frontiers | Frontiers in Education

TYPE Original Research
PUBLISHED 14 March 2024
DOI 10.3389/feduc.2024.1287744

OPEN ACCESS

EDITED BY
Luisa Losada-Puente,
University of A Coruña, Spain

REVIEWED BY
Kevin Mark Roessger,
University of Arkansas, United States
Pedro Simão Mendes,
University Institute of Lisbon, Portugal

*CORRESPONDENCE
Juana M. Ortega-Tudela
✉ jmortega@ujaen.es

RECEIVED 02 September 2023
ACCEPTED 27 February 2024
PUBLISHED 14 March 2024

CITATION
Lechuga MT, Ortega-Tudela JM and Gómez-Ariza CJ (2024) Retrieval-based concept mapping makes a difference as a retrieval practice activity: a study with high school students.
Front. Educ. 9:1287744.
doi: 10.3389/feduc.2024.1287744

Retrieval-based concept mapping makes a difference as a retrieval practice activity: a study with high school students

María Teresa Lechuga[1], Juana M. Ortega-Tudela[2]* and Carlos J. Gómez-Ariza[1]

[1]Department of Psychology, University of Jaén, Jaén, Spain, [2]Department of Pedagogy, University of Jaén, Jaén, Spain

Introduction: While specific sequences of retrieval-based activities have been shown to have a powerful effect on learning, no previous study has examined this issue in children or adolescents. Here, we aimed to determine whether the benefit of concept mapping as an initial retrieval activity observed previously with college students in a lab setting may also be found in younger and less experienced students in a more naturalistic school setting.

Methods: After a short training in concept mapping, participants (N = 60) read an educational text and then engaged in a sequence of two learning activities that required them to retrieve relevant ideas from the text. One of the activities involved free recall by writing down as many ideas as possible and the other involved creating a concept map, both of them in the absence of educational texts. Critically, we manipulated the order in which the activities were performed.

Results: A mediation analysis with success during retrieval practice as the mediator revealed both a direct effect of retrieval sequence and an indirect effect. Creating a concept map first and then freely recalling by writing paragraphs significantly improved performance on a 2-week delayed learning test, as compared to performing the same activities in the inverse order, even when doing concept mapping first led to lower success rates during practice.

Discussion: These results support the idea that concept learning from instructional texts can be modulated by the sequence of retrieval activities performed and point to the educational value of retrieval-based concept mapping as a first learning activity within a series across different learners and educational settings.

KEYWORDS

concept mapping, retrieval-based learning, teaching strategies, adolescents, testing effect

1 Introduction

Findings from a good number of studies show that retrieval-based study activities can be highly effective in promoting long-term learning across learner characteristics, materials, and contexts (Karpicke, 2017; Moreira et al., 2019; Yang et al., 2021). In retrieval-based learning, students engage in activities that require them to repeatedly recall previously studied information (e.g., answering a teacher's questions or self-created flashcards) once study materials are set aside. Importantly, it has been shown that retrieval practice may lead to better

student performance (the so-called "testing effect") on subsequent learning tests than do other traditional learning activities such as rereading or concept mapping (Adesope et al., 2017, for a meta-analysis; Nunes and Karpicke, 2015, for a review), and that this benefit can be observed either in laboratory or natural school settings, including when genuine curriculum materials are used (Moreira et al., 2019; Lamotte et al., 2021, for reviews; see also Ortega-Tudela et al., 2021).

A common retrieval strategy that has been systematically considered in studies is free recall [i.e., participants are asked to write down all of the information they can remember once learning material has been removed; see reviews by Nunes and Karpicke (2015) and Karpicke (2017)], which has typically been compared with extra study (i.e., rereading) of materials. Roediger and Karpicke (2006), for example, compared two retrieval practice conditions (one involving 3 study periods and 1 retrieval period, and another comprising 1 period of study and 3 of retrieval) with rereading (4 periods of study), matching total time of learning activities across the different conditions. Participants' learning was tested either 5 min or 1 week after they undertook the learning activities. Although there was a greater benefit from the rereading condition in the immediate test, it was the group that had engaged in repeated retrieval (free recall) that exhibited better retention in the delayed final assessment.

In addition to rereading, free recall has also been compared to more elaborative study strategies such as concept mapping (Karpicke and Blunt, 2011; Lechuga et al., 2015; O'Day and Karpicke, 2021; for other elaborative strategies see Fritz et al., 2007; Karpicke and Smith, 2012). Elaborative processes are thought to boost the encoding of relational information as well as the distinctiveness of specific ideas and facts (Craik and Tulving, 1975; Hunt, 2012), and concept map creation is considered to rely to a large extent on the contribution of such processes to meaningful learning (Novak and Gowin, 1984; Novak and Cañas, 2006). Indeed, over the course of recent years, concept maps, which are node-and-link diagrams that hierarchically represent relevant concepts and relations among them within a knowledge domain, have gained extraordinary popularity among educators and students (Chevron, 2014; Kinchin, 2014; Karpicke, 2018; Schroeder et al., 2018). Strikingly, studies that have compared retrieval practice (i.e., free recall by writing paragraphs) to concept mapping have mostly failed to show the latter to be superior as a learning tool (Karpicke and Blunt, 2011; Lechuga et al., 2015; O'Day and Karpicke, 2021). By using educational texts as materials, Karpicke and Blunt (2011; see also O'Day and Karpicke, 2021, Exp. 2) showed that repeated free recall produced better learning than creating a concept map (while viewing the text) on a final test 1 week later. Lechuga et al. (2015) essentially replicated this effect, even though they also showed that it was less pronounced (only affected verbatim questions, not inferential questions) for participants with experience in using concept maps than it was for those who undertook a short training in concept mapping.

A matter that has recently begun to draw researchers' attention is the identification of the most effective ways to enhance student learning by combining different learning (including retrieval-based) activities (e.g., see McDaniel, 2023; Roelle et al., 2023, for recent reviews). Thus, for example, Karpicke and Smith (2012) examined whether combining an imagery-based elaborative strategy (the keyword method; McDaniel and Pressley, 1984) with either restudy or retrieval practice improved the learning of uncommon words. The results revealed that practicing retrieval was the most effective strategy to learn new words (as measured on a 1-week delayed final test) regardless of whether it was coupled with the elaborative strategy. In other words, the combination of elaborative encoding with repeated retrieval failed to produce a benefit beyond that of simply practicing retrieval (see Fritz et al., 2007, for a related finding with 12–13 year-old children). However, and of special relevance here, research has highlighted that distinct learning activities can offer unique benefits to the learning process (see McDaniel, 2023; Roelle et al., 2023). While generative learning activities could foster the development of coherent conceptual representations that are well integrated with prior knowledge, retrieval practice activities could promote knowledge consolidation in memory. The idea that generative activities, such as concept mapping, enhance meaningful learning is well supported by a number of studies (see Fiorella and Mayer, 2016). Thus, because concept mapping seems to rely on elaborative processing and enhance the organizational representation of concepts in memory (Novak and Gowin, 1984; Nesbit and Adesope, 2006; Chevron, 2014; see Karpicke, 2018 for a critical analysis), it is thus not surprising that research is being conducted to determine whether the coupling of concept mapping and retrieval practice may be particularly effective to promote long-term learning.

To date, although research on this issue is still limited, retrieval and concept mapping activities have been combined in two different ways. On the one hand, one study examined the effectiveness of engaging students in concept mapping and free recall, as separate activities, within a sequence of learning events. Thus, O'Day and Karpicke (2021; Exp. 2) compared retrieval practice (5-min initial study followed by free recall for 20 min, restudy for 5 min, and free recall for 10 min; 30 min of total time) against a combination of concept mapping and retrieval practice (initial study for 5 min followed by concept mapping for 20 min, free recall for 10 min, restudy for 5 min, and free recall for 10 min; 50 min of total time). One reason to expect that the combination of creating a concept map first and then practicing retrieval might be particularly effective is that it is assumed that concept mapping enhances students' relational representation of concepts within materials. Hence, this organizational representation might subsequently provide more elaborated retrieval structures that would guide search processes in memory and facilitate access to relevant information (either in subsequent retrieval practice periods or in learning tests) (Hunt, 2012; Kahana, 2017; Roelle et al., 2022, 2023; McDaniel, 2023). One week after the learning session, participants took a short answer test comprising verbatim and inference questions. Results revealed that coupling concept mapping and free recall did not improve learning as compared to repeated retrieval alone, despite the fact that the combined condition significantly increased (in 20 min) the total time spent on the learning session. Again, adding an elaborative study strategy (concept mapping in this case) in a sequence of learning activities did not produce a benefit beyond that of simply carrying out retrieval practice.

