IN PRESS in INSTRUCTIONAL SCIENCE

A specific benefit of retrieval-based concept mapping to enhance learning from texts

Juana M. Ortega-Tudela, M. Teresa Lechuga, and Carlos J. Gómez-Ariza
University of Jaén, Spain

Research has shown that retrieval activities, that is, actively recalling previously studied information, may substantially contribute to learning from complex educational materials, sometimes more so than other more popular techniques such as rereading and elaborative study. In this context, recent studies (Blunt and Karpicke 2014) have reported that two different retrieval formats (free recall by writing down as many ideas as possible and creating a concept map in the absence of texts) are equally effective as learning tools. Given the benefits frequently attributed to concept mapping and the potential practical implications of this finding, we aimed to further examine the relative effectiveness of both retrieval-based activities. In Experiment 1, we conceptually replicated the main finding from Blunt and Karpicke's study to show that the two formats may lead to similar learning outcomes. In Experiment 2, we coupled both retrieval formats but manipulated the order in which the activities were performed. Results revealed that retrieval-based concept mapping before free recall by means of paragraph writing resulted in better learning on a 2-week delayed test than performing the same activities the other way round. These findings contradict the general idea that it is retrieval itself, regardless of the activity it is embedded in, what promotes learning. From a more applied standpoint, our results join others from recent studies to show that combining different retrieval activities when dealing with educational materials might be particularly effective

---

Learning from texts often requires the use of different techniques such as repeated reading, note taking, and creating visual representations (e.g., diagrams and concept maps). These are all well-known approaches and, to a greater or lesser extent, recognized as useful tools to prompt meaningful learning (see, for example, Dunlosky et al. 2013). Less popular among students and teachers, however, is retrieval practice, which Francis Bacon even suggested played a role in learning (Bacon 1620/2000). Thus, only recently has research shown that retrieval practice may substantially contribute to learning from complex educational materials, sometimes more so than other more popular techniques such as rereading (Roediger and Karpicke 2006a), note taking (McDaniel et al. 2009), elaborative study (Karpicke and Smith 2012) and, of particular importance here, concept mapping (Karpicke and Blunt 2011; see also Lechuga et al. 2015).

In a standard retrieval practice experiment, participants start by studying a set of materials. Afterwards, some participants engage in activities that require them to retrieve the previously studied materials, while others partake in control activities that do not involve retrieval (i.e., rereading). Lastly, all

participants take a final learning test which may be scheduled anywhere from a few minutes post-retrieval to several weeks or months later (for a review, see Karpicke 2017). In short, retrieval-based learning requires students to actively recall information that they have previously studied, with this practice usually leading to enhanced performance on subsequent learning tests (Nunes and Karpicke 2015).

Despite the robust effect of retrieval on learning, students rarely think of retrieval-based activities as relevant learning tools (Dunlosky et al. 2013). For example, when asked to rank their exam preparation techniques, the most frequently cited strategy among college students is rereading the material (notes or textbooks), with retrieval-based activities lagging far behind several other strategies (Karpicke et al. 2009). Moreover, a number of studies have revealed that when asked to make predictions about their own performance in an upcoming learning test, students who performed retrieval-based activities tend to predict worse performance than those who reread the material or created concept maps (e.g., Grimaldi and Karpicke 2012; Karpicke and Blunt 2011; Lechuga et al. 2015; Roediger and Karpicke 2006a). It has been suggested that this failure to recognize the benefits of retrieval strategies (as well as others; see Rohrer and Pashler 2010) may stem from the fact that these strategies usually produce more errors during the learning sessions, even though they tend to be more effective at enhancing performance during subsequent learning assessments (Schmidt and Bjork 1992). Hence, a straightforward challenge for researchers working in this field is to help raise students' and teachers' awareness of the learning benefits of retrieval-based strategies.

The benefits of retrieval for learning are usually interpreted from two different viewpoints. On the one hand, repeated retrieval is seen to promote semantic elaboration (Carpenter 2009, 2011). The idea is that retrieval cues (provided or present in the environment) activate related information in long-term memory, which is encoded along with the memory traces of target and cues. Hence, if these elaborated knowledge structures become activated on a criterion test, target memories are more easily retrieved (see Rawson et al. 2015 for recent evidence supporting this elaborative retrieval hypothesis). On the other hand, the episodic context account (Karpicke et al. 2014) suggests that when learners retrieve knowledge from memory, they reinstate the prior context in which the information was learned. If retrieval is successful, the context associated with the retrieved items is updated and may be used as a powerful cue to access previously encoded information (Lehman et al. 2014). Regardless of the cognitive processes underlying their benefits, however, it is now clear that retrieval activities may play a relevant role in educational contexts, and that further research is needed to determine how retrieval-based tasks should be implemented to improve learning (e.g., Bae et al. 2018; Karpicke 2012; Karpicke and Grimaldi 2012; Roediger and Karpicke 2006b).

Along these lines, Karpicke and Blunt (2011) showed in college students that repeated retrieval produced better learning of educational texts than concept mapping on a 1-week delayed test. At first sight, this finding might be surprising given that concept map creation is thought to entail mental activities (e.g., relational processing, making inferences) that underpin meaningful learning (Novak and Cañas 2006; Novak and Gowin 1984). In fact, concept mapping is a popular technique in educational settings (Hay et al. 2008; Quinn et al. 2003; Vanides et al. 2005). However, the main results observed by Karpicke and Blunt (2011) were basically replicated by Lechuga et al. (2015), who also found that retrieval practice led to better overall performance than concept mapping, even though this effect was less

pronounced for students with experience in using concept maps than it was for those who underwent a short training session in concept mapping.

Blunt and Karpicke (2014) recently conducted two experiments comparing the effectiveness of two different retrieval practice conditions: the paragraph format (the usual way of practicing whereby participants recall by writing down as many ideas as they can); and the concept map format (which requires participants to recall by creating a concept map). Although students usually create concept maps with the aid of texts (as it was the case in the studies by Karpicke and Blunt 2011 and Lechuga et al. 2015), an alternative approach to concept mapping is to draw up maps in the absence of materials, which would necessarily require them to retrieve information from memory. Hence, one could think of this latter learning activity as being especially effective, since it would combine the benefits of concept mapping (i.e., relational processing) and retrieval practice (i.e., prior context reinstatement). In their first experiment, Blunt and Karpicke (2014) had students read a brief science text and then retrieve information from the text using either the paragraph or concept map format (both in the absence of texts). Specifically, after a 5-min reading period, students recalled the text for 10 min, reread it for 5 min, and recalled it again for 10 min. Students completed this first stage of the experiment by predicting their performance on a 1-week delayed test. The participants returned to the laboratory one week later to respond to a final short-answer test that included 10 verbatim and 5 inference questions about the studied texts. The results revealed that there was no difference between the two retrieval formats for either actual performance at test or participants' learning judgments. In their second experiment, the authors also failed to observe differences between the two retrieval formats but showed that both concept mapping and writing down ideas in the absence of texts enhanced long-term retention more than completing the same activities with the original materials made available. Importantly, these findings led Blunt and Karpicke (2014) to conclude that concept mapping and paragraph writing were equally effective as retrieval practice activities. From a more theoretical stance, they suggested that the benefit promoted by the organizational/relational processing involved in concept mapping might be redundant with the processing associated with repeated retrieval. Hence, and according to Blunt and Karpicke (2014), the locus of the observed learning effect is the very act of retrieving information.

