic, three of the questions assessed literal comprehension, which targeted important ideas explicitly mentioned in a single document. For example, one of the questions was "*Why is electricity so useful?*". To answer this, the student had to recall the information located in the text "Electricity", specifically "*Electricity is very useful because it is easily transported and it can be transformed into any other kind of energy*". Three additional questions assessed inter-text integration, as they required to infer the relations between relevant information from different documents. For example, the question: "*Which is the difference between renewable energies and not renewable energies?*" required students to combine gist information from the document called "Renewable energies" and from the document "Non-renewable energies". Complete and correct questions were given 1 point, and incorrect or incomplete questions a 0. The first and last author assessed 20% of the responses for each topic, using a previously developed rubric. Raters reached a significant agreement in both questionnaires (Cohen's $\kappa = .90$ for Energy and .83 for Arthropods). Conflicts were solved after discussion. The first author completed the evaluation of the rest of the questions. Although our original goal was to use separate scores for each comprehension type (i.e. literal and integration), our attempts during piloting and the final study revealed low reliability for such separated scales. This is why we decided to combine all responses from each topic in a single comprehension scale, with moderate reliability, with Cronbach's $\alpha = .61$ for "Energy" and .59 for "Artrophods".

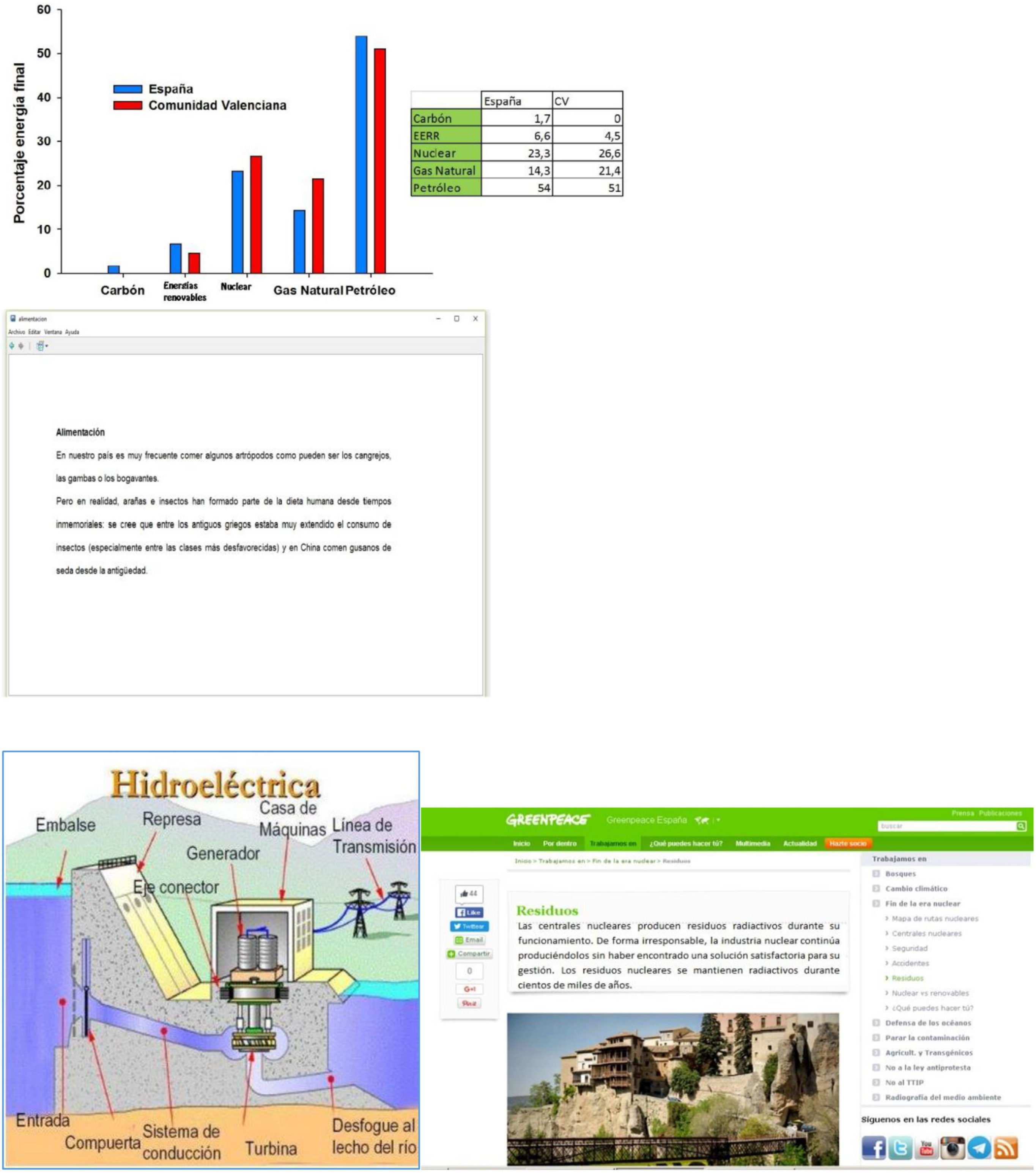

Figure 1. Examples of the documents used for the topic "Energy".

2.3.3. Delayed integration summary

The day after completing the map with each set of documents, students had to summarize the contents from the documents studied. Students were given 5 min to complete the task, and they had no access to the map or to the documents studied. On average, students wrote rather short summaries for the "Energy", $M = 27.6$ words, $SD = 13.8$, and "Artrophods", $M = 24.7$ words, $SD = 14.0$, topics. To evaluate students' integration we constructed an ad-hoc rubric. Each summary could get up to three points: 1 point if the student explicitly identified the two main categories developed in each topic (e.g. "There are two types of energies: renewable and non-renewable"); 1 point if the student established a particular relationship between ideas from different documents ("Petroleum and coal are non-renewable energies that can't be replaced once they are used, but renewable energies will last"); and 1 point if the

Table 1. Description of solved example video #1.

Goal: reading all node titles and extract the main idea		
What the narrator says	What the narrator makes with computer mouse	Where the narrator looks at (only for EMME condition)
"First of all, I am going to look around the information provided. Oh, I see that there are twelve different nodes… I am going to read the titles to get an overview of the topic".		A circle representing the eyes looks sequentially at all titles.
"Let's go, I am going to start reading by going to the most important node"		