In other cases, retrieval practice has been combined with concept mapping by embedding the act of recall in the process of creating a concept map. While students typically construct concept maps in the presence of materials (the texts aid them in identifying main concepts and relations; as in O'Day and Karpicke's study, for example), an alternative way of elaborating a concept map is to draw it up in the absence of materials, which necessarily requires students to engage in retrieving information from memory. Interestingly, Blunt and

Karpicke (2014) tested the potential effectiveness of retrieval-based concept mapping across two experiments. In their first one, and after a 5-min reading period, participants practiced retrieval of information from a short science text. Importantly, half of participants did so by writing down as many ideas as they could, whereas the other half were required to recall by creating a concept map. In both cases retrieval was done in the absence of texts and with two periods of practice. One week later, the participants returned to the lab to complete a short-answer learning test that assessed verbatim and inferential knowledge. The results showed performance from both retrieval formats on the final test to be virtually equivalent. In the second experiment, the researchers factorially crossed retrieval format and availability of texts. Again, no difference between the two retrieval formats was observed, even though both concept mapping and writing down ideas in the absence of materials produced better learning than completing the same activities with the materials at hand.

From their findings, Blunt and Karpicke (2014) argued that retrieval-based concept mapping is as effective as freely recalling by writing paragraphs, thus suggesting that the benefit attributed to both the relational and item-specific processing required in concept mapping is redundant with the processing associated with repeated retrieval. Altogether, the results of the above-mentioned studies that combined concept mapping (and also other elaborative strategies such as Fritz et al., 2007; Karpicke and Smith, 2012 did) with retrieval practice would in fact seem to indicate that no additional benefit should be expected beyond the one provided by the very fact of practicing retrieval. Hence, an interim conclusion that might be drawn from the described evidence is that it is the very act of retrieving information that promotes learning the most; accordingly, freely recalling (by writing down, for example) might be the simplest and best choice when it comes to learning from educational texts, and concept mapping (either retrieval-based or not) might not be a worthwhile learning activity to engage in since it is effortful and time-consuming (Quinn et al., 2003; Mintzes et al., 2011). The results of a recent study by Ortega-Tudela et al. (2019), however, suggest a different picture.

In their initial experiment, Ortega-Tudela et al. (2019) replicated the main finding from Blunt and Karpicke (2014), such that paragraph writing (PW) and concept mapping (CM), both as retrieval-based activities, led to similar levels of performance in a 2-week (instead of 1-week) delayed learning test. After successfully replicating Blunt and Karpicke's main finding, Ortega-Tudela et al. (2019) designed a second experiment to further test the idea that the relational and item-specific processing involved in concept mapping is redundant with retrieval practice. Their prediction was straightforward: If both concept mapping and paragraph writing down, as retrieval-based activities, engage redundant cognitive processing with equivalent learning outcomes (as Blunt and Karpicke, 2014 suggested), then one would expect two groups of participants doing retrieval practice in both formats but performing the activities in a different order (CM + PW vs. PW + CM) to exhibit similar performance on a learning test. One might predict, however, that if the fact of elaborating a concept map provides a unique opportunity for later retrieval [i.e., because of the organizational/relational processing that it is assumed to demand; see O'Day and Karpicke (2021) for a similar hypothesis when it comes to creating a concept map with materials made available], then performance on the final learning test should be dependent on the position that retrieval-based concept mapping takes in the sequence of learning activities. Specifically, it might be expected that the sequence CM + PW would give rise to better learning than PW + CM. This was, precisely, the main result observed by Ortega-Tudela et al. (2019). On average, those participants who first elaborated concept maps outperformed the participants who first retrieved ideas by writing them down as paragraphs, with this difference not being attributable to the quality of the concept maps created by participants from both groups. Interestingly, these results are in line with what Roelle et al. (2023) call "construction-before-consolidation," suggesting the potential advantages of practicing generative learning prior to retrieval practice activities (see also Roelle and Nückles, 2019). Nevertheless, results from other recent studies seem to favor retrieval-practice-first sequences (e.g., Roelle et al., 2022), highlighting the need for further exploration. As suggested by Roelle et al. (2023), there may not be a singular optimal learning sequence, and the effects of generative-first or retrieval-practice-first sequences could be influenced by moderating variables. In any case, the sequencing effect obtained by Ortega-Tudela et al. (2019) might be of particular interest in gaining an understanding of effective combinations of learning strategies and merits further investigation. Although it stands in stark contrast to results from other (still scarce) studies showing that concept mapping does not contribute to retrieval practice as a learning tool (Karpicke and Blunt, 2011; Lechuga et al., 2015; O'Day and Karpicke, 2021), it lends empirical support to the widespread idea that concept mapping relies on cognitive operations that may be especially effective for learning from educational texts (Novak and Gowin, 1984; Chevron, 2014; Karpicke, 2018). In addition, the finding points to the need to further consider the nature of distinct retrieval activities when it comes to boosting learning, and is particularly suggestive of the potential educational value of creating concept maps in the absence of educational texts as a first stage in a sequence of learning activities.

The educational implications of this sequencing effect could be of special relevance in the case of primary and secondary education students, in whom performance on learning tasks has a greater potential for improvement. However, to date, no other studies have addressed the issue of combining and sequencing in respect of retrieval-based concept mapping, aside from the original work by Ortega-Tudela et al. (2019). Thus, in the present study, we aimed to determine whether the sequencing effect observed in Ortega-Tudela et al. (2019) with college students may also be found in learners attending high school. The replication of the main finding obtained by Ortega-Tudela et al. (2019)—with a sample of younger and less experienced students—in a more naturalistic school setting would contribute significantly to identifying how retrieval-based activities (especially concept mapping) may be used more effectively and in distinct educational contexts.

While we expected to find the above-mentioned sequencing effect in 16–17-year-old students, we also considered that several factors could modulate such an effect in the present study. Firstly, episodic memory functioning and its neural underpinnings undergo changes throughout adolescence and young adulthood, particularly regarding strategic components of encoding and retrieval (Shing et al., 2008, 2010; Mechie et al., 2021). Thus, we could anticipate limited benefits of elaborating concept maps first in our sample due to the lower success rates during retrieval practice. Furthermore, as self-regulation abilities improve with age across the middle and high school grades, so does the spontaneous deployment of retrieval practice as a learning

activity (e.g., Tullis and Maddox, 2020). Consequently, the benefit of concept mapping first observed among college students may be less prominent in younger learners. On the other hand, conducting the study in a naturalistic educational context (the entire research process was conducted in classrooms as part of academic activities) poses important technical challenges that could reduce the sensitivity to observe experimental effects. However, we took advantage of an outstanding opportunity to conduct research beyond the laboratory because the Spanish Education System includes a cross-curricular competence in the 11th and 12th grades curricula that focuses on 'learning to learn', which allowed us to more easily justify our presence in the classrooms. Thus, following closely the procedure used in the second study experiment from Ortega-Tudela et al. (2019), but embedding it into the curricular content of one academic activity oriented to enhance the students' skills in preparing for exams, in the present study we directly compare the effect of the sequence of retrieval activities on a learning test taken 2 weeks later. If creating a concept map first in the absence of texts facilitates the retrieval of relevant isolated concepts, with a focus on relations among the reached concepts, thus increasing their accessibility and organization over future retrieval attempts, then better learning should be observed in students who are able to create concept maps of sufficient quality as the first retrieval activity. Although research among young students might not support the preference for concept mapping as a learning tool, primary and secondary education students as well as adults have been shown to benefit from concept mapping training and guidance (Habók, 2012; Gallenstein, 2013; Roessger et al., 2018). Thus, in the present study a brief training in concept mapping was provided to ensure understanding and practice on why and how to elaborate concept maps. Hence, our prediction was that adolescents attending high school would also benefit more from the sequence CM + PW than from the sequence PW + CM.