Blunt and Karpicke's findings may have relevant practical consequences. Should the benefits of retrieval practice guide the selection of learning strategies in educational settings (Nunes and Karpicke 2015), and to the extent to which both retrieval formats (paragraph and concept map) lead to similar performance, it seems reasonable to suggest the utilization of the simplest and easiest retrieval-based activity for learning. Since concept mapping appears to be a complex activity taking a long time to master (Mintzes et al. 2011; Quinn et al. 2003), designing learning activities based on retrieving in paragraph format might be the best choice. However, before drawing strong conclusions about the possible equivalence between the two above-mentioned retrieval formats, more evidence to support such an idea seems necessary. In addition, further research should examine potential differences between the benefits provided by recalling in paragraph format and retrieval-based concept mapping. Previous memory research has shown that task demands may determine the nature of retrieval processes and their potential effects on performance (Rugg and Wilding 2000; Stenberg et al. 2006). Thus, for example, the behavioural effects of orienting retrieval towards conceptual information are different from those that

stem from orienting retrieval towards perceptual information (Stenberg et al. 2006). However, this issue remains largely unexplored in relation to retrieval-based learning. Hence, the experiments reported here aimed to a) compare the two retrieval formats used by Blunt and Karpicke (2014) under slightly different conditions and b) explore the learning effects of coupling both formats (for a related attempt using other retrieval activities, see Bae et al. 2018).

Experiment 1 was essentially geared towards replicating Blunt and Karpicke's first experiment by comparing the effectiveness of doing retrieval by writing down paragraphs and recalling by creating a concept map in the absence of materials, even though we introduced some procedural changes in order to maximize the possible differences between the two retrieval conditions. Specifically, we eliminated the rereading activity between retrieval periods so that learning, if any, was uniquely based on retrieval activities (i.e., Lechuga et al. 2015; Roediger and Karpicke 2006a). This is an important step in efforts to disentangle the direct effects of retrieval (those derived from retrieval itself) from the so-called mediated effects of retrieval, which may be promoted if learners are given the opportunity to reread or receive feedback after performing retrieval activities (see Karpicke 2017). In addition, we increased delay between retrieval practice and learning assessment (2 weeks in the present experiments; 1 week in the study by Blunt and Karpicke 2014), given that several studies have demonstrated that longer periods may optimize the benefits of retrieval practice (Dunlosky et al. 2013; Kornell et al. 2011; Roediger and Karpicke 2006b).

Experiment 2 aimed to put to a strong empirical test the idea that the two retrieval formats are similarly effective. Specifically, we had two groups of participants do retrieval practice in both formats but performing the activities in a different order. Thus, a group first engaged in concept mapping and then moved onto a paragraph-based activity, whereas the other group did the opposite. If both retrieval formats engage comparable cognitive processing with parallel effects on learning (as suggested by Blunt and Karpicke 2014), then one would expect the order of the retrieval activities to be irrelevant and the two groups to show indiscernible performance on the final test (as expected in Experiment 1). On the contrary, if creating a concept map provides a special context for retrieval (for example, by virtue of the relational processing that it is thought to involve; Novak and Cañas 2006), then the effectiveness of recalling in paragraphs, which would entail distinct cognitive processing, could depend on whether this retrieval activity is done before or after retrieval-based concept mapping. In other words, depending on the degree to which both formats recruit differential processing, interplays between them could be expected on the basis of which retrieval format is performed first time round. Although we were theoretically sceptical about the possible effect of combining the two retrieval formats, this was an empirical question to address given the previous findings from Blunt and Karpicke's (2014) study and its potential practical implications.

Experiment 1

Method

Participants

The sample comprised 60 undergraduate students (30 women; M_{age} = 26.76, SD_{age} = 3.67) studying for a degree in Education at the University of Jaén (south-central Spain), who were randomly assigned to one of the two experimental conditions. Participants obtained course credits in exchange for their voluntary engagement in the study. They were all provided with partial information about the study and they gave written informed consent before participating. The university's research ethics committee approved the study.

Materials

Here we used the same text as Lechuga et al. (2015), which was selected from a list of suggested material for concept mapping training (Repetto et al. 2004). The text, which contained 427 words and included 25 ideas or propositions on the use of fibres and the manufacturing of fabrics, had a clear-cut organizational structure and educationally relevant content corresponding to the first year of the Spanish Secondary Education (ESO) system.

The learning test comprised 15 short-answer questions: 5 inference questions that elicited inferences from the studied content and 10 verbatim questions that required specific recall of content directly presented in the text. Even though the text included 25 propositions and the test comprised just 15 short-answer questions, most of the ideas were conveyed in the questions since the inference questions required more than one idea when it came to providing a correct answer.

One example of an inference question is "Which kind of fibres have been the most recently discovered?", with "synthetic" being the correct answer. In order to answer the question, information must be integrated from different sentences in the text ("There are three kinds of fibres that can be used for the production of fabrics: Natural, synthetic and mixed" and "Until 1900 all fibres were natural, but now there are more artificial than natural fibres"). One example of a verbatim question is "What kind of fibres can be used for the production of fabrics?", with the correct answer being "natural, synthetic and mixed", which explicitly appeared in the text.

Procedure

The experiment took place over two independent sessions with a 14-day delay between them. In the first session, participants completed a short form of a questionnaire to assess their use of study techniques (Hábitos y Técnicas de Estudio, CHTE; Álvarez and Fernández 1999). The scores pertaining to this scale range from 1 (never) to 4 (always) according to the frequency of use of different techniques. This was immediately followed by participants receiving the same text for a 5 min reading. This was the only time that participants (from both groups) spent reading the text. From this point onwards, each group performed two retrieval trials that differed according to the assigned group: paragraph format or concept map format.

After reading the text, the paragraph group engaged in retrieval by using a paragraph format. They were given a blank piece of paper and asked to write down everything they could recall about the text for 10 minutes. When this time was up, the piece of paper was collected and there was a 1-min break between each trial. Finally, another blank piece of paper was handed to the students for them to write down everything they could recall again for a further 10 min.