"'Renovable energies' is really important, and it include water, son, wind, electric power…"	Computer mouse moves to title 'Renovable energies'	Looks at specific nodes as she mentions them
"Oh no, but nuclear power, radioactive material and all those are not included under 'Renovable energies'"		Looks at specific nodes as she mentions them
"I am going to continue Reading… Ops! Energy! 'Energy' includes them all.	Computer mouse moves to title 'Energy'	Looks at node 'Energy'
"Let me double check… yes, 'Energy' includes all sections"	Computer mouse stays over title 'Energy'	Looks quickly at most nodes
"Then, I will start by reading this"	Clicks on node 'Energy'	Reads completely the section 'Energy'
Task after the video: *In which phase of the process of constructing concept maps are we? Why do you think we are in this phase?*		

students further develops an idea from at least one of the documents that were related). 20% of the summaries from each topic were rated by the first and third authors, reading a 100% of agreement. The first author completed the evaluation of the rest of the summaries.

2.3.4. Reading comprehension test

We used the Test of Comprehension Strategies (Vidal-Abarca et al., 2007), a standardized comprehension test in Spanish, composed of two expository texts and ten multiple-choice questions per text. The test included literal and inferential questions, which had to be answered without accessing the texts.

2.4. Procedure

The procedure included four sessions of approximately 50 min each, and they took place in consecutive days (see Figure 2). Students were randomly assigned to the experimental ($n = 34$) or active control ($n = 32$) group. This ensured that any preexisting difference between participants was randomly distributed across the control and experimental group. The initial session served as pre-test. First, students were instructed in how to use CMapsTool, a free software dedicated to build concept maps with digital documents. When they felt confident on the basic use of the tool, they were introduced to the first set of multiple documents. Specifically, they were provided with a CMaps consisting of twelve randomly organized documents about "Energy" (Figure 2). They were instructed to build a concept map combining the documents and creating and labeling links between them. Once they finished, they answered the comprehension questions. Students had no access to the documents while answering the questions. At the beginning of the second session, students were requested to summarize the contents from the "Energy" documents studied the previous day, without being able to access the map or the documents. Afterwards, the instructional program was provided. In the third session, which served as post-test, students build a map about the topic "Arthropoda", starting from a set of twelve documents. Once they had finished, they answered the open-ended comprehension questions. Finally, in the fourth session, students summarized the contents from the "Arthropoda" documents studied the previous day, following the same procedure as in the pre-test. Next, they completed a standardized test to measure their reading comprehension skills. Finally, they were debriefed and an open discussion was held with the students to discuss their perceived usefulness of the program.