In addition to performance (either on the retrieval practice periods or the learning test), we assessed participants' own outcome expectations as well as academic self-concept and overall use of learning strategies, in order to control for their potential effect on final learning. Because the format of the retrieval activity and prior learning strategy experience may bias participants' metacognitive judgements and expectations about their final performance, we asked them to provide judgments of learning immediately following their completion of the first retrieval trial, in any format. Although retrieval-based activities generally do not generate high expectations for learning outcomes (Rivers, 2021), we intended to explore whether such judgments were reliant on the type of retrieval activity that learners performed. Additionally, we included a short questionnaire to assess academic self-concept due to its reciprocal relationships with the use of learning strategies and achievement (McInerney et al., 2012). Finally, we also assessed the frequency with which participants utilized various study techniques to control for potential differences between the groups in this regard (Miyatsu et al., 2018).

2 Method

2.1 Participants

The study was conducted at a public high school located in a middle-class neighborhood in the city of Jaén (Spain), as part of a school time activity organized by the Educational Orientation Department. Partial information about the study was first presented to 101 students from the first and second years of the bachelor's degree in the Spanish educational system (Grades 11 and 12), who were participating at that moment in a series of academic activities on learning strategies (as part of the transversal curricular content promoted by the school)[1]. Ultimately, 64 students agreed to participate in the study. However, we detected 4 outliers (two participants from each experimental group) who were no longer considered for analyses. These were students who produced less than two ideas (out of 25) from the studied text in any of the two retrieval conditions during the first practice trial. Thus, the final sample consisted of 60 participants. None of them had specific learning difficulties. The school's team approved the activities carried out in the study as part of the academic program, and the University's Research Ethics Committee approved the study. Informed consent and assent form were obtained prior to participation.

2.2 Materials and procedure

The experiment was carried out over three independent sessions (one training session and two assessment sessions) with a delay of 14 days between them. In the first session, participants were informed in general about the study and required to provide their informed consent to participate in the activities.

Before the training session started, the academic self-concept scale and the frequency of using study strategies were completed. An academic self-concept scale validated for use with Spanish-speaking adolescents was selected (Academic Self-Concept for Adolescents, ASCA; Ordaz-Villegas et al., 2013). The ASCA scale consists of 28 Likert items ranging from 1 (never) to 5 (always), loading on four different factors: self-regulation, general intellectual abilities, motivation, and creativity. With the original validation sample ($n = 347$; ranging 14–18 years old), Ordaz-Villegas et al. (2013) reported a global Cronbach's alpha of 0.828. In our sample, the global Cronbach's alpha was 0.750 (specific alphas: self-regulation = 0.555; general intellectual abilities = 0.517; motivation = 0.478, and creativity = 0.569).

The students also filled out a short questionnaire about their use of diverse learning strategies (Hábitos y Técnicas de Estudio, CHTE; Álvarez and Fernández, 1999). The CHTE comprises 12 Likert-type items (scoring frequency of use from 1, never, to 4, always) assessing various student practices. We specifically selected those that directly relate to common study strategies for preparation of subsequent learning test. These include underlining, schema creation,

1 Although we were aware of the difficulty of collecting data from a large sample in a naturalistic setting, the ideal sample size was determined in advance by assuming a medium effect size (like the one observed in Ortega-Tudela et al., 2019). A power analysis performed with G*Power (Faul et al., 2007) determined that a minimal total sample size of 98 participants would be necessary to obtain a statistically significant effect from the sequencing of the retrieval-based learning activities (power=0.80 and α=0.05; ANOVA: Repeated measures, between factor).

summarizing, concept mapping, and self-generation of questions. Only information on frequency of use of these strategies was collected.

After completing the questionnaires, students were provided with a short training on concept mapping in their natural classroom groups (see also Lechuga et al., 2015; Ortega-Tudela et al., 2019). The program consisted of three steps. First, a general explanation of concept maps was given as well as a brief description of their usefulness for students. Then, participants were taught how to elaborate a concept map. They were given a text (about different types of forests) and one of the researchers created a concept map from the text by drawing on the blackboard, emphasizing the various characteristics and possibilities. Finally, the participants were given a second text (on factors to be considered when carrying out an academic task) and, in groups of two, they were instructed to work on creating their own concept map from this particular text. Two of the researchers supervised the students to ensure that the maps were constructed according to the instructions given and the researchers also answered any questions on the process. The students' usual teacher was also involved in supervising and in answering questions. Both texts and procedures used during the training were adapted from Repetto et al.'s (2004) program, as used in Lechuga et al. (2015). Session 1 lasted approximately 55 min.

The second session took place 2 weeks later and lasted about 25 min. Only students who expressed interest in participating attended sessions 2 and 3. Students belonging to the same class performed the study session as a natural group. All of them were encouraged to read a text for 5 min. The text (taken from Repetto et al.'s (2004) program and also used in previous studies: Lechuga et al., 2015; Ortega-Tudela et al., 2019), consisted of 25 ideas or propositions on the use of fibers and the manufacturing of fabrics. This is an expository text with clear organizational structures, educationally relevant content in the natural sciences, and corresponding to the first year of the Spanish Secondary Education (ESO) system (7th grade in the American system).

Subsequently, the students undertook two 10-min retrieval trials in the absence of the previously studied text. Critically, participants were randomly assigned to one of the two experimental conditions involving the same retrieval activities but in reverse order: Paragraph writing first and then concept map (PW + CM) or concept map first and then paragraph writing (CM + PW). This assignment took place upon arrival at the rooms designated for the study, regardless of participants' class affiliation. Both groups were similar in gender distribution and age (Group CM + PW: $n = 32$, 18 women, 3 unidentified; $M_{age} = 16.65$, $SD_{age} = 0.55$; Group PW + CM: $n = 28$, 12 women, 4 unidentified; $M_{age} = 17.00$, $SD_{age} = 0.78$).

Those participating in the PW + CM group were given a blank piece of paper and asked to write down everything they could recollect about the text (paragraph format). After 10 min the pieces of paper were collected and the students had a 1-min break. Then, they were provided with a new blank piece of paper and asked to elaborate a concept map in the way that they learned in the first session. In the CM + PW group, however, the participants first engaged in concept mapping and then they were moved onto the paragraph-based activity. Immediately after the 1-min break, and before taking the second retrieval trial, participants from both groups were given a sheet of paper and asked to make judgments of learning, a subjective estimate on a scale ranging from 0 to 100 of how much they would be able to recall from the text in a final test to be administered 2 weeks later.

The third experimental session took place 2 weeks later. In it, the students took a learning test that comprised 15 short-answer questions to be answered in writing in the space available. Ten of the questions required students to recall content directly presented in the text (verbatim). An example of a verbatim question is, "What is the origin of the natural fibers that are used to create fabrics?," with "animal and vegetables" being the correct answer that explicitly appeared in the text. The other five questions elicited inferences about the studied content (inference). An example of an inference question is, "Why are China and Egypt good countries for cotton cultivation?," with "Due to the hot and humid climate" as the correct answer. Different sentences from the text must be integrated in order to answer the question ("Thousands of years ago, the Chinese and the Egyptians knew the method of making fiber from the hairs of cotton seeds." and "Most cotton is grown in countries that have a hot, humid climate"). The students were given 15 min to answer the questions.

The responses to the learning test were scored by using a correction template containing the ideas of the text in addition to the acceptable responses for each question. This template has been widely used across studies and made the scoring process straightforward, highly reliable and free of ambiguity (see Lechuga et al., 2015; Ortega-Tudela et al., 2019). A graduate Psychology student was briefly trained in the application of the template on the basis of examples of hits and errors from previous studies. This student, who was blind to all aspects of the study, scored the responses to verbatim and inferential questions from the learning test (up to 1 point per question). One of the authors (J.M.O-T.) supervised the correct application of the template.

As in related studies (Lechuga et al., 2015; Ortega-Tudela et al., 2019), two judges (graduate students who were also totally blind to the aims of the study) were asked to score the number of ideas from the text (1 point per unit) that were recalled during the retrieval trials in either format (paragraphs or concept maps). They also used the aforementioned template with the 25 ideas from the text to score the responses. Since this scoring procedure did not lead to significant disagreements between the two raters (see also Lechuga et al., 2015; Ortega-Tudela et al., 2019), it was decided to proceed collaboratively after scoring a few participants. Thus, both judges scored each participants' response face-to-face, and a consensus on scoring was reached when disagreements arouse (less than 1% of responses). Finally, these two judges rated the quality of the participants' concept maps on a 10-point scale regarding 5 dimensions: number of represented primary ideas, number of represented secondary ideas, hierarchical structure, use of connectors, and concept organization. The global intraclass correlation coefficient (ICC) between raters was 0.844. The ICCs for the dimensions were: primary ideas = 0.758, secondary ideas = 0.796, hierarchical structure = 0.759, use of connectors = 0.694, and concept organization = 0.753.