As for the concept-mapping group, they first received a brief summary of the characteristics of concept maps and then the experimenter gave them an example following the procedure outlined by Blunt and Karpicke (2014; see also Blunt and Karpicke 2011 and Lechuga et al. 2015). After this, they read the text for 5 min and then were asked to engage in retrieval practice by creating a concept map (in the absence of the text). They were also given a blank piece of paper for 10 min and asked to write/draw on it. After a 1-min break, they were handed a new blank piece of paper and asked to create a second concept map for 10 min. Following that the participants were scheduled for a second experimental session.

In the second experimental session, two weeks later, the participants took the learning test. Like in related previous studies (i.e., Blunt and Karpicke 2011; Blunt and Karpicke 2014; Lechuga et al. 2015), participants were told to spend the time they needed to answer the questions, which came to approximately 15 min.

Results

To ensure minimal engagement in the learning tasks and to maximize the effect of retrieval [a meta-analysis by Rowland (2014) concluded that retrieval practice effects become more robust when initial retrieval is greater than 75%], participants who recalled less than 6 ideas (25% of the text content) during retrieval practice were no longer considered for analysis. Because our main aim was to better understand the potential effect of retrieval on learning, we also wanted to ensure minimal homogeneity in retrieval performance across both groups during the learning sessions. One participant from the paragraph group and 5 from the concept-mapping group were eliminated from the data set. Hence, the final sample sizes were 29 for the paragraph group (12 women; M_{age} = 21.07, SD_{age} = 3.39) and 25 (15 women; M_{age} = 20.40, SD_{age} = 4.01) for the concept-mapping group. For the sake of completeness, results of analyses including all participants are presented as Supplementary Material. Importantly, the general findings that emerge from both analytical approaches are essentially the same, which allows for a better comparison between our results and those by Blunt and Karpicke (2014; Exp. 1) who did not report having used an exclusion criterion.

	Experiment 1		Experiment 2	
Technique	**P+P**	**CM+CM**	**P+CM**	**CM+P**
Underlining	2.69 (.60)	2.80 (.50)	2.72 (.59)	2.66 (.69)
Schemes	1.79 (1.01)	1.92 (1.00)	2.13 (.94)	1.78 (.89)
Summaries	1.86 (.83)	1.84 (1.02)	2.07 (.77)	1.84 (.84)
Concept Maps	1.21 (.98)	1.08 (1.00)	1.16 (.90)	1.10 (.89)

Note: P+P = Paragraph + Paragraph; CM+CM= Concept Map + Concept Map; P+CM = Paragraph + Concept Map; CM+P = Concept Map + Paragraph.

Table 1. Mean scores (and standard deviations) on the CHTE for the frequency of use (1=never, 2=occasionally, 3=frequently, 4=always) of different study techniques as a function of group in Experiments 1 and 2.

One-way analyses of variance (ANOVA) on CHTE scores (obtained at the very start of the first experimental session) revealed that both groups were similar in frequency of use of different study techniques (*Fs* < 1). As shown in Table 1, underlining was the most utilized technique, whereas concept mapping was hardly used.

Two judges (second-year PhD students), who were blind to the study aim, scored the number of ideas from the text (1 point per unit) that were either recalled (retrieval group) or represented (concept-mapping group) during the retrieval trials. These judges also scored the responses on the final learning test (correct responses were also awarded 1 point). They used a template to score the participants' responses, which made the procedure relatively simple and free of ambiguity. The (Pearson's) correlation between the scores from the two evaluators was always over .91. Potential discrepancies (less than 5% of all cases; Cohen's Kappa = .93) were discussed until 100% consensus was reached.

A 2 (retrieval group) x 2 (retrieval trial) mixed ANOVA was carried out on the proportion of ideas recalled from the text (out of the total of 25) during the first experimental (learning) session (see Table 2). The analysis revealed that there was a main effect of retrieval trial, $F(1,52) = 13.77$, $Mse = .003$, $p < .01$, $G\eta^2 = .03$, with more ideas being recalled in the first trial (.40) than in the second retrieval trial (.37). Neither the main effect of group, $F(1,52) = 2.65$, $Mse = .02$, $p = .11$, $G\eta^2 = .04$, nor the interaction, $F(1,52) < 1$, $Mse = .003$, $p > .50$, $G\eta^2 < .01$, showed to be reliable.

Retrieval	Experiment 1		Experiment 2	
Attempt	**P+P**	**CM+CM**	**P+ CM**	**CM+P**
Trial 1	.39 (.11)	.42 (.09)	.42 (.09)	.45 (.09)
Trial 2	.34 (.12)	.39 (.09)	.38 (.09)	.45 (.08)

Note: P+P = Paragraph + Paragraph; CM+CM= Concept Map + Concept Map; P+CM = Paragraph + Concept Map; CM+P = Concept Map + Paragraph.

Table 2. Proportion of ideas units produced in each retrieval attempt at learning period in Experiments 1 and 2. Standard Deviations are shown in parentheses.

To determine participants' knowledge of the text ideas two weeks after doing retrieval practice, we performed a mixed 2 (retrieval group) x 2 (question type) ANOVA on the proportion of correct responses (out of the total of 15, 10 verbatim + 5 inference, questions) in the learning test (see Figure 1). The analysis failed to show a reliable effect of group that had a very small effect size, thus revealing that both retrieval conditions led to comparable performance, $F(1,52) < 1$, $Mse = .04$, $p = .44$, $G\eta^2 = .01$. The effects of question type, $F(1,52) < 1$, $Mse = .02$, $p > .50$, $G\eta^2 < .01$, and its interaction with group, $F(1,52) < 1$, $Mse = .02$, $p > .50$, $G\eta^2 < .01$, also failed to reach statistical significance.

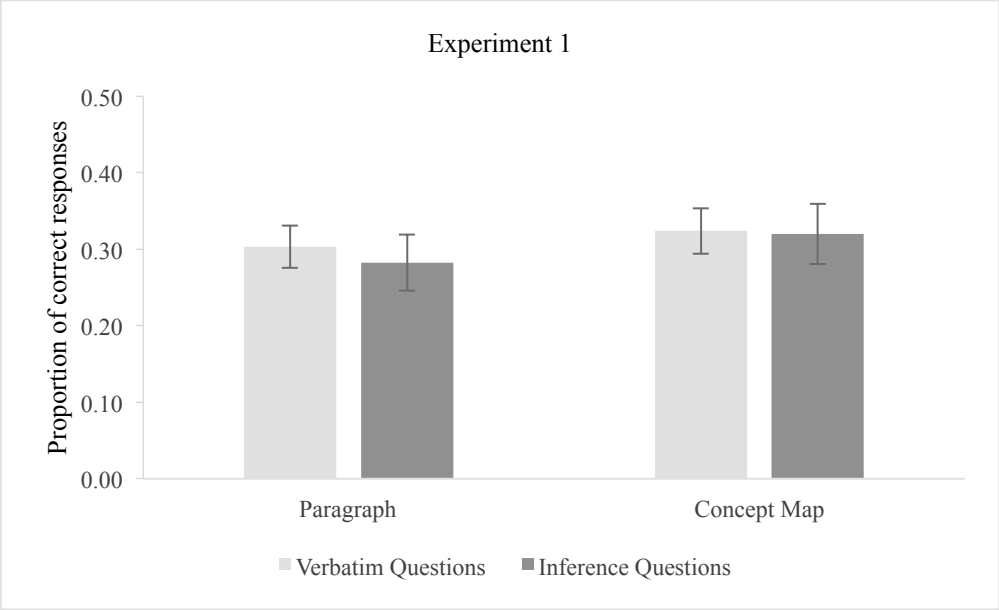

Figure 1. Performance in the final test of Experiment 1 as a function of type of questions. Error bars represent standard errors of the means.