2.5. Design

We used a mixed design with Time (pre vs. post test) as repeated measures factor, and Instruction Group (control vs. EMME) as between-participants factor. As dependent variables, we used scores on the knowledge test about how to construct a concept map, comprehension scores (i.e. number of correct responses to the open-ended comprehension questions), and summary scores (i.e. sum of points obtained in the summary with the integration rubric). Results were analyzed with generalized linear mixed effects models, that were constructed following the theoretical model described in the introduction. Specifically, we tested the effect of our EMME intervention against a control group, with a pre-post test design. Pre-existing individual differences were introduced as covariates to control for their potential effects, mimicking the design of similar previous studies (Hilbert and Renkl, 2009; Salmerón et al., 2020; Salmerón and Llorens, 2019). Although in principle other more complex models are possible (e.g. interaction between individual differences and intervention effects), the lack of solid theoretical arguments in the field prevented us from exploring other options.

3. Results

Although students were assigned to groups randomly, we first ensured that groups didn't differ in terms of performance measures at pre-test (for descriptive scores see Table 2). Student t comparisons revealed that there were no group differences on pre-test text comprehension

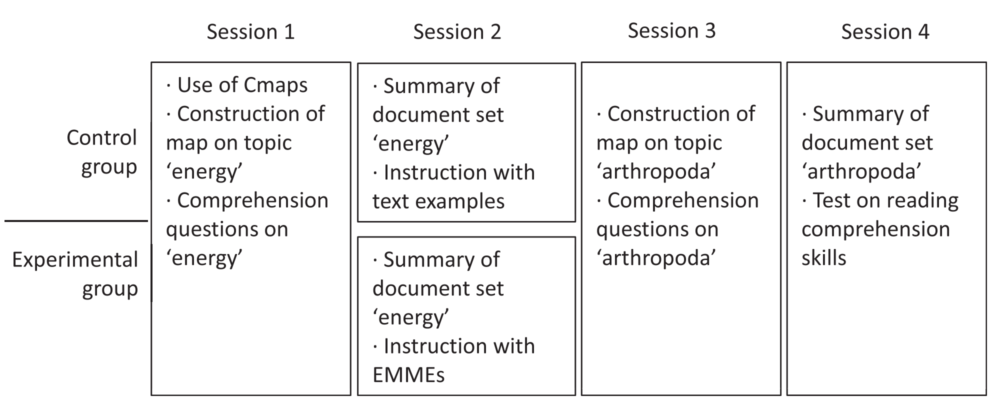

Figure 2. Procedure followed in the study.

Table 2. Descriptive statistics for all measured variables, by Time and Instruction group.

	Pre-test		Post-test	
	Control	EMMEs	Control	EMMEs
Map construction knowledge scores			5.2 (0.9)	5.6 (1.1)
Comprehension questions	50.0 (24.1)	50.0 (23.0)	31.9 (22.8)	29.0 (18.7)
Delayed summary:				
0 points	36.7%	41.9%	31.0%	21.9%
1 point	36.7%	29.0%	44.8%	40.6%
2 points	16.7%	29.0%	6.9%	28.1%
3 points	10.0%	0%	17.2%	9.4%

scores or pre-test delayed integration summary (all $t < 1$). In addition, scores from the standardized reading comprehension test indicated that all students showed a grade-appropriate competence level.