3 Results

Mediation analyses, independent samples *t*-tests, analyses of variance, and correlation/regression analyses were conducted on the data from the self-reported (use of learning strategies, academic self-concept, and judgments of learning; Table 1), and performance measures (ideas produced at retrieval practice and correct responses on the learning test; Table 2; and quality of concept maps; Table 3). All

TABLE 1 Mean proportion (and standard deviations) of ideas produced in each trial of retrieval and correct responses in the final test as a function of retrieval sequence.

Retrieval sequence	Retrieval practice		Learning (final) test			
	Trial 1	Trial 2	Overall	Verbatim	Inferential	Overall
PW + CM	0.34 (0.17)	0.22 (0.10)	0.28 (0.13)	0.17 (0.12)	0.20 (0.21)	0.18 (0.13)
CM + PW	0.24 (0.09)	0.30 (0.18)	0.27 (0.13)	0.24 (0.17)	0.26 (0.20)	0.25 (0.16)
Overall	0.28 (0.14)	0.26 (0.15)		0.21 (0.15)	0.23 (0.21)	

CM, concept mapping; PW, paragraph writing. For the PW + CM group, Trial 1 refers to PW and Trial 2 refers to CM. For the CM + PW group, Trial 1 refers to CM and Trial 2 refers to PW.

TABLE 2 Mean scores (and standard deviations) of self-reported measures (learning strategies, as measured with CTHE; academic self-concept as measured with ASCA, and judgments of learning) as a function of retrieval sequence.

		Retrieval sequence			
		PW + CM	CM + PW	*p*	Cohen's *d*
Learning strategies (0–4)	Underlining	3.09 (1.04)	3.32 (1.01)	0.41	0.23
	Schemas creation	2.08 (1.04)	2.29 (1.13)	0.50	0.19
	Summarizing	2.79 (1.25)	2.90 (1.16)	0.73	0.09
	Concept mapping	1.46 (0.65)	1.42 (0.67)	0.83	−0.06
	Self-generating questions	3.17 (0.91)	3.23 (0.88)	0.81	0.07
Academic self-concept (0–20)	Self-regulation	13.04 (2.91)	12.29 (3.19)	0.37	−0.24
	Cognitive ability	14.21 (2.78)	13.68 (2.96)	0.50	−0.18
	Motivation	14.17 (3.73)	13.58 (2.96)	0.52	−0.18
	Creativity	11.65 (2.81)	11.45 (2.34)	0.80	−0.07
Learning judgment (0–100)		38.57 (20.85)	33.75 (19.96)	0.36	−0.24

CM, concept mapping; PW, paragraph writing.

TABLE 3 Mean scores (and standard deviations in a 0–10 scale) in the considered dimensions on the quality of the concept maps as a function of retrieval sequence.

	Retrieval sequence			
Assessed dimensions	PW + CM	CM + PW	*p*	Cohen's *d*
Primary ideas	6.15 (1.90)	6.22 (1.81)	0.88	0.04
Secondary ideas	5.63 (1.78)	5.48 (2.19)	0.80	−0.07
Hierarchical structure	5.85 (1.77)	5.83 (1.79)	0.96	−0.01
Use of connectors	4.71 (2.65)	5.78 (2.40)	0.59	0.15
Conceptual organization	6.04 (1.74)	6.09 (1.51)	0.91	0.03
Overall quality (mean score)	5.67 (1.79)	5.74 (1.71)	0.89	0.04

CM, concept mapping; PW, paragraph writing.

analyses were performed by using JASP software (JASP team, 2023, version 0.17.3). The full data set is available at https://osf.io/82acz/.

We report analyses of performance on retrieval activities (paragraph writing and concept mapping) at the group level and their potential association with performance on the learning test. Subsequently, we report analyses aimed at determining whether the sequence of retrieval-based activities had a direct effect on learning outcomes in the 2-week delayed test and whether this effect was mediated by retrieval success during the first retrieval activity. Finally, we describe the results of analyses aimed at discovering: (a) whether students' judgments of learning were predictive of actual performance on the learning test; and (b) whether the use of learning strategies and the academic self-concept modulated performance either in the retrieval activities or in the learning test.

3.1 Success during retrieval practice

A mixed analysis of variance (ANOVA) on the proportion of ideas produced (out of the total of 25) during retrieval practice was performed with retrieval sequence (PW + CM vs. CM + PW) as the between-participants factor and retrieval period as the within-participant factor (trial 1 vs. trial 2). Neither sequence [$F(1, 58) < 1$, $p\eta^2 < 0.01$] nor retrieval trial reached statistical significance, [$F(1, 58) = 3.37$, $p = 0.07$, $p\eta^2 = 0.05$]. However, the interaction retrieval sequence × trial did, [$F(1, 58) = 38.08$, $p < 0.01$, $p\eta^2 = 0.40$]. Follow-up analyses revealed that during trial 1 the PW + CM group showed better performance ($M = 0.34$; $SD = 0.16$) than the CM + PW group [$M = 0.24$; $SD = 0.10$; $F(1, 58) = 8.38$, $p < 0.01$, $p\eta^2 = 0.13$]. During trial 2, it was the CM + PW group that produced more ideas ($M = 0.30$; $SD = 0.18$) than the PW + CM group [$M = 0.22$; $SD = 0.10$; $F(1, 58) = 4.75$, $p = 0.03$, $p\eta^2 = 0.08$]. In other words, writing paragraphs led participants to recall more ideas than concept mapping regardless of the order in which activities were done (see Table 1). No reliable difference between the groups emerged either when writing

paragraphs or when elaborating concept maps (both with $F < 1$ and $p\eta^2 < 0.01$).

We also examined whether performance during the retrieval practice periods predicted final learning considering the whole set of participants. First, we wanted to learn whether overall success during retrieval practice (collapsing performance across retrieval periods) correlated with learning outcomes (averaging performance in both types of test questions) 2 weeks later. This analysis revealed a reliable positive Pearson correlation ($r = 0.64$, $p < 0.001$). Then, we examined correlations for each type of retrieval activity. Both retrieval success in paragraph writing ($r = 0.62$, $p < 0.001$) and concept mapping ($r = 0.58$, $p < 0.001$) predicted learning outcomes. In addition, we conducted separate stepwise multiple regression analyses over both types of test questions to identify the best predictor in each case. Interestingly, the analyses revealed that success in paragraph writing periods uniquely predicted performance [31% of variance; $t(59) = 5.22$, $p < 0.01$] on verbatim questions, whereas success in concept mapping periods uniquely predicted performance [29% of variance, $t(59) = 4.95$, $p < 0.01$] on inferential questions.

