Retrieval Practice "versus" Elaborative Studying With Concept Mapping – Both Promote New Learning With Related Prose Materials

Verena M. Kriechbaum, Karl-Heinz T. Bäuml

Department of Experimental Psychology, Regensburg University, Germany

To appear in:

Journal of Applied Research in Memory and Cognition

Correspondence Address:

Karl-Heinz T. Bäuml

Department of Experimental Psychology, Regensburg University,

93040 Regensburg, Germany.

Email: karl-heinz.baeuml@ur.de

Phone: +49-941-943-3818

Abstract

Retrieval practice on earlier studied material can enhance recall of later studied, new material, a finding termed the forward testing effect. Here, we examined whether a similar effect can arise in response to elaborative studying with concept mapping. Participants studied three prose passages and, between study of the single passages, practiced the preceding passage by virtue of restudy, retrieval practice, or the creation of a concept map. Concept mapping enhanced recall of the final critical passage to a similar degree than retrieval practice when related passages but not when unrelated passages were employed. Moreover, with related passages, both interim tasks enhanced recall even compared to when no preceding passages were studied, and improved performance on a final inference test. The findings suggest that, with related materials, the learning of the later studied material can be aided by interim learning tasks that promote one`s knowledge on the earlier studied material.

Keywords

retrieval practice
concept mapping
forward testing effect
prose material
prior knowledge

General Audience Summary

When students face the study of lengthy textbook chapters, learning of the chapters` later parts can be particularly hard considering the time-related decline in learning. Because textbook chapters often comprise contents that are highly related to one another, here, we addressed the issue of whether learning of the later parts can be aided by interim learning tasks that can promote the students` knowledge on the chapters` earlier parts. In three experiments, participants studied three prose passages that were either related or unrelated in content, and, immediately after study, took a recall test on the final critical passage. In-between the study of the three passages, participants engaged in restudy practice, retrieval practice, or practice by creation of a concept map about the immediately preceding passage. The results showed that when related passages were studied, retrieval practice and concept mapping were both effective in enhancing recall of the final passage, while when unrelated passages were studied, recall was enhanced in response to retrieval practice only. Strikingly, with related passages, both types of tasks enhanced recall of the final passage even compared to when the final passage was studied only, and both types of tasks improved performance on a final test assessing deeper comprehension of the studied passages. Taken together, the findings suggest that when students face the study of a chapter comprising highly related content, learning of later parts can be aided by interspersing study with retrieval practice or concept mapping, both of which can promote the students` knowledge on the earlier practiced parts of the chapter. When faced with a chapter comprising unrelated content, however, retrieval practice only seems to aid learning of the later parts, but not elaborative studying with concept mapping.

Retrieval Practice "versus" Elaborative Studying With Concept Mapping – Both Promote New Learning With Related Prose Materials

Learning with lengthy textbook chapters often poses challenges to students. In fact, learning tends to decline over longer study sessions (e.g., Lindquist & McLean, 2011; Risko et al., 2012), suggesting that it can be hard for students to learn information presented in the later parts of the chapters. While there is growing research on which learning activities are useful with such materials (see, for example, Dunlosky et al., 2013), it has been well established that, with complex educational materials, prior knowledge is a driving factor in learning (e.g., Boscolo & Mason, 2003; Bransford & Johnson, 1972; Buchin & Mulligan, 2023; McNamara et al., 1996). If so, prior knowledge may also be helpful in aiding learning of a textbook chapter`s later information. Unfortunately, as knowledge in a given text`s domain varies considerably across students, poor learning outcomes can result when adequate prior knowledge lacks (e.g., McNamara et al., 1996).

Knowledge relevant for later parts of chapters may not only derive from a student`s prior knowledge base, however. Indeed, contents discussed in textbook chapters are often highly related and build off one another, such that information discussed in earlier parts of a chapter may aid learning related information occurring in the chapter`s later parts. Thus, whenever a learning activity is effective in promoting one`s knowledge on the earlier parts, it may also be effective in facilitating learning of the subsequently studied, later parts. The primary goal of the present study was to test this hypothesis by employing two learning activities assumed to aid learning for complex educational materials: retrieval practice and elaborative studying with concept mapping.

Prior research has repeatedly demonstrated that retrieval practice, compared to restudy, of previously studied material can yield mnemonic benefits for the practiced material itself (for reviews, see Agarwal et al., 2021; Karpicke, 2017). With more complex materials, like prose

passages or video lectures, this so-called testing effect has not only been shown on measures of direct retention (e.g., free recall; Roediger & Karpicke, 2006), but also on measures of deeper comprehension (e.g., inference questions; Butler, 2010; Hinze et al., 2013), suggesting that retrieval practice can indeed promote one`s knowledge on the practiced material. Retrieval practice, compared to restudy, has also been shown to promote the learning of material encountered after practice of the previously studied material (e.g., Szpunar et al., 2008). The finding is often referred to as the forward testing (FT) effect and has typically been demonstrated on measures of direct retention, but rarely on measures of deeper comprehension (for reviews, see Chan, Meissner, & Davis, 2018; Pastötter & Bäuml, 2014; Yang et al., 2018).