Discussion

Results of the present experiment, like those obtained by Blunt and Karpicke (2014; Exp. 1), show that two different retrieval-based activities, paragraph writing and concept map creation, led to similar levels of performance in a learning test that included literal and inferential questions. On the one hand, we found that the two retrieval formats are similar in learning outcomes even after two weeks (in Blunt and Karpicke's study the participants' memory was tested one week after the learning session). On the other hand, we observed a similar performance pattern with a learning procedure that did not include rereading activities (in Blunt and Karpicke's study, participants reread the text between the two retrieval trials). Thus, our findings coincide with previous literature documenting how the potential benefits of concept mapping overlap those provided by the act of retrieving from memory. Blunt and Karpicke (2014) concluded that the organizational/relational processing that is traditionally associated with concept mapping might be redundant with the cognitive processing that underlies repeated retrieval. While suggestive, such a claim contrasts with those deriving from educational views that attribute valuable benefits to concept mapping and endorse its use to promote meaningful learning (Chevron 2014; Nesbit and Adesope 2006). Though experimental research on the cognitive processes underlying concept mapping is quite sparse, concept map creation is thought to require students to engage in relational processing by focusing on how a set of concepts within a specific domain relate one to another (Karpicke 2018; Vanides et al. 2005), which would facilitate the acquisition of structured knowledge. Hence, and regardless of the role that retrieval plays in promoting learning, one would expect retrieval-based concept mapping to make a difference relative to performing retrieval in a paragraph format.

In a further attempt to test whether these two retrieval formats rely on precisely the same cognitive activities, in Experiment 2 we had two groups of participants do retrieval practice by coupling both formats (for a similar approach with other activities see Bae et al. 2018). Critically, however, we

manipulated the order in which each retrieval format was used. Thus, one group performed the first retrieval by using a paragraph format and then created a concept map in the absence of the text, whereas the other group first engaged in concept mapping and then retrieved in paragraphs. The idea is that if the two retrieval-based activities are cognitively redundant, the order in which participants engage in these activities should be irrelevant and similar learning would result from carrying out the two retrieval activities one after the other, regardless of the order. Hence, the observation of differential performance at the learning test as a function of the order in which the activities were performed would suggest that 1) distinct cognitive operations underlie retrieval in both formats and 2) effectiveness of retrieval for learning may depend on how different activities are coupled.

Experiment 2

Method

Participants

The initial sample comprised 110 undergraduate students (65 women; M_{age} = 21.03, SD_{age} = 4.14) from two natural (first and second year) groups studying for a degree in Education at the University of Jaén. All students voluntarily participated in the study in exchange for course credits and were provided with information about the study; they gave written informed consent before participating. The university's research ethics committee approved the study. Participants were randomly assigned to one of the two experimental conditions: paragraph first and then concept map (P+CM), and concept map first and then paragraph (CM+P).

Materials and procedure

Materials (text and learning test) were the same as those used and described in Experiment 1. The experiment also took place over two independent sessions with a 14-day delay between them. Importantly, the two experimental groups performed the same activities, albeit in reverse order. Specifically, and after being informed and signing the corresponding consent, all participants were given the text for an initial 5-min reading period. Immediately after, they undertook two 10-min retrieval trials with a 1-min break between each trial. However, the type of retrieval activity to be done first depended on the group. Thus, the P+CM group first engaged in retrieving as much as possible from the previously read text by writing paragraphs, and then was told to create a concept map about the content (but in the absence) of the text. In contrast, the CM+P group first engaged in concept mapping (also in the absence of the text) and then was told to retrieve and write down as many ideas as possible from the text. Participants took the learning test two weeks later. This lasted approximately 15 min.

Results

As in Experiment 1, the CHTE questionnaire revealed that both groups were similar in frequency of use of different study techniques (*Fs* < 1; See Table 1). Also, as in the previous experiment, participants who did not retrieve at least 6 ideas (25% contained in the text) at the first retrieval trial were

eliminated (see Supplementary Material for analyses including these participants). This criterion left us with 50 participants (31 women) in the P+CM group and 43 (27 women) in the CM+P group.

To check for differences between the groups, we first performed a mixed ANOVA on the proportion of retrieved ideas during the first experimental session, with order of activities as the between-participants factor (P+CM vs. CM+P) and retrieval trial as the within-participant factor (Trial 1 vs. Trial 2) (see Table 2). The analysis failed to show a main effect of retrieval trial, $F(1,91) = 1.60$, $Mse = .01$; $p = .21$, $G\eta^2 = .01$, thus revealing that, on average, participants came up with a similar proportion of ideas in the first (.43) and second (.42) retrieval activity. There was, however, a main effect of the order of the activities, $F(1,91) = 12.43$, $Mse = .01$; $p < .01$, $G\eta^2 = .08$. The CM+P group recalled more ideas (.45) from the text than the P+CM group (.40). This effect was qualified by an interaction, $F(1,91) = 3.76$, $Mse = .01$; $p = .05$, $G\eta^2 = .01$. Although both groups retrieved a comparable number of ideas at the first trial, $F(1,91) = 2.64$, $Mse = .01$; $p = .11$, $G\eta^2 = .03$, the difference between them was reliable at the second retrieval trial, $F(1,91) = 17.22$, $Mse = .01$; $p < .01$, $G\eta^2 = .16$. Those participants who wrote paragraphs in their second retrieval activity recalled more relevant information than those who had to create a concept map.

We also compared the quality of the concept maps constructed by the participants (see also Lechuga et al. 2015). First, we asked an independent judge, a PhD student who was enrolled in educational research, to score each concept map. The evaluator, who was totally blind to the aims of the study, was told to score according to 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 (see Table 3). As shown in Table 3, the ANOVAs failed to reveal reliable differences between the two groups in any of the quality dimensions.