3.1. *Effects on knowledge of map construction phases*

To test the effects of the instruction on students' knowledge of map construction, we conducted a generalized linear mixed effects model (GLME) using the *glmer* function from the lme4 package (Bates et al., 2015) in R version 1.2.5024 (R Core Team, 2020). The fixed factors in the model were Instruction Group (dummy coding: control = 0 vs. EMME = 1) and Reading Comprehension Skills (raw scores). As random factors we included Study partnership and Test item. Please note that as this task was performed in pairs, study partnership was included as random factor, but not participant. The dependent variable was either 0 for incorrect responses or 1 for correct responses for each of the seven solved examples videos regarding concept map construction phases. Accordingly, we used a binomial distribution. The fitted model and complete results are presented in Table 3. Results showed a main effect of Instruction, indicating that students in the EMME group scored higher than those in the control group. Of note is that the effect estimate was rather low, and actually a 95% confidence intervals indicated that the results were not significant. Reading comprehension skills did not influence students' scores.

3.2. *Effects on comprehension*

To test the effects of the instruction on comprehension, we first run a GLME model (*glmer* function with binomial distribution) with the following fixed factors: Time (dummy coding: pre = 1 vs. post test = 0), Instruction Group (dummy coding: control = 0 vs. EMME = 1), the interaction Time and Instruction, and Reading Comprehension Skills (raw scores). As random factors we included Participant, and Test item. Please note that as this task was done individually, we included participant and not study partnership as random factor. The dependent variable was either 0 for incorrect responses or 1 for correct responses to each comprehension question. The model fitted and complete results are presented in Table 4.

Overall, means range from 29 to 50% (see Table 2), which indicates a modest to medium ability to comprehend and integrate information across documents. In interpreting these results, please keep in mind that students responded to open-ended comprehension questions, and therefore chance level can be considered 0. Furthermore, results showed non-significant effects of Time, Instruction, or its interaction (see Table 4). Only the factor Reading comprehension skills resulted in a significant effect. Good comprehenders score higher in the open-ended comprehension questions administered immediately after the completion of each map.

Second, to analyze the effects on delayed summaries, we run an additional GLME model (*glmer* function with binomial distribution) with Time (dummy coding: pre = 1 vs. post test = 0), Instruction Group (dummy coding: control = 0 vs. EMME = 1), the interaction Time and Instruction, and Reading Comprehension Skills (raw scores) as fixed factors; and Participant and Summary category as random factors. Please note that as this task was done individually, we included participant and not study partnership as factor. As dependent variable we used the scores on the delayed integration summary, with a score of 0 or 1 in each of the three categories assessed in the rubric (see method). The model fitted and complete results are presented in Table 5.

Regarding the quality of the delayed summaries, the number of students who obtained the lower score surpassed those who got the highest score (see Table 2), a pattern which speaks about the difficulty of the task. Results from the last model showed non-significant effects of Time, Instruction, or its interaction. Finally, the factor Reading comprehension skills was significant, indicating that good comprehenders wrote better summaries one day after having worked with the documents (see Table 5).

4. Discussion

This study contributes to research on instructions to foster integrated understanding, and demonstrates that a combination of EMMEs and self-explanation prompts can be used to teach 6^{th} grade students to build

Table 3. Model for the effects on knowledge of map construction phases.

Fitted model:

map_construction_score ~ instruction_group + reading_comprehension + (1|partnership)+ (1|map_construction_item), family = binomial

Random effects:

Groups Name	Variance	Std.Dev.
Partnership (intercept)	0.67	0.82
Map_construction_item (intercept)	2.16	1.47

Fixed effects:

	Estimate	Std. Error	z value	Pr(>|z|)	95% CI
(Intercept)	0.84	0.79	1.06	0.29	[0.31, 2.80]
Instruction_group	0.81	0.42	1.93	0.05 *	[-0.01, 1.62]
Reading_comprehension	0.08	0.05	1.50	0.12	[-0.07, 0.59]

Table 4. Model for the effects on comprehension questions.

Fitted model:

comprehension_score ~ instruction_group * time + reading_comprehension + (1|participant)+ (1|comprehension_item), family = binomial

Random effects:

Groups Name	Variance	Std.Dev.
Participant (intercept)	0.37	0.61
Comprehension item (intercept)	1.88	1.37

Fixed effects:

	Estimate	Std. Error	z value	Pr(>|z|)	95% CI
(Intercept)	-1.99	0.59	-3.37	0.001 ***	[-2.24, -0.29]
Instruction group	-0.02	0.27	-0.07	0.95	[-0.55, 0.51]
Time	0.74	0.67	1.11	0.27	[-0.56, 2.05]
Reading comprehension	0.08	0.03	2.41	0.02 *	[0.05, 0.46]
Instruction group x Time	0.28	0.30	0.95	0.34	[-0.30, 0.87]

Table 5. Model for the effects on delayed summaries.