3.2 Effect of sequence of retrieval activities on learning test performance

Since the main goal of the present study was to determine whether retrieval sequence had an effect on students' learning, and following the main analysis performed by Ortega-Tudela et al. (2019), we firstly conducted the originally-planned mixed ANOVA on correct responses in the learning test with retrieval sequence (between-groups) and type of question (within-participant) as factors. The analysis failed to show statistically significant effects [retrieval sequence: $F(1,58) = 2.55$, $p = 0.12$, $p\eta^2 = 0.04$; question type: $F(1,58) < 1$, $p\eta^2 = 0.01$; interaction: $F(1,58) < 1$, $p\eta^2 < 0.01$]. Importantly, however, the fact that in the present study with adolescents (a) the two retrieval activities had different retrieval success rates, with paragraph writing leading to the production of more ideas, especially in trial 1, which is an important deviation from the results previously found with college students, and that (b) success during retrieval practice was predictive of final learning, clearly pointed to the need of considering performance in the first trial of retrieval practice as a relevant factor when accounting for learning in the final test. Differences in success rates (stemming from distinct factors) during retrieval practice may modulate learning outcomes (e.g., Minear et al., 2018; Leggett et al., 2019; Racsmány et al., 2020; Ortega-Tudela et al., 2021; see also Karpicke, 2017). Specifically, in the present study elaborating retrieval-based concept maps first seemed to have very different effects on retrieval success (fewer ideas produced during retrieval practice) and performance on the delayed learning test (more correct responses), but both retrieval sequence (more benefit from concept mapping first) and retrieval success (more benefit from paragraph writing first) might have opposing effects on learning. If so, failing to account for such facing effects could make it difficult to observe differences between the two sequences in the final test. In other words, the expected direct effect of sequencing the retrieval activities may have been masked by the indirect larger effect of retrieval success during paragraph writing. To deal with this possibility, we conducted a mediation analysis to examine whether the lack of effect of retrieval sequence was mediated by the effect of success during retrieval practice on performance on the learning test. If this were the case, a reliable direct effect of retrieval sequence should be found after accounting for the mediating (indirect) effect of retrieval practice success on learning. The mediation analysis was conducted by introducing retrieval sequence (CW + PW vs. PW + CW) as the predictor, success during retrieval practice as the mediator, and correct responses to verbatim and inferential questions as the outcomes. Statistical significance was tested by using 5,000 (bias-corrected percentile) bootstrap samples. The analysis confirmed direct and indirect effects. As can be seen in Figure 1, retrieval sequence had a direct effect on the proportion of correct responses to both verbatim (c'_1: $B = -0.12$ [95% CI, −0.19, −0.05], $p = 0.001$, $\beta = -0.77$) and inference questions (c'_2: $B = -0.13$, 1.12 [95% CI, −0.23, −0.03], $p = 0.01$, $\beta = -0.63$) that was independent of success during retrieval practice. Thus, CM + PW led to better learning scores than PW + CM. There was also a (smaller) indirect effect that was mediated by success during the first retrieval trial on verbatim (ab_1: $B = 0.05$ [95% CI, 0.02, 0.10], $p = 0.02$, $\beta = 0.34$) and inference questions (ab_2: $B = 0.07$ [95% CI, 0.02, 0.14], $p = 0.02$, $\beta = 0.34$). This mediation effect indicates that the benefit of paragraph writing for learning relies on its capability to facilitate retrieval, although this effect was smaller than the direct effect produced by concept mapping first. Confirming the differential influence of concept mapping and paragraph writing as first retrieval activities on learning, the direct and indirect effects had estimates of a different sign, which accounts for the lack of a statistically significant total effect (please, note that total effect = direct effect + indirect effect) of retrieval sequence on correct responses to verbatim (c_1: $B = -0.07$ [95% CI, −0.14, 0.01], $p = 0.08$, $\beta = -0.44$) and inferential questions (c_2: $B = -0.06$ [95% CI, −0.16, 0.05], $p = 0.25$, $\beta = -0.29$).

3.3 Learning strategies and academic self-concept

Only the self-reported frequency of use of two learning strategies showed to be predictive of successful performance. On the one hand, greater frequency of use of underlining was shown to be inversely correlated with the proportion of ideas recalled when writing paragraphs ($r = -0.34$, $p = 0.01$). On the other hand, summarizing negatively correlated with the number of ideas recalled either when writing paragraphs ($r = -0.37$, $p < 0.01$) or creating concept maps ($r = -0.40$, $p < 0.01$), and with the overall score on the final learning test ($r = -0.36$, $p < 0.01$)[2].

2 Since underlining and summarizing are considered scarcely valuable learning techniques (although students may have the ability to implement them effectively, they do not always do so; Dunlosky et al., 2013), these associations might be simply reflecting individual differences in motivation or general abilities to engage in academic tasks (such as the ones being required in our study). Hence, those participants who reported more use of underlining and summarizing as study techniques could have devoted less attention while reading the text or when told to think back on its main ideas during the retrieval periods. Because there were no differences between the groups in the use of these techniques, this finding is not further discussed here.

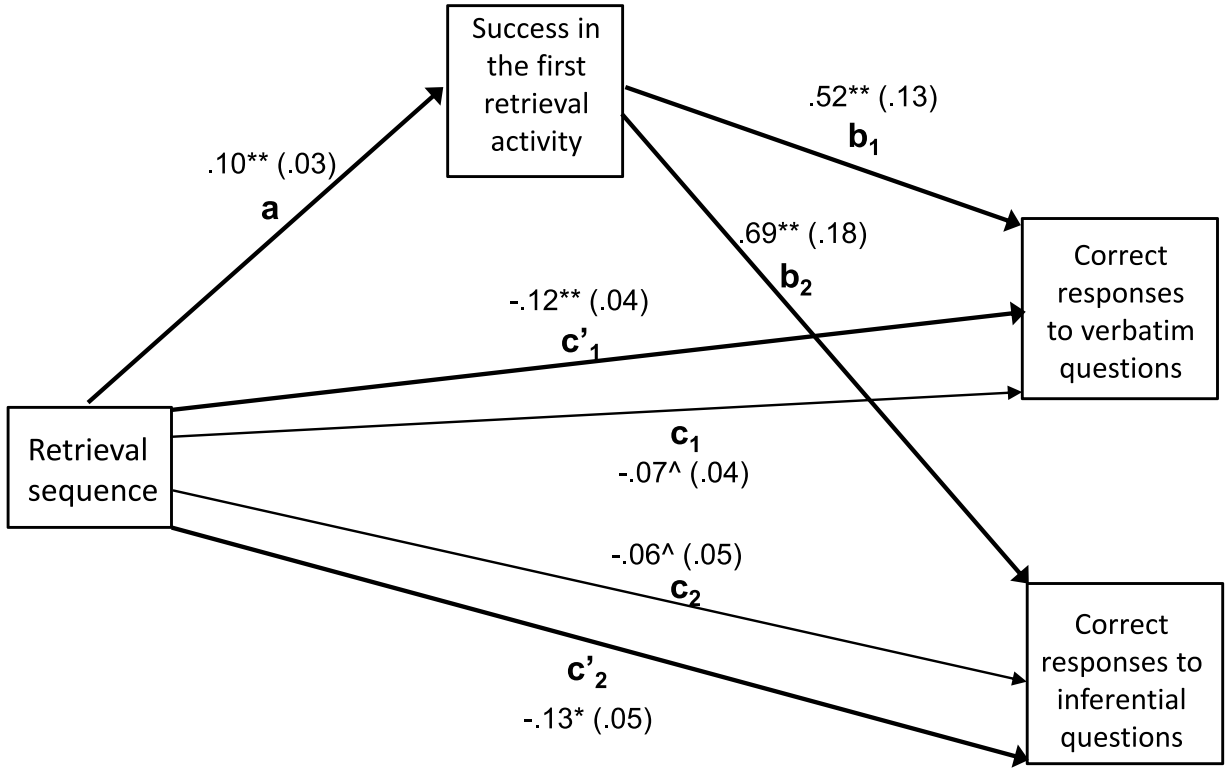

FIGURE 1
Path model diagram with the results of the mediation analysis. Total effect (paths c_1 and c_2) represents the effect of retrieval sequence on final test performance with no mediator in the model. Direct effect (paths c'_1 and c'_2) represents the effect of retrieval sequence on performance that is independent of retrieval success as a mediator in the model. Indirect effect (paths ab_1 and ab_2) represents the effect of retrieval on performance that is mediated by retrieval success. The figure shows unstandardized coefficients (and standard errors). See main text for further information. *p = 0.01, **$p < 0.01$, ^ non-significant.

No associations emerged between any of the dimensions of the ASCA scale and student performance either in the retrieval activities or in the learning test (all Pearson coefficients with $p > 0.19$). For the sake of completeness, correlations between the self-reported measures are reported in Supplementary material, although they will be not further discussed.

Table 2 reports means and standard deviations of learning strategies and the academic self-concept scales for both groups of participants. Two-tailed t-tests for independent samples failed to show differences between the two groups in frequency of use of learning strategies. While concept mapping was reported as the least used study strategy, underlining and self-generation of questions were the strategies that students reported using the most. No group differences were also observed when it came to responding to items regarding the different dimensions included in the ASCA scale (see Table 2).

3.4 Judgment of learning

A bivariate Pearson correlation between the judgments of learning made by the students after the first retrieval activity and their overall performance on the (2-weeks delayed) learning test showed no reliable association ($r = 0.22$, $p = 0.08$; partial correlation after controlling for overall success at retrieval practice: $r = 0.07$, $p = 0.60$). A two-tailed t-test failed to show a difference in the judgments of learning made by the groups (see Table 2).