Assessed dimensions	P + CM	CM + P	*F*	*p*
Primary ideas	7.89 (1.00)	8.15 (1.08)	1.37	.245
Secondary ideas	6.90 (1.24)	7.21 (1.08)	1.63	.205
Hierarchical structure	6.41 (1.69)	6.34 (1.95)	<1	.851
Use of connectors	5.59 (1.99)	5.66 (2.34)	<1	.873
Organization	6.18 (1.53)	6.11 (1.55)	<1	.815
Mean score	6.59 (1.02)	6.69 (1.41)	<1	.707

Note. P+CM = Paragraph + Concept Map; CM+P = Concept Map + Paragraph.

Table 3. Mean scores (and standard deviations in a 0-10 scale) in the five assessed dimensions on the quality of the concept maps in Experiment 2 as a function of group.

Performance on the final learning test (two weeks later) was assessed using a mixed ANOVA with order (P+CM vs. CM+P) as the between-participants factor and question type (verbatim vs. inferential) as the within-participant variable. Like in the previous experiment, the dependent variable was

the proportion of correct responses (to the 10 verbatim and 5 inference questions). Because the scoring procedure showed to be straightforward in Experiment 1 (see also Lechuga et al. 2015), only one of the judges who participated in the previous experiment scored the responses in Experiment 2. The analysis showed a reliable performance difference between the two groups, $F(1,91) = 8.43$, $Mse = .04$; $p < .01$, $G\eta^2 = .05$ (see Figure 2). The group that completed the concept map first (.38) outperformed the group that first retrieved ideas by writing them down as paragraphs (.30). There was no other statistically significant source of variability [type of question: $F(1,91) = 2.06$, $Mse = .03$; $p = .16$, $G\eta^2 = .01$; interaction: $F(1,91) < 1$, $Mse = .03$; $p = .36$, $G\eta^2 < .01$].

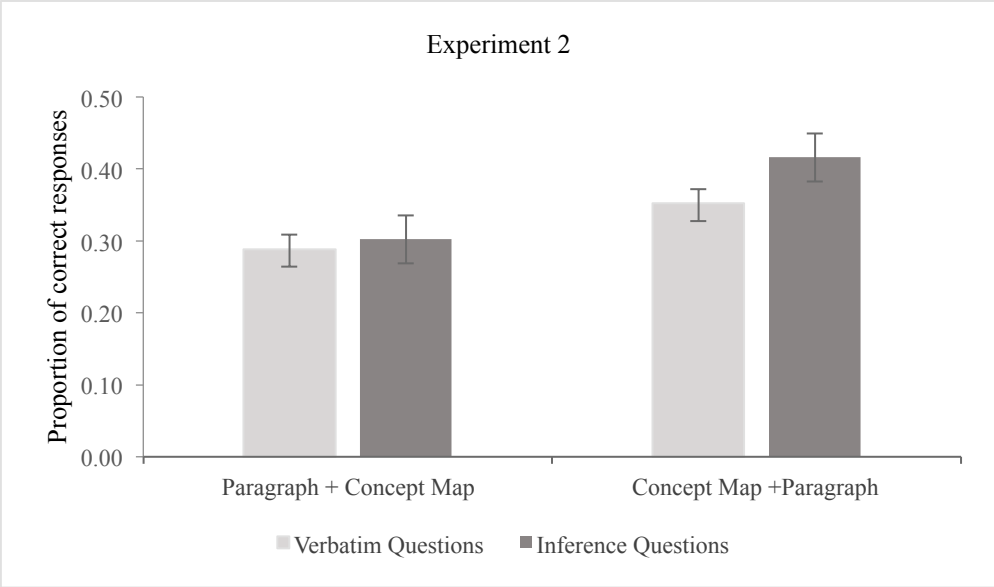

Figure 2. Performance in the final test of Experiment 2 as a function of type of questions. Error bars represent standard errors of the means.

Discussion

Experiment 2 straightforwardly shows that the learning that follows retrieval practice may be modulated by the sequence of retrieval-based activities that are performed. Specifically, elaborating a retrieval-based concept map before performing retrieval from questions by using a paragraph format produced better learning performance than carrying out the same activities in the reverse order. Of relevance, the advantage of the group that created the concept maps first cannot be attributed to the quality of their creations, since the two groups showed to be comparable regarding the quality of the concept maps they constructed. In a similar vein, the number of ideas produced during concept mapping cannot easily account for the benefit of performing this activity in the first place, since both groups produced a similar number of ideas during the first retrieval activity.

However, it was during the second activity that the participants who wrote down paragraphs exhibited better performance, which was also observable on the final learning test. It is worth noting that this benefit of concept mapping as the first learning activity over the next one was not observed in the

previous experiment, since the group that worked on retrieval-based concept maps did not exhibit better performance in the second retrieval activity than the group that did retrieval by using paragraphs. Hence, the present results are clearly indicative of interplay between the type of retrieval activity and its position in a sequence of retrieval attempts.

It seems therefore reasonable to suggest that the two types of retrieval formats used here may (at least partially) involve distinct cognitive operations that, in a row, lead to different learning outcomes as a function of the order they are performed. This contrasts with the results from Experiment 1 (and those from Blunt and Karpicke, 2014), which revealed that performing the same retrieval activity twice (either concept mapping or writing down paragraphs) leads to similar learning performance. Rather, the present result is suggestive of how successful the combination of different retrieval activities may become.

General discussion

The general aim of the present work was to contribute to the field by further exploring the idea that it is retrieval itself, regardless of the activity it is embedded in, what promotes learning. Blunt and Karpicke (2014) provided initial support for this idea by showing that practicing retrieval either by writing down ideas in paragraphs or by creating a concept map led to similar performance on a 1-week delayed learning test. Given the potential educational implications of this finding, our first experiment sought to replicate this by introducing some relevant procedural changes. Indeed, the results of Experiment 1 confirmed that the two retrieval formats may lead to similar outcomes 2 weeks after the learning sessions, even when no rereading opportunities are given to learners. Hence, our conceptual replication of Blunt and Karpicke's (2014) findings would seemingly lend support to the idea that concept mapping and writing down paragraphs are similarly effective as retrieval practice activities, and that the organizational/relational processing involved in creating a concept map might be redundant with the one involved in repeatedly retrieving from memory. However, before strongly claiming the indistinguishable effectiveness of concept mapping and paragraph writing as retrieval activities, we aimed to further test this idea by having students use both retrieval formats one after the other. While we had no a priori hypothesis regarding potential performance differences as a function of the sequential combination of both formats, observing such differences would speak against the idea that memory retrieval, regardless of the activity in which it is embedded, is the main factor to consider when it comes to planning retrieval-based learning activities. Experiment 2 showed that creating a concept map in the absence of materials and then engaging in free recall by writing paragraphs results in better learning than doing these activities the other way round. In fact, constructing a concept map without materials first led to better performance even during the next learning activity (paragraph writing). Hence, this result does not fit well with the idea that both retrieval activities involve redundant cognitive processing (Blunt and Karpicke 2014). If this were the case and the core of both activities was the act of retrieval itself, no learning differences between the two sequences should have been observed.