Fitted model:

summary_score ~ instruction_group * time + reading_comprehension + (1|participant)+ (1|summary_item), family = binomial

Random effects:

Groups Name	Variance	Std.Dev.
Participant (intercept)	0.83	0.91
Summary item (intercept)	0.32	0.56

Fixed effects:

	Estimate	Std. Error	z value	Pr(>|z|)	95% CI
(Intercept)	-1.59	0.67	-2.39	0.02 *	[-1.54, 0.19]
Instruction group	0.29	0.41	0.71	0.48	[-0.52, 1.10]
Time	-0.24	0.36	-0.68	0.50	[-0.94, 0.46]
Reading comprehension	0.10	0.05	1.94	0.05 *	[0.00, 0.67]
Instruction group x Time	-0.46	0.49	-0.93	0.35	[-1.42, 0.50]

concept maps from multiple documents. Of note is that this is one of the few studies that has used EMMEs in a natural educational context, which allowed us to test the ecological validity of this method. In addition, it evidences that 6th grade students show a modest ability to integrate complementary information from curricular topics provided in different documents.

Uniquely to our design is that the control group received an instruction similar to the experimental group, with the only exception of the way the examples were provided. While the active control group read and discussed solved examples written in text, the experimental group saw and discussed the audiovisual EMMEs. Thus, our design allowed to disentangle the unique contribution of a single instructional method, something that is rare in previous intervention studies (Barzilai et al., 2018). Specifically, the effect of EMMEs on the acquisition of knowledge about concept map construction phases could be attributed to the support provided in guiding students' attention during their learning towards relevant parts of the process of constructing a conceptual map. These results converge with recent studies demonstrating that EMMEs can be an effective instructional tool to teach complex reading strategies (Krebs et al., 2019; Mason et al., 2015, 2016; Salmerón and Llorens, 2019). These findings contributed to the design of more effective concepts mapping tasks highlighting that learners need procedural guidance to understand how to build more elaborated maps. Audiovisual EMMEs appear as a promising guidance that help learners to focus more on the relevant phases of concept mapping.

Unexpectedly, no evidence confirmed that EMMEs exerted any effect on integrated understanding. A potential reason for the lack of effect on comprehension of contents is that we used a combined scale with both questions targeting literal and inter-text integrated comprehension, to avoid having two separated but highly unreliable scales. Thus, the comprehension questions used don't uniquely assess integrated understanding, and the intervention is not aimed to improve comprehension of single documents. Another explanation of the lack of results of EMMEs on comprehension could be that the mapping task was not really supportive for comprehension. Learners could have focused their attention of the mapping task more than to comprehension. The positive effect of EMME observed on knowledge about the phases to construct concept maps and not on comprehension supports this interpretation. These results are finally in line with recent works from Colliot and Jamet (2018a, 2018b) that underlined that concept mapping can be infective for learning even when learners produce relevant maps. Guidance can be sometimes ineffective and even detrimental for learning when it diverts attention from contents or require demanding cognitive processing to interpret guidance, like it was sometimes observed with attentional signaling and cueing in learning from animations (Berney and Bétrancourt, 2016).

Moreover, consistent with the previous point, in interpreting the average scores in the open-ended comprehension questions and delayed summaries, we must consider the considerable level of difficulty of the tasks used. The pattern of results suggests that 6th grade students show a modest ability to integrate ideas from multiple documents (Martínez et al., 2015; Wissinger & De La Paz, 2016). Students' reading comprehension skills predict their document comprehension and integration as measured by an immediate test, but it doesn't play a significant role in the delayed summaries.