3.5 Quality of the concept maps

Two-tailed t-tests for independent samples showed that there were no differences between both groups in any of the dimensions considered when assessing the quality of the concept maps (see Table 3). Finally, we also examined whether the quality of concept maps (overall score) was predictive of performance in the final test. Pearson correlation analyses showed positive and reliable associations between overall quality and proportion of correct responses to verbatim ($r = 0.37$, $p < 0.01$) and inferential questions ($r = 0.32$, $p = 0.02$).

4 Discussion

With the aim of further investigating the potential benefit of elaborating a concept map (in the absence of texts) as an initial learning activity within a series, in the present study, we assessed the effect of the sequence of such retrieval-based activities in a sample of high school students. While there is an increasing interest in determining the effectiveness of distinct combinations of learning

activities (with an emphasis on those relying on retrieval practice to boost outcomes), to the best of our knowledge no previous study has examined the effect of such combinations on learners other than college students.

Extending previous findings among college students, our results with high-school students support the idea that conceptual learning from educational texts can be modulated by the sequence of retrieval-based activities performed. Specifically, these results reveal that creating a concept map in the absence of educational texts followed by free recall by writing paragraphs led to better responses to verbatim and inferential questions (on a learning test that took place 2 weeks later) than performing the same activities in the reverse order. Additionally, our results indicate that this effect is not related to prior experience with the use of study techniques or academic self-concept. Hence, that specific benefit, originally observed in college students in a lab setting (Ortega-Tudela et al., 2019), has been shown to be generalizable in a more naturalistic educational setting to learners who are less experienced as students than college adults and who were, in fact, less successful at retrieving relevant ideas from the text and less skilled at elaborating concept maps[3]. A few aspects of these results merit particular consideration and will be discussed below.

Aside from the age of the participants, a remarkable difference between the present study and the one with college learners (Ortega-Tudela et al., 2019) is that the set of research-related tasks (concept mapping training, retrieval practice, and learning assessment) were all implemented in the classroom as part of a series of academic activities oriented to enhance the students' skills in preparing for exams. Hence, all these activities were carried out in a genuine educational setting[4]. Despite these differences with the original study, our results largely replicate the finding of better learning outcomes resulting from the sequence that involves concept mapping first, which suggests that such a sequencing effect might work across a variety of learners and contexts. Although the latter is an empirical question that would be worth addressing in future studies, the present results contribute to this idea.

It has turned out to be of special practical significance that we observed the benefit of concept mapping first with learners who reported that they hardly ever created concept maps to deal with exams (concept mapping was reported as the least used learning strategy by participants) and who, as a matter of fact, elaborated concept maps of rather low-moderate quality during their retrieval periods in our study. Thus, from our own findings with adolescents it would seem that the value of retrieval-based concept mapping does not depend on expertise in concept mapping. However, we acknowledge that as concept mapping is a developmental skill that improves with repeated practice and feedback (Roessger et al., 2018), the potential benefits of retrieval-based concept mapping as a learning activity may increase with training and practice. This is also an important question to address in future studies. Studies that have compared concept mapping with other learning strategies (e.g., retrieval practice) have typically introduced concept map creation as a one-off activity (Karpicke and Blunt, 2011; O'Day and Karpicke, 2021) which means that learners did not necessarily master the required skill, which could have minimized the potential benefits of concept mapping (see Lechuga et al., 2015).

Instead, and also of practical relevance, the benefit of retrieval-based concept mapping would seem to depend on the nature of the retrieval activity that follows in the sequence of learning tasks. As Blunt and Karpicke (2014, Exp. 2; see also Exp. 1 from Ortega-Tudela et al., 2019) showed, two consecutive periods of concept mapping in the absence of materials (i.e., CM + CM) failed to improve learning as compared to either open-book concept mapping or repeated free recall by writing paragraphs (i.e., PW + PW). Therefore, it seems to be the specific configuration of concept mapping plus free recall by writing (in this order) that it has been shown to be particularly effective in improving learning outcomes. It remains to be explored whether other learning activities (retrieval-based or not) beyond free recall might also make a difference when it comes to learning from educational materials. In addition, it would also be necessary to show that CM + PW improves learning more than PW + PW, given the difficulty of constructing concept maps with closed books (especially for adolescents and children) and the apparent ease of writing paragraphs. While an indirect (cross-experiment) comparison between both sequences is possible from the data of Ortega-Tudela et al. (2019) with younger adults, which reveals better learning after CM + PW than after PW + PW, future studies should examine this issue more closely. In any case, the present results, together with previous findings, argue against the general idea that all retrieval-based activities are similarly effective simply because they involve the fact of recalling (Blunt and Karpicke, 2014; O'Day and Karpicke, 2021). Rather, different retrieval demands may interact with each other and with other cognitive processes underpinning the specific learning task at hand to modulate knowledge acquisition and long-term retention. Creating a concept map without materials might promote a more conceptually-guided retrieval orientation and the generation of more elaborated mental models that, in turn, would facilitate the adoption of more effective retrieval strategies at subsequent attempts to access relevant knowledge (Ortega-Tudela et al., 2019). This interpretation of the benefit of CM + PW over PW + CM as retrieval activities would align with the elaborative retrieval account (Carpenter, 2009, 2011; see also Hinze et al., 2013). According to this view, when retrieval attempts are demanding and the learners have to get engaged in reconstructive process reliant on additional information (e.g., during concept mapping in the absence of texts), these additional pieces of information are then integrated with existing memory traces (Carpenter, 2011). In other words, elaborated memories may be created during demanding retrieval episodes that might help to re-organize or supplement initially encoded information, thus facilitating future access to these memories. Nevertheless, we would like to emphasize that these findings do not

3 Cross-experiment two-tailed *t*-tests confirmed that the (adult) participants in Ortega-Tudela et al.'s study (Exp. 2): (a) recalled more ideas in the first retrieval activity (CM+PW: $M=0.45$; $SD=0.09$; PW+CM: $M=0.42$; $SD=0.09$) than the (adolescent) participants in the present study (CM+PW: $M=0.24$; $SD=0.09$; PW+CM: $M=0.34$; $SD=0.17$), both tests with $p<0.01$, and (b) elaborated concept maps of overall better quality ($M=6.60$; $SD=1.45$) than the younger participants ($M=5.71$; $SD=1.73$), $p<0.001$.

4 Closely following the definition of classroom research considered by Agarwal et al. (2021) in their recent meta-analysis on retrieval practice in educational settings, in the present study the materials and activities to be performed on them were directly related to the assigned course materials. In addition, retrieval practice was individual, not collaborative, so that all participants engaged in either concept mapping or free recall individually under the supervision of instructors. Finally, in all cases retrieval-based activities were performed without the use of external learning aids.

necessarily indicate a general superiority of performing elaborative/generative learning tasks first within a sequence. As a matter of fact, a notable difference between the present study and previous research with young adults is that high school students recalled fewer ideas when using concept mapping as a retrieval task compared to paragraph writing. This suggests that concept mapping with closed books may be a challenging retrieval task for younger students. It is worth noting that adults showed similar retrieval success across retrieval formats in Ortega-Tudela et al. (2019). Furthermore, in the present study, success in the first retrieval activity played a role as a mediator in the relationship between retrieval sequence and learning scores. This implies that, in addition to the inherent benefit of creating concept maps first, which was evident from the reliable direct effect of the retrieval sequence favoring CM + PW over PW + CM, the rate of success during retrieval practice may modulate its effectiveness. While this fact does not come as a surprise (e.g., Leggett et al., 2019), from a more practical perspective it points to the need to consider the utilization of procedures (e.g., providing hints during concept mapping) that ensure acceptable success rates for students facing challenging retrieval tasks. Additionally, there is evidence that having learners engaged in retrieval practice (without elaborative components) before engaging in a generative learning activity can also lead to better learning outcomes (Roelle et al., 2022). Hence, the main question to be answered is not which specific learning sequence is most effective, but under which conditions a specific sequence can deliver superior learning outcomes. Further research is clearly necessary on this issue.