So what else does the present finding tell us about the two retrieval formats under analysis? A reasonable approach to answering this question might start with the consideration of the possible cognitive processes that underlie concept mapping. While few studies have specifically addressed this issue (i.e., Grimaldi et al. 2015; Karpicke 2018), it is widely assumed that concept mapping entails

semantic/relational processing by virtue of requiring students to hierarchically organize related information (Novak and Cañas 2006). In fact, there is evidence in support of a better knowledge organization after elaborating concept maps (Chevron 2014; Novak and Gowin 1984). Hence, we favour the idea that when the creation of a concept map strongly relies on retrieval (i.e., in the absence of texts), the relational structure of the concepts in the text might gain special prominence and guide access to relevant information in memory. Research has largely demonstrated that semantic associations reliably influence organization during episodic memory search (e.g., Morton and Polyn 2016). Thus, asking learners to construct a concept map in the absence of materials would make a difference as opposed to instructing them to only recall and write down ideas from the text. It is worth noting here that this proposal calls for the consideration of varying retrieval processes that might be induced by the two formats used in the present study. Research into cognitive psychology and neuroscience has shown that task demands determine the specific processing of cues that takes place during retrieval, which has been called retrieval orientation (Herron and Wilding 2004; Rugg and Wilding 2000). Interestingly, some studies have provided evidence that the adoption of a certain retrieval orientation may modulate performance (Bridger et al. 2009; Rosburg et al. 2014). For example, orienting the retrieval towards either perceptual or conceptual information leads to different behavioural effects and brain signatures (Stenberg et al. 2006). Hence, because the two retrieval activities performed by our participants differed in their demands (one of them necessarily required relational processing: "elaborate a concept map with everything you remember from the text", whereas the other did not: "write down everything you remember from the text"), retrieval orientation during such activities might be contributing to our main finding. Thus, we posit that, in addition to the relevant isolated concepts, concept mapping in the absence of texts might primarily direct retrieval[1] towards specific subsets of information (i.e., relations among the retrieved concepts), which could thus increase the accessibility of the central (and related) ideas contained in the text to a greater extent than via other retrieval activities (such as free recall by writing down).

It should be noted, however, that only considering concept mapping would not give a complete picture of its potential benefit as a retrieval-based learning activity. It becomes evident across experiments that it was only the coupling of concept mapping with paragraph format (in this order only) what led to the best performance in the learning test two weeks later. This finding is suggestive of how different retrieval demands may interplay with each other to modulate knowledge acquisition and representation. It is possible that creating a concept map without the text and then freely recalling by writing paragraphs facilitates the adoption of a more effective retrieval strategy (i.e., by highlighting the relations among concepts), which in turn could lead to more elaborated knowledge representations. If we assume that the two practice formats may orient retrieval towards different memory traces (e.g., gist-based ones in the case of concept mapping vs. more linear verbatim ones in the case of recalling by writing paragraphs; for a related interpretation, see Bouwmeester and Verkoeijen 2011), it would seem reasonable to expect differential downstream influences on subsequent learning and performance as a function of their arrangement. Thus, doing the final learning test two weeks later after "P+CM" could draw on somewhat

[1] Note that this contrasts with the generalized idea that concept mapping offers an advantage by providing better organization at encoding (Chevron 2014; Hay et al. 2008; Novak and Gowin 1984; Vanides et al. 2005). Interestingly, our results indicate that elaborating a concept map may also be useful for organizing relevant information at retrieval, which allows us to suggest that concept mapping might enhance learning in a variety of circumstances.

linear mental models about the text ideas, since relational retrieval during concept mapping would have been influenced by the previous rather verbatim retrieval induced by paragraph writing. Performing the learning test after "CM+P", however, might rely more heavily on elaborated mental models that would stem from the relational processing involved in creating the concept maps first, with this processing also influencing subsequent, presumably less elaborative, retrieval activities (i.e., paragraph writing). While speculative, this line of reasoning aligns with the widespread idea that concept mapping facilitates the acquisition of structured knowledge (Karpicke 2018; Vanides et al. 2005) and might explain the results of Experiment 2.

Going back to the recent accounts of retrieval-based learning mentioned in the introduction, at first sight the main finding in Experiment 2 would seem to lend support to the elaborative retrieval account proposed by Carpenter and colleagues (2009, 2011; Carpenter and Yeung 2017). Essentially, it is argued that people generate several pieces of semantically related information during retrieval that are integrated with target information to form richer memory traces, which will become more accessible in future retrieval attempts (i.e., an exam). Hence, at least on the basis of the arguments made above, our finding in Experiment 2 coincides with the idea that certain retrieval activities may be more effective for learning when performed first by virtue of their greater demands on semantic processing. We do, however, recognize that these studies were not specifically designed to test predictions arising from any of the general accounts addressing the benefits of retrieval-based activities, which prevents us from drawing strong conclusions on this issue.

A point that also merits attention is the fact that we do not observe an effect of question type in the final test. Because inference questions are thought to rely more on relational processing and deeper comprehension (Grasser et al. 1994), they should be more sensitive to the conceptual learning that is expected to stem from concept mapping, or any other elaborative techniques. In partial support of this expectation, in Experiment 2 the mean proportion of correct responses to inference questions was numerically higher than the mean proportion of correct responses to verbatim questions in the concept map plus paragraph group (the one with the best general performance at final test). Moreover, a previous study that used the same materials (text and questions in the final test) as we employed in the present work revealed that inferential questions were sensitive to learning after concept mapping (Lechuga et al. 2015), which endorses the utility of such questions to measure relational comprehension here. Taken together, the above-mentioned evidence allows us to suggest that coupling retrieval-based concept mapping and free recall in paragraphs only had a subtle effect on those questions requiring inferential processing. Perhaps such an effect could be maximized in future studies by including a higher number of inference questions in the learning test to better sample the relational contents described in the text (more complex study materials may be required).

The present work is not without limitations. First of all, we recognize that the design of Experiment 1 might not be suitable for detecting small, but still meaningful, effects. Hence, we call for more studies geared towards comparing the effectiveness of free recall and retrieval-based concept mapping. A second limitation is the relatively short interval (2 weeks) between learning activities and learning assessment. While this learning-test delay is longer than that (1 week) used in most previous studies (e.g., Bae et al. 2018; Blunt and Karpicke 2014), it is a truism that to more precisely determine the

effectiveness of learning techniques to be adopted in educational settings requires the consideration of more realistic intervals (i.e., more weeks or even a few months). An additional drawback of the present study is that neither the educational material (a text on fabrics and fibres) nor the learning/assessment sessions were part of the genuine educational setting of the students who participated in the experiments, which necessarily tempers the strength of the conclusions to be drawn from our main findings. As for the educational material used, we assumed that our undergraduate students had little prior knowledge about the subject; we therefore chose not to perform a pretest on this knowledge. While any potential effect of this factor on learning should have been minimized due to the fact that participants were randomly assigned to conditions, a pretest of this kind could have helped to more precisely determine if this was indeed the case. Finally, and like many of the previous studies in the field, our second experiment had an empirical rather than a theoretical focus, which places constraints on how to interpret the findings.