4.1. Limitations and future research

Of course, our study comes with certain limitations. First, different reasons may explain the lack of positive effect on comprehension of EMME and of mapping (pre- and posttest), as well as the rather small effect on students' knowledge of map construction phases. The study aimed at examining the effects of EMMEs on relevant procedures for mapping that was supposed to improve comprehension. However, the study did not ensure that concept mapping was particularly adapted for the learning aims. In spite of a positive effect of the EMME on knowledge of map construction phases, it did not support higher comprehension. Therefore, further studies should try to disentangle effects on comprehension that are due to the nature of the task (i.e. level of active learning) and those that are due to the relevance of the procedures used by students. Because concept mapping may have been demanding for students even with the help of EMME, further studies should compare concept mapping with less active learning task such as reading and recall tasks (Blunt and Karpicke, 2014), and should also investigate how guidance of relevant procedure to perform the task contribute to learning. Even for more traditional and less elaborative tasks as reading, students may need to acquire and use adapted procedures. Moreover, it could be supposed that effective mapping for comprehension would require more automatic knowledge of map construction phases that could be trained by a use on longer period. Finally, the instruction imposed to students the use of a software that may also have required cognitive resources for students. Their lack of practices with the tool could have been detrimental for their learning by playing the role of an additional task. Using a software implies declarative knowledge of functions and features as well as procedural knowledge to perform the task by using the software functions. In any case, in our study students performed the task once they felt confident on the basic use of the tool.

Second, the results of the present study could also shed light on the need of investigating more clearly how improving EMMEs and measuring their real effect on online behavior and mental representation. Regarding the mental representation that learners build of EMMEs, it might be possible that students interpreted the function of EMME in ways different than those expected by the experimenters. Showing an example of an effective mapping procedure could have led students to construct a task model based of the reproduction of such a procedure. Their representation of the goal consisting in the replication of the procedure could have diverted student from the core task, that is to say, processing texts to construction coherent mental representation of information. More investigations should be conducted on the learners' mental representations of the task and EMMEs meaning for their task.

Finally, future research should expand the feasibility of using concept maps with primary school students to build relationships between documents that go beyond the documents' ideas, by including source information or rhetorical relationships between the documents. Recent research by Barzilai and Ka'adan (2017) with 9th grade students has proved that building maps may support those document-level processing. Our study aims to stimulate research on the potentialities and limitations of constructing concept maps to boost integrating understanding in the last years of primary school.

Declarations

Author contribution statement

Victoria García: Conceived and designed the experiments; Performed the experiments; Analyzed and interpreted the data; Contributed reagents, materials, analysis tools or data; Wrote the paper.

Franck Amadieu: Conceived and designed the experiments; Wrote the paper.

Ladislao Salmerón: Conceived and designed the experiments; Analyzed and interpreted the data; Contributed reagents, materials, analysis tools or data; Wrote the paper.

Funding statement

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Data availability statement

Data will be made available on request.

Declaration of interests statement

The authors declare no conflict of interest.

Additional information

No additional information is available for this paper.

References

Amadieu, F., Salmerón, L., 2014. Concept maps for comprehension and navigation of hypertexts. In: Hanewald, R., Ifenthaler, D. (Eds.), Digital Knowledge Maps in Education. Springer, New York, pp. 41–59.

Amadieu, F., Salmerón, L., Cegarra, J., Paubel, P.-V., Lemarié, J., Chevalier, A., 2015. Learning from concept-mapping and hypertext: an eye tracking study. Educ. Technol. Soc. 18, 100–112.

Bandura, A., 1986. Social Foundations of Thought and Action: A Social Cognitive Theory. Prentice-Hall, Inc.

Barzilai, S., Ka'adan, I., 2017. Learning to integrate divergent information sources: the interplay of epistemic cognition and epistemic metacognition. Metacognition Learn. 12, 193–232.

Barzilai, S., Zohar, A.R., Mor-Hagani, S., 2018. Promoting integration of multiple texts: a review of instructional approaches and practices. Educ. Psychol. Rev. 30, 973–999.

Bates, D., Maechler, M., Bolker, B., Walker, S., 2015. Fitting linear mixed-effects models using lme4. J. Stat. Software 67, 1–48.

Berney, S., Bétrancourt, M., 2016. Does animation enhance learning? A meta-analysis. Comput. Educ. 101, 150–167.

Blunt, J.R., Karpicke, J.D., 2014. Learning with retrieval-based concept mapping. J. Educ. Psychol. 106 (3), 849–858.

Britt, M.A., Rouet, J.-F., Perfetti, C.A., 1996. Using hypertext to study and reason about historical evidence. In: Rouet, J.-F., Levonen, J.J., Dillon, A., Spiro, R.J. (Eds.), Hypertext and Cognition. Lawrence Erlbaum Associates, Mahwah, NJ, pp. 43–72.