Finally, an unexpected, but not surprising, finding in the present study was a retrieval activity-based dissociation when it came to responding to factual and inference questions 2 weeks after doing retrieval practice. While it was performance during the periods of paragraph writing that better predicted the proportion of correct responses to verbatim questions on the final test, retrieval success during concept mapping was a better predictor of correct responses to inference questions (thought to demand more elaborative processes) on the learning test. Indeed, this finding is suggestive of the elaborative/relational nature of the cognitive processing that underlies concept mapping, with particular consequences for learning, and is consistent with results from recent studies that compared the effects of generative learning and retrieval practice (e.g., Roelle and Nückles, 2019). The benefit of generative/elaborative tasks (in the present case, concept mapping) is thought to be reliant on the quantity or quality of mental processes that make sense of the information provided, and one would expect such a benefit to be particularly evident in inference or comprehension questions (Fiorella, 2023). Nevertheless, our participants did not appear to be aware of the potential impact of retrieval-based concept mapping as an initial learning activity, as both groups made similar judgments of learning but had different learning outcomes.

5 Limitations and recommendations

A limitation of our results, even when converging with previous findings concerning students of a different nature, is the reliance on the restricted educational materials worked on by the students. It would be of practical relevance to determine in future studies the effectiveness of initial retrieval-based concept mapping with more varied and curriculum-related materials. A second limitation is that the present study included a relatively small sample of students, so that the present results should be interpreted with caution even when they are in line with results from previous findings. Naturalistic studies usually come at the expense of reduced number of participants even when ideal sample size was estimated in advance. In addition, we recognize that our participants' prior knowledge of the subject matter of the learning text was not assessed. While the random assignment of students to experimental conditions is expected to control for prior knowledge, differences between groups on this variable may still be possible. A fourth limitation concerns the quality of the concept maps. Although we assessed a number of relevant dimensions, we did not consider the number of connections that a given concept had with other concepts, which has been shown to be a relevant parameter mediating immediate and delayed recall (Roessger et al., 2021). Finally, our study only focused on the students. However, given the role that teachers may play when designing and implementing retrieval-based activities in the classroom (Ortega-Tudela et al., 2021), considering their expectations about the effectiveness of these activities as well as their feelings on the benefit–cost ratio after implementing them would contribute to our understanding of the practical value of such activities.

6 Conclusion

Our study with high school students provides evidence, consistent with the results from some previous studies, for the importance of learning tasks that promote the construction of coherent mental representations of to-be-learned materials (Roelle et al., 2023). Specifically, the present findings lend support to the benefit of concept mapping as a retrieval-based learning activity and should help instructors and students to think of tests as knowledge modifiers and learning rather than just as assessment tools (for recent approaches see Bae et al., 2019; Roelle et al., 2022). Specifically, the present results contribute to the identification of an effective combination of retrieval practice activities that may be implemented among high school students with little experience of concept maps. Future studies should identify potential factors that may moderate the efficacy of such a combination.

Data availability statement

The original contributions presented in the study are publicly available. This data can be found here: https://osf.io/82acz/.

Ethics statement

The study was approved by Comité de Bioética de la Universidad de Jaén and conducted in accordance with the local legislation and institutional requirements. Written informed consent for participation in this study was provided by the participants' legal guardians/next of kin.

Author contributions

ML: Conceptualization, Data curation, Investigation, Methodology, Writing – original draft, Writing – review & editing. JO-T: Conceptualization, Data curation, Funding acquisition, Investigation, Methodology, Resources, Writing – review & editing. CG-A: Conceptualization, Data curation, Formal analysis, Funding acquisition, Methodology, Writing – original draft, Writing – review & editing.

Funding

The authors declare that financial support was received for the research and publication of this article from Consejería de Economía, Conocimiento, Empresas y Universidades from Junta de Andalucía, Spain [Programa Operativo FEDER Andalucía 2014–2020. Proyectos_FEDER_1262768].

Acknowledgments

We thank the staff and children from the IES Fuente de la Peña (Jaén, Spain) for their implication and help.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The author(s) declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

Publisher's note

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/feduc.2024.1287744/full#supplementary-material

References

Adesope, O. O., Trevisan, D. A., and Sundararajan, N. (2017). Rethinking the use of tests: a meta-analysis of practice testing. *Rev. Educ. Res.* 87, 659–701. doi: 10.3102/0034654316689306

Agarwal, P. K., Nunes, L. D., and Blunt, J. R. (2021). Retrieval practice consistently benefits student learning: a systematic review of applied research in schools and classrooms. *Educ. Psychol. Rev.* 33, 1409–1453. doi: 10.1007/s10648-021-09595-9

Álvarez, M., and Fernández, R. (1999). *Cuestionario de Hábitos y Técnicas de Estudio (CHTE)*. Madrid: TEA.

Bae, C. L., Therriault, D. J., and Redifer, J. L. (2019). Investigating the testing effect: retrieval as a characteristic of effective study strategies. *Learn. Instr.* 60, 206–214. doi: 10.1016/j.learninstruc.2017.12.008

Blunt, J. R., and Karpicke, J. D. (2014). Learning with retrieval-based concept mapping. *J. Educ. Psychol.* 106, 849–858. doi: 10.1037/a0035934

Carpenter, S. K. (2009). Cue strength as a moderator of the testing effect: the benefits of elaborative retrieval. *J. Exp. Psychol. Learn.* 35, 1563–1569. doi: 10.1037/a0017021

Carpenter, S. K. (2011). Semantic information activated during retrieval contributes to later retention: support for the mediator effectiveness hypothesis of the testing effect. *J. Exp. Psychol. Learn.* 37, 1547–1552. doi: 10.1037/a0024140

Chevron, M. P. (2014). A metacognitive tool: theoretical and operational analysis of skills exercised in structured concept maps. *Persp. Sci.* 2, 46–54. doi: 10.1016/j.pisc.2014.07.001

Craik, F. I., and Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. *J. Exp. Psychol. Gen.* 104, 268–294. doi: 10.1037/0096-3445.104.3.268

Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J., and Willingham, D. T. (2013). Improving students' learning with effective learning techniques: promising directions from cognitive and educational psychology. *Psychol. Sci. Publ. Int.* 14, 4–58. doi: 10.1177/1529100612453266

Faul, F., Erdfelder, E., Lang, A. G., and Buchner, A. (2007). G* power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav. Res. Methods* 39, 175–191. doi: 10.3758/BF03193146

Fiorella, L. (2023). Making sense of generative learning. *Educ. Psychol. Rev.* 35:50. doi: 10.1007/s10648-023-09769-7

Fiorella, L., and Mayer, R. E. (2016). Eight ways to promote generative learning. *Educ. Psychol. Rev.* 28, 717–741. doi: 10.1007/s10648-015-9348-9

Fritz, C. O., Morris, P. E., Acton, M., Voelkel, A. R., and Etkind, R. (2007). Comparing and combining retrieval practice and the keyword mnemonic for foreign vocabulary learning. *Appl. Cognit. Psych.* 21, 499–526. doi: 10.1002/acp.1287

Gallenstein, N. (2013). Concept mapping for learners of all ages. *J. Educ. Teach. Train.* 4, 59–72.

Habók, A. (2012). Evaluating a concept mapping training programme by 10- and 13-year-old students. *Int. Electr. J. Element. Ed.* 4, 459–472.

Hinze, S. R., Wiley, J., and Pellegrino, J. W. (2013). The importance of constructive comprehension processes in learning from tests. *J. Mem. Lang.* 69, 151–164. doi: 10.1016/j.jml.2013.03.002

Hunt, R. R. (2012). Distinctive processing: the co-action of similarity and difference in memory. *Psychol. Learn. Mot.* 56, 1–46. doi: 10.1016/B978-0-12-394393-4.00001-7

Kahana, M.J. (2017). *Memory search, in learning and memory: a comprehensive reference*, Oxford, England: Academic Press, *2*, 181–200

Karpicke, J. D. (2017). "Retrieval-based learning: A decade of progress, in cognitive psychology of memory" in *Learning and memory: a comprehensive reference*. ed. J. Wixted (Oxford: Elsevier)

Karpicke, J. D. (2018). "Concept mapping" in *The SAGE encyclopaedia of educational research, measurement and evaluation*. ed. B. Frey (Thousand Oaks, CA: SAGE Publishing. Inc.), 351–354.