Despite its limitations, we believe the present research makes a relevant contribution to the field and has implications for practice and future research. Thus, the present findings lend partial support to the idea that concept mapping is an effective strategy for learning from educational texts (Chevron 2014; Novak and Gowin 1984). While standard concept mapping (in the presence of texts) has not shown to be better for learning than performing retrieval-based activities (i.e., Karpicke and Blunt 2011; Lechuga et al. 2015), the educational relevance of creating concept maps in the absence of texts as well as of combining it with distinct retrieval activities deserve further attention. Identifying specific couplings of learning and teaching strategies with a high degree of effectiveness, as well as helping teachers and students to think of memory tests as knowledge modifiers and learning (not just assessment) tools, should have practical priority. A recent study (Bae et al. 2018) has shown that coupling free recall (the equivalent to our paragraph condition) with practice quizzing using multiple-choice questions (e.g., Dunlosky et al. 2013) led to better learning than studying by using other combinations of retrieval-based activities. Hence, coupling concept mapping with practice quizzing and free recall activities (via different arrangements) could represent an exciting new direction for forthcoming studies.

Finally, we do think that further research should continue focusing on combining retrieval-based concept mapping and free recall by writing paragraphs, even with educational materials and participants (i.e., high school students) different to those selected in the present experiments. Building on this idea, future research efforts on retrieval-based learning should be carried out under more realistic circumstances. Thus, researchers could, for example, take advantage of the very materials students cover in their courses as well as embed the learning/testing sessions into classroom activities. While this approach would sacrifice internal validity for external validity, it would also help to design better interventions in genuine educational settings by involving teachers in the process. After all, teachers (like students, educational psychologists, and parents) are a target audience of the intended message that emerges from research on retrieval-based learning. On a more theoretical note, more work is needed to clarify the precise mental processes that different types of retrieval activities entail. Only in the past few years have researchers started to systematically pay attention to the neurocognitive mechanisms underpinning retrieval-based learning (see Karpicke 2017). In this sense, our findings point to the need for more study on the nature of the different retrieval processes that might be induced by different retrieval activities, as well as on the knowledge representations they might elicit.

REFERENCES

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

Bacon, F. (1620;2000). Francis Bacon: *The New Organon.* Cambridge University Press.

Bae, C.L., Therriault, D.J., & Redifer, J.L. (2018). Investigating the testing effect: Retrieval as a characteristic of effective study strategies. *Learning and Instruction; https://doi.org/10.1016/j.learninstruc.2017.12.008*

Blunt, J. R., & Karpicke, J. D. (2014). Learning with retrieval-based concept mapping. *Journal of Educational Psychology,* 106, 849-858. doi: 10.1037/a0035934

Bouwmeester, S., & Verkoeijen, P. P. J. L. (2011). Why do some chil- dren benefit more from testing than others? Gist trace processing to explain the testing effect. *Journal of Memory and Language*, 65, 32–41.

Bridger, E. K., Herron, J. E., Elward, R. L., & Wilding, E. L. (2009). Neural correlates of individual differences in strategic retrieval processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 35, 1175–1186.

Carpenter, S. K. (2009). Cue strength as a moderator of the testing effect: the benefits of elaborative retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *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. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 37,*1547-1552.

Carpenter, S.K., & Yeung, K.L. (2017). The role of mediator strength in learning from retrieval. *Journal of Memory and Language,* 92, 128–141. http://dx.doi.org/10.1016/j.jml.2016.06.008.

Chevron, M.P. (2014). A metacognitive tool: Theoretical and operational analysis of skills exercised in structured Concept Maps. *Perspectives in Science,* http://dx.doi.org/10.1016/j.pisc.2014.07.001

Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J., & Willingham, D. T. (2013). Improving students' learning with effective learning techniques promising directions from cognitive and educational psychology. *Psychological Science in the Public Interest,* 14, 4-58. doi:10.1177/1529100612453266

Graesser, A. C., Singer, M. & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review*, 101(3), Jul 1994, 371-395

Grimaldi, P. J., & Karpicke, J. D. (2012). When and why do retrieval attempts enhance subsequent encoding? *Memory & Cognition, 40*, 505-513.

Grimaldi, P.J., Poston, L., & Karpicke, J.D. (2015). How does creating a concept map affect item-specific encoding? *Journal of Experimental Psychology: Learning, Memory, and Cognition,* 42, 1049-1061.

Hay, D., Kinchin, I., & Lygo-Baker, S. (2008). Making learning visible: the role of concept mapping in higher education. *Studies in Higher Education*, *33*(3), 295-311.

Herron, J.E., & Wilding, E.L. (2004). An electrophysiological dissociation of retrieval mode and retrieval orientation. *NeuroImage,* 22, 1554-1562.

Karpicke, J. D. (2012). Retrieval-based learning active retrieval promotes meaningful learning. *Current Directions in Psychological Science, 21*, 157-163. doi: 10.1177/0963721412443552**.**

Karpicke, J. D., Butler, A.C., & Roediger, H. L. (2009). Metacognitive strategies in student learning: Do students practice retrieval when they study on their own? *Memory, 17*, 471-479.

Karpicke, J. D., & 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., & Grimaldi, P. J. (2012). Retrieval-based learning: A perspective for enhancing meaningful learning. *Educational Psychology Review*, 24, 401 – 418. doi: 10.1007/s10648-012-9202-2

Karpicke, J. D., & Smith, M. A. (2012). Separate mnemonic effects of retrieval practice and elaborative encoding. *Journal of Memory and Language*, *67*, 17-29. doi: 10.1016/j.jml.2012.02.004.

Karpicke, J. D., Lehman, M., & Aue, W. R. (2014). Retrieval-based learning: An episodic context account. In B. H. Ross (Ed.), *Psychology of Learning and Motivation, Vol. 61* (pp. 237-284). San Diego, CA: Elsevier Academic Press.

Karpicke, J.D. (2017). Retrieval-based learning: A decade of progress. In J. Wixted (Ed.), *Cognitive Psychology of Memory*. Vol. 2 of Learning and Memory: A comprehensive reference (J.H.vByrne, Series Ed.)

Karpicke, J. D. (2018). *Concept mapping*. In B. Frey (Ed.), The SAGE Encyclopedia of Educational Research, Measurement, and Evaluation. (pp. 351-354). Thousand Oaks, CA: SAGE Publishing, Inc.

Kornell, N., Bjork, R. A., & Garcia, M. A. (2011). Why tests appear to prevent forgetting: A distribution-based bifurcation model. *Journal of Memory and Language*, *65*, 85-97.