Cáliz, S., 2009. Conocimiento del medio, 6 Primaria. Ediciones SM.

Cho, B.Y., Afflerbach, P., 2017. An evolving perspective of constructively responsive reading comprehension strategies in multilayered digital text environments. In: Israel, S.E. (Ed.), Handbook of Research on reading Comprehension, second ed. Guilford, New York, pp. 109–134.

Colliot, T., Jamet, É., 2018a. Does self-generating a graphic organizer while reading improve students' learning? Comput. Educ. 126, 13–22.

Colliot, T., Jamet, É., 2018b. How does adding versus self-generating a hierarchical outline while learning from a multimedia document influence students' performances? Comput. Hum. Behav. 80, 354–361.

de Koning, Jarodzka, H., 2017. Attention guidance strategies for supporting learning from dynamic visualizations. In: Lowe, R., Ploetzner, R. (Eds.), *Learning from dynamic visualization*. Springer, pp. 255–278.

Faul, F., Erdfelder, E., Buchner, A., Lang, A.-G., 2009. Statistical power analyses using G*Power 3.1: tests for correlation and regression analyses. Behav. Res. Methods 41 (4), 1149–1160.

Fesel, S.S., Segers, E., Clariana, R.B., Verhoeven, L., 2015. Quality of children's knowledge representations in digital text comprehension: evidence from pathfinder networks. Comput. Hum. Behav. 48, 135–146.

Fesel, S.S., Segers, E., De Leeuw, L., Verhoeven, L., 2016. Strategy training and mind-mapping facilitates children's hypertext comprehension. Writ. Lang. Lit. 19, 131–156.

Gegenfurtner, A., Lehtinen, E., Jarodzka, H., Säljö, R., 2017. Effects of eye movement modeling examples on adaptive expertise in medical image diagnosis. Comput. Educ. 113, 212–225.

Hilbert, T.S., Renkl, A., 2009. Learning how to use a computer-based concept-mapping tool: self-explaining examples helps. Comput. Hum. Behav. 25, 267–274.

Jarodzka, H., Balslev, T., Holmqvist, K., Nyström, M., Scheiter, K., Gerjets, P., Eika, B., 2012. Conveying clinical reasoning based on visual observation via eye-movement modeling examples. Instr. Sci. 40 (5), 813–827.

Jarodzka, H., van Gog, T., Dorr, M., Scheiter, K., Gerjets, P., 2013. Learning to see: guiding students' attention via a model's eye movements fosters learning. Learn. InStruct. 25, 62–70.

Krebs, M.C., Schüler, A., Scheiter, K., 2019. Just follow my eyes: the influence of model-observer similarity on Eye Movement Modeling Examples. Learn. InStruct. 61, 126–137.

List, A., Alexander, P.A., 2018. Postscript: in pursuit of integration. Learn. InStruct. 57, 82–85.

Lorch Jr., R.F., 1989. Text-signaling devices and their effects on reading and memory processes. Educ. Psychol. Rev. 1, 209–234.

Magliano, J.P., McCrudden, M.T., Rouet, J.F., Sabatini, J., 2018. The modern reader: should changes to how we read affect theory and research? In: Schober, M.F., Rapp, D.N., Britt, M.A. (Eds.), Handbook of Discourse Processes, second ed. Routledge, New York, pp. 343–361.

Martínez, R., 2005. Los antrópodos del planeta Tierra. Ministerio de Educación.

Martinez, L., Mateos, M., Martín, E., Rijlaarsdam, G., 2015. Learning history by composing synthesis texts: effects of an instructional program on learning, reading and writing processes, and text quality. J. Writ. Res. 7, 275–302.

Mason, L., Pluchino, P., Tornatora, M.C., 2015. Eye-movement modeling of integrative reading of an illustrated text: effects on processing and learning. Contemp. Educ. Psychol. 41, 172–187.

Mason, L., Pluchino, P., Tornatora, M.C., 2016. Using eye-tracking technology as an indirect instruction tool to improve text and picture processing and learning. Br. J. Educ. Technol. 47, 1083–1095.

Naumann, J., Richter, T., Christmann, U., Groeben, N., 2008. Working memory capacity and reading skill moderate the effectiveness of strategy training in learning from hypertext. Learn. Indiv Differ 18, 197–213.