Karpicke, J. D., and Blunt, J. R. (2011). Retrieval practice produces more learning than elaborative studying with concept mapping. *Science* 331, 772–775. doi: 10.1126/science.1199327

Karpicke, J. D., and Smith, M. A. (2012). Separate mnemonic effects of retrieval practice and elaborative encoding. *J. Mem. Lang.* 67, 17–29. doi: 10.1016/j.jml.2012.02.004

Kinchin, I. M. (2014). Concept mapping as a learning tool in higher education: a critical analysis of recent reviews. *J. Cont. High. Ed.* 62, 39–49. doi: 10.1080/07377363.2014.872011

Lamotte, M., Izaute, M., and Darnon, C. (2021). Can tests improve learning in real university classrooms? *J. Cogn. Psychol.* 33, 974–992. doi: 10.1080/20445911.2021.1956939

Lechuga, M. T., Ortega-Tudela, J. M., and Gómez-Ariza, C. J. (2015). Further evidence that concept mapping is not better than repeated retrieval as a tool for learning from texts. *Learn. Instr.* 40, 61–68. doi: 10.1016/j.learninstruc.2015.08.002

Leggett, J. M. I., Burt, J. S., and Carroll, A. (2019). Retrieval practice can improve classroom review despite low practice test performance. *Appl. Cognit. Psychol.* 33, 759–770. doi: 10.1002/acp.3517

McDaniel, M. A. (2023). Combining retrieval practice with elaborative encoding: complementary or redundant? *Educ. Psychol. Rev.* 35:75. doi: 10.1007/s10648-023-09784-8

McDaniel, M. A., and Pressley, M. (1984). Putting the keyword method in context. *J. Educ. Psychol.* 76, 598–609. doi: 10.1037/0022-0663.76.4.598

McInerney, D. M., Cheng, R. W., Mok, M. M. C., and Lam, A. K. H. (2012). Academic self-concept and learning strategies: direction of effect on student academic achievement. *J. Adv. Acad.* 23, 249–269. doi: 10.1177/1932202X12451020

Mechie, I. R., Plaisted-Grant, K., and Cheke, L. G. (2021). How does episodic memory develop in adolescence? *Learn. Mem.* 28, 204–217. doi: 10.1101/lm.053264.120

Minear, M., Coane, J. H., Boland, S. C., Cooney, L. H., and Albat, M. (2018). The benefits of retrieval practice depend on item difficulty and intelligence. *J. Exp. Psychol. Learn.* 44, 1474–1486. doi: 10.1037/xlm0000486

Mintzes, J. J., Canas, A., Coffey, J., Gorman, J., Gurley, L., Hoffman, R., et al. (2011). Comment on retrieval practice produces more learning than elaborative studying with concept mapping. *Science* 334:453. doi: 10.1126/science.1203698

Miyatsu, T., Nguyen, K., and McDaniel, M. A. (2018). Five popular study strategies: their optimal implementation and pitfalls. *Perspect. Psychol. Sci.* 13, 390–407. doi: 10.1177/1745691617710510

Moreira, B. F. T., Pinto, T. S. S., Starling, D. S. V., and Jaeger, A. (2019). Retrieval practice in classroom settings: a review of applied research. *Front. Educ.* 4:5. doi: 10.3389/feduc.2019.00005

Nesbit, J. C., and Adesope, O. O. (2006). Learning with concept and knowledge maps: a meta-analysis. *Rev. Educ. Res.* 76, 413–448. doi: 10.3102/00346543076003413

Novak, J. D., and Cañas, A. J. (2006). The origins of the concept mapping tool and the continuing evolution of the tool. *Inform. Visual.* 5, 175–184. doi: 10.1057/palgrave.ivs.9500126

Novak, J.D., and Gowin, D.B. (1984). *Learning how to learn*. New York: Cambridge University Press

Nunes, L. D., and Karpicke, J. D. (2015). "Retrieval-based learning: research at the interface between cognitive science and education" in *Emerging trends in the social and behavioral sciences*. eds. R. A. Scott and S. M. Kosslyn (Chichester: Wiley), 1–16.

O'Day, G. M., and Karpicke, J. D. (2021). Comparing and combining retrieval practice and concept mapping. *J. Educ. Psychol.* 113, 986–997. doi: 10.1037/edu0000486

Ordaz-Villegas, G., Acle-Tomasini, G., and Reyes-Lagunes, L. I. (2013). Development of an academic self-concept for adolescents (ASCA) scale. *J. Beh. Health Soc. Iss.* 5, 117–130. doi: 10.5460/jbhsi.v5.2.42304

Ortega-Tudela, J. M., Lechuga, M. T., Bermúdez-Sierra, M., and Gómez-Ariza, C. J. (2021). Testing the effectiveness of retrieval-based learning in naturalistic school settings. *SAGE Open* 11:215824402110615. doi: 10.1177/21582440211061569

Ortega-Tudela, J. M., Lechuga, M. T., and Gómez-Ariza, C. J. (2019). A specific benefit of retrieval-based concept mapping to enhance learning from texts. *Instr. Sci.* 47, 239–255. doi: 10.1007/s11251-018-9476-y

Quinn, H., Mintzes, J., and Laws, R. A. (2003). Successive concept mapping: assessing understanding in college science classes. *J. Coll. Sci. Teach.* 33, 12–17.

Racsmány, M., Szőllősi, Á., and Marián, M. (2020). Reversing the testing effect by feedback is a matter of performance criterion at practice. *Mem. Cogn.* 48, 1161–1170. doi: 10.3758/s13421-020-01041-5

Repetto, E., Sutil, I., and Manzano, N. (2004). *Comprender y aprender en el aula. Programa para la integración en el curriculum de las estrategias de comprensión lectora*. Madrid: UNED.

Rivers, M. L. (2021). Metacognition about practice testing: a review of learners' beliefs, monitoring, and control of test-enhanced learning. *Educ. Psychol. Rev.* 33, 823–862. doi: 10.1007/s10648-020-09578-2

Roediger, H. L., and Karpicke, J. D. (2006). Test-enhanced learning: taking memory tests improves long term retention. *Psychol. Sci.* 17, 249–255. doi: 10.1111/j.1467-9280.2006.01693.x

Roelle, J., Endres, T., Abel, R., Obergassel, N., Nückles, M., and Renkl, A. (2023). Happy together? On the relationship between research on retrieval practice and generative learning using the case of follow-up learning tasks. *Educ. Psychol. Rev.* 35:102. doi: 10.1007/s10648-023-09810-9

Roelle, J., Froese, L., Krebs, R., Obergassel, N., and Waldeyer, J. (2022). Sequence matters! Retrieval practice before generative learning is more effective than the reverse order. *Learn. Instr.* 80:101634. doi: 10.1016/j.learninstruc.2022.101634

Roelle, J., and Nückles, M. (2019). Generative learning versus retrieval practice in learning from text: the cohesion and elaboration of the text matters. *J. Educ. Psychol.* 111, 1341–1361. doi: 10.1037/edu0000345

Roessger, K. M., Daley, B. J., and Hafez, D. A. (2018). Effects of teaching concept mapping using practice, feedback, and relational framing. *Learn. Instr.* 54, 11–21. doi: 10.1016/j.learninstruc.2018.01.011

Roessger, K. M., Moy, G. R., Phan, T. V., Campbell, H., and Parker, D. (2021). Multi-level factors affecting immediate and delayed concept recall in concept mapping with adult learners. *Educ. Psychol.* 41, 224–244. doi: 10.1080/01443410.2020.1806211

Schroeder, N. L., Nesbit, J. C., Anguiano, C. J., and Adesope, O. O. (2018). Studying and constructing concept maps: a meta-analysis. *Educ. Psychol. Rev.* 30, 431–455. doi: 10.1007/s10648-017-9403-9

Shing, Y. L., Werkle-Bergner, M., Brehmer, Y., Müller, V., Li, S., and Lindenberger, U. (2010). Episodic memory across the lifespan: the contributions of associative and strategic components. *Neurosci. Biobehav. R.* 34, 1080–1091. doi: 10.1016/j.neubiorev.2009.11.002

Shing, Y. L., Werkle-Bergner, M., Li, S.-C., and Lindenberger, U. (2008). Associative and strategic components of episodic memory: a life-span dissociation. *J. Exp. Psychol. Gen.* 137, 495–513. doi: 10.1037/0096-3445.137.3.495

Tullis, J. G., and Maddox, G. B. (2020). Self-reported use of retrieval practice varies across age and domain. *Metacogn. Learn.* 15, 129–154. doi: 10.1007/s11409-020-09223-x

Yang, C., Luo, L., Vadillo, M. A., Yu, R., and Shanks, D. R. (2021). Testing (quizzing) boosts classroom learning: a systematic and meta-analytic review. *Psychol. Bull.* 147, 399–435. doi: 10.1037/bul0000309