Lechuga, M. T., Ortega-Tudela, J. M., & Gómez-Ariza, C. J. (2015). Further evidence that concept mapping is not better than repeated retrieval as a tool for learning from texts. *Learning and Instruction*, *40*, 61-68.

Lehman, M., Smith, M. A., & Karpicke, J. D. (2014). Toward an episodic context account of retrieval-based learning: Dissociating retrieval practice and elaboration. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 40*, 1787-1794.

McDaniel, M.A., Howard, D.C., & Einstein, G.O. (2009). The Read-Recite-Review study strategy: Effective and portable. *Psychological Science, 20*, 516–522.

Mintzes, J.J., Canas, A., Coffey, J., Gorman, J., Gurley, L., Hoffman, R., McGuire, S.Y., Miller, N., Moon, B., Trifon, J., & Wandersee, J. (2011). Comment on "Retrieval Practice produces more learning than elaborative Studying with Concept Mapping". *Science* 334, 453. doi: 10.1126/science.1203698

Morton, N.W., & Polyn, S.M. (2016). A predictive framework for evaluating models of semantic organization in free recall. *Journal of Memory and Language,* 86, 119-140. Doi: 10.1016/j.jml.2015.10.002.

Nesbit, J. C., & Adesope, O. O. (2006). Learning with concept and knowledge maps: A meta-analysis. *Review of Educational Research, 76*, 413–448.

Novak, J. D., & Cañas, A. J. (2006). *The theory underlying concept maps and how to construct them.* Technical Report IHMC Cmap Tools. Florida Institute for Human and Machine Cognition. http://cmap.ihmc.us/Publications/ResearchPapers/TheoryUnderlyingConceptMaps.pdf Retrieved 02/18

Novak, J. D., & Gowin, D. B. (1984). *Learning how to learn*. Cambridge Univ. Press, New York.

Nunes, L. D., & Karpicke, J. D. (2015). Retrieval-based learning: Research at the interface between cognitive science and education. In R. A. Scott & S. M. Kosslyn (Eds.), *Emerging Trends in the Social and Behavioral Sciences* (pp. 1-16). John Wiley & Sons, Inc

Quinn, H., Mintzes, J., & Laws, R. A. (2003). Successive concept mapping: Assessing understanding in college science classes. *Journal of College Science Teaching*, 33, 12–17.

Rawson K. A., Vaughn, K. E., & Carpenter, S. K. (2015). Does the benefit of testing depend on lag, and if so, why? Evaluating the elaborative retrieval hypothesis. *Memory and Cognition*, 43, 619-633. doi: 10.3758/s13421-014-0477-z

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

Roediger, H. L., & Karpicke, J. D. (2006a). Test-enhanced learning: Taking memory tests improves long-term retention. *Psychological Science, 17*, 249–255. doi: 10.1111/j.1467-9280.2006.01693.x

Roediger, H. L., & Karpicke, J. D. (2006b). The power of testing memory: Basic research and implications for educational practice. *Perspectives on Psychological Science*, *1*, 181-210. doi: 10.1111/j.1745-6916.2006.00012.x

Rohrer, D., & Pashler, H. (2010). Recent Research on Human Learning challenges conventional Instructional Strategies. *Educational Researcher,* 39 (5), 406-412. DOI: 10.3102/0013189X10374770.

Rosburg, T., Johansson, M., Sprondel, V., & Mecklinger, A. (2014). Retrieving self-vocalized information: An event-related potential (ERP) study on the effect of retrieval orientation. *Brain and Cognition, 92C*, 123–132.

Rowland, C.A. (2014). The effect of testing versus restudy on retention: a meta-analytic review of the testing effect. *Psychological Bulletin,* 140 (6), 1432–1463. http://dx.doi.org/10.1037/ a0037559.

Rugg, M.D., & Wilding, E.L. (2000). Retrieval processing and episodic memory. *Trends in Cognitive Science.* 4, 108-115.

Schmidt, R. A., & Bjork, R. A. (1992). New conceptualizations of practice: Common principles in three paradigms suggest new concepts for training. *Psychological Science*, *3*, 207–217.

Stenberg, G., Johansson, M. & Rosén, I. (2006). Conceptual and perceptual memory: Retrieval orientations reflected in event-related potentials. *Acta Psychologica,* 122, 174-205.

Vanides, J., Yin, Y., Tomita, M., & Ruiz-Primo, M.A. (2005). Using concept maps in the science classroom. *Science Scope,* 28, 8, 27-31.

SUPPLEMENTARY MATERIAL

Analyses including all participants

Experiment 1

Performance in the learning session

The 2 (retrieval group) x 2 (retrieval trial) mixed ANOVA revealed that there was a main effect of retrieval trial, $F(1,58) = 11.89$, $Mse < .00$, $p < .001$, $G\eta^2 = .02$, with more ideas being produced in the first trial (.38) than in the second retrieval trial (.35). Neither the main effect of group, $F(1,58) < 1$, $Mse = .03$, $p > .50$, $G\eta^2 < .003$, nor the interaction, $F(1,58) = 1.56$, $Mse = .002$, $p = .217$, $G\eta^2 = .006$, showed to be reliable.

Performance in the final test

The 2 (retrieval group) x 2 (question type) ANOVA showed a no reliable effect of group, $F(1,58) < 1$, $Mse = .04$, $p = .48$, $G\eta^2 = .01$. The effects of question type, $F(1,58) < 1$, $Mse = .03$, $p > .50$, $G\eta^2 < .01$, and its interaction with group, $F(1,58) < 1$, $Mse = .03$, $p = .46$, $G\eta^2 < .01$, also failed to reach statistical significance and had tiny effect sizes.

Experiment 2

Performance in the learning session

The ANOVA failed to show a main effect of retrieval trial, $F(1,108) < 1$, $Mse = .007$; $p > .50$, $G\eta^2 < .01$, thus revealing that participants came up with a similar proportion of ideas in the first (.40) and second (.41) retrieval activity. There was, however, a main effect of the order of the activities, $F(1,108) = 5.4$, $Mse = .02$; $p = .02$, $G\eta^2 = .04$. The CM+P group recalled more ideas (.43) from the text than the P+CM group (.38). There was no effect of the interaction, $F(1,108) = 1.52$, $Mse = .01$; $p = .219$, $G\eta^2 < .01$.

Performance in the final test

The ANOVA showed a reliable effect of group, $F(1,108) = 7.25$, $Mse = .04$; $p = .008$, $G\eta^2 = .04$. The group that completed the concept map first (.36) outperformed the group that first retrieved ideas by writing them down as paragraphs (.29). There was no other statistically significant source of variability [type of question: $F(1,108) = 3.23$, Mse = .03, $p = .08$, $G\eta^2 = .01$; interaction: $F(1,108) < 1$, Mse = .03; $p > .50$, $G\eta^2 < .01$].