Potelle, H., Rouet, J.F., 2003. Effects of content representation and readers' prior knowledge on the comprehension of hypertext. Int. J. Hum. Comput. Stud. 58, 327–345.

R Development Core Team, 2020. R: A Language and Environment for Statistical Computing. Vienna: R Foundation for Statistical Computing, Vienna, Austria.

Rouet, J.-F., Britt, M.A., 2014. Learning from multiple documents. In: Mayer, R.E. (Ed.), Cambridge Handbook of Multimedia Learning, second ed. Cambridge University Press, Cambridge, MA, pp. 813–841.

Salmerón, L., Llorens, A., 2019. Instruction of digital reading strategies based on eye-movements modeling examples. J. Educ. Comput. Res. 57, 343–359.

Salmerón, L., Baccino, T., Cañas, J.J., Madrid, R.I., Fajardo, I., 2009. Do graphical overviews facilitate or hinder comprehension in hypertext? Comput. Educ. 53, 1308–1319.

Salmerón, L., Gil, L., Bråten, I., Strømsø, H.I., 2010. Comprehension effects of signalling relationships between documents in search engines. Comput. Hum. Behav. 26, 419–426.

Salmerón, L., Strømsø, H.I., Kammerer, Y., Stadtler, M., van den Broek, P., 2018. Comprehension processes in digital Reading. In: Barzillai, M., Thomson, J., Schroeder, S., van den Broek, P. (Eds.), Learning to Read in a Digital World. John Benjamins, Amsterdam, pp. 91–120.

Salmerón, L., Delgado, P., Mason, L., 2020. Using eye-movement modelling examples to improve critical reading of multiple webpages on a conflicting topic. J. Comput. Assist. Learn. 36, 1038–1051.

Sanchiz, M., Lemarié, J., Chevalier, A., Cegarra, J., Paubel, P.V., Salmerón, L., Amadieu, F., 2019. Investigating multimedia effects on concept map building: Impact on map quality, information processing and learning outcome. Educ. Inform. Technol. 24, 3645–3667.

Schroeder, N.L., Nesbit, J.C., Anguiano, C.J., Adesope, O.O., 2018. Studying and constructing concept maps: a meta-analysis. Educ. Psychol. Rev. 30, 431–455.

Sizmur, S., Osborne, J., 1997. Learning processes and collaborative concept mapping. Int. J. Sci. Educ. 19, 1117–1135.

Stadtler, M., Bromme, R., 2008. Effects of the metacognitive computer-tool met.a.ware on the web search of laypersons. Comput. Hum. Behav. 24, 716–737.

van den Broek, P., 1997. Discovering the cements of the universe: the development of event comprehension from childhood to adulthood. In: van den Broek, P., Bourg, T. (Eds.), Developmental Spans in Event Comprehension: Bridging Fictional and Actual Events. Erlbaum, Mahwah, NJ, pp. 321–342.

van Gog, T., Jarodzka, H., Scheiter, K., Gerjets, P., Paas, F., 2009. Attention guidance during example study via the model's eye movements. Comput. Hum. Behav. 25 (3), 785–791.

van Marlen, T., van Wermeskerken, M., Jarodzka, H., van Gog, T., 2016. Showing a model's eye movements in examples does not improve learning of problem-solving tasks. Comput. Hum. Behav. 65, 448–459.

van Marlen, T., van Wermeskerken, M., Jarodzka, H., van Gog, T., 2018. Effectiveness of Eye Movement Modeling Examples in problem solving: the role of verbal ambiguity and prior knowledge. Learn. InStruct. 58, 274–283.

Vidal-Abarca, E., Gilabert, R., Martínez, T., Sellés, P., Abad, N., Ferrer, C., 2007. TEC. Test de Estrategias de Comprensión (*Test of Comprehension Strategies*). ICCE, Madrid.
Wissinger, D.R., De La Paz, S., 2016. Effects of critical discussions on middle school students' written historical arguments. J. Educ. Psychol. 108, 43–59.
Xie, H., Zhao, T., Deng, S., Peng, J., Wang, F., Zhou, Z., 2021. Using eye movement modelling examples to guide visual attention and foster cognitive performance: a meta-analysis. J. Comput. Assist. Learn.