Only few studies have yet addressed the question of whether, with related complex materials, interim retrieval practice of earlier studied materials can promote the learning of later studied materials via facilitated comprehension. In one study, Jing et al. (2016) had subjects study a video lecture on public health, which was divided into several parts containing highly related information. Retrieval practice between study of the single parts was not only found to improve recall of the critical final part relative to a restudy condition, reflecting an FT effect, but also to improve integration of information within and across the different parts of the lecture (see also Wissman & Rawson, 2015), suggesting that the FT effect may have been driven by facilitated comprehension due to knowledge integration across the lecture. Indeed, interim retrieval practice has been proposed to make the practiced material more readily accessible during study of the new material, which may aid the integration of related information from different parts of the studied material (Wahlheim, 2015). In two other studies (Kriechbaum & Bäuml, 2024; Wissman et al., 2011), the FT effect with related prose passages has been found to improve recall of a critical final passage relative to a condition in which the final passage was studied only. The finding of such proactive facilitation may also point to facilitated comprehension induced by interim retrieval practice.

Besides retrieval practice, another – and presumably still more popular – learning activity in educational settings is elaborative studying with concept mapping. Concept mapping refers to the creation of node-link-node diagrams that represent key terms and relations among the terms within a set of materials (Novak & Cañas, 2008; Novak & Gowin, 1984). When applied during the study of complex materials, concept mapping is thought to aid knowledge integration, requiring relational processing of the to-be-studied material. In fact, prior research has shown that concept mapping can enhance learning, compared to, for instance, simply listening to the material (for reviews, see Nesbit & Adesope, 2006; Schroeder et al., 2018). To the best of our knowledge, however, no study has yet examined whether concept mapping can also enhance the learning of later materials studied after concept mapping on the earlier materials has taken place.

The Present Study

Three experiments are reported aimed to examine whether, when faced with the study of prose materials, the learning of later parts of such materials can be enhanced by interim learning activities that promote one`s knowledge on earlier parts of the material. In all three experiments, participants were asked to study three prose passages and, immediately after study of passage 3, took a free-recall test on this critical final passage. Experiments 1 and 2 employed different materials, but within each experiment the passages were highly related in content and built off one another. Experiment 3 employed passages that were unrelated to one another and had no overlapping content. After study of passages 1 and 2, participants in each experiment were either asked to study the preceding passage again (restudy condition), freely recall the preceding passage (retrieval practice condition), or create a concept map about the preceding passage (concept mapping condition). In Experiments 1 and 2, participants in all three conditions took a final inference test on passages 1–3, which is thought to assess knowledge integration (e.g., Graesser et al., 1994). All three experiments additionally included a condition

in which participants were asked to just study passage 3 (passage-3 only condition). Inclusion of such condition permits a test of proactive facilitation and can thus be indicative of induced facilitated comprehension.

Three research questions were addressed. First, for related passages, we examined whether interim retrieval practice enhances learning of the critical final passage, as evidenced by improved free recall (Experiments 1 and 2). On the basis of previous findings, we expected to find an FT effect relative to the restudy condition and proactive facilitation relative to the passage-3 only condition (Kriechbaum & Bäuml, 2024; Wissman et al., 2011). In order to more directly address the role of facilitated comprehension for the FT effect, we further examined whether interim retrieval practice enhances knowledge integration across the passages, as measured by the final inference test (Experiments 1 and 2). Because interim retrieval practice has been shown to foster integration across the to-be-studied material (Jing et al., 2016; Wahlheim, 2015), we expected to find improved performance on the final inference test in response to interim retrieval practice.

Second, for related passages, we examined whether similar effects on learning can arise when, instead of interim retrieval practice, elaborative studying with concept mapping is applied as interim learning activity (Experiments 1 and 2). If concept mapping can aid the integration of knowledge by promoting relational processing of practiced material (Novak & Cañas, 2008; Novak & Gowin, 1984), concept mapping may also be effective in facilitating learning of subsequently studied, related material, very similar to what retrieval practice may achieve by making the practiced material more accessible. Interim concept mapping thus may not only improve performance on the final inference test, but also induce an FT-like effect and proactive facilitation. It was the primary goal of the present study to more directly examine the role of facilitated comprehension for the FT effect by testing whether parallel effects can arise in the retrieval practice and concept mapping conditions.

Third, we examined whether concept mapping still induces an FT-like effect if the study material is changed to unrelated passages. Prior research has shown that the FT effect induced by retrieval practice also arises with unrelated passages, but the effect with this type of material only improved recall to a level that was similar to the level in a final-passage-only condition, which contrasts with the finding of proactive facilitation with related passages (Kriechbaum & Bäuml, 2024). The difference in findings suggests that facilitated comprehension mediated the FT effect with related but not with unrelated passages (see also General Discussion). If the FT-like effect induced by concept mapping is strictly tied to facilitated comprehension, concept mapping may produce an FT-like effect with related but not with unrelated passages, because knowledge effects on learning typically arise for domain-related material only (e.g., Hunt & Rawson, 2011; Witherby & Carpenter, 2022). We addressed the issue by changing study material from related passages in Experiments 1 and 2 to unrelated passages in Experiment 3.

Experiment 1

Method

Ethical Considerations

All reported experiments were carried out in accordance with the provisions of the World Medical Association`s Declaration of Helsinki. Each participant in Experiments 1, 2, and 3 provided informed consent prior to participation.

Participants

Sample size in Experiments 1, 2, and 3 was determined on the basis of a power analysis using the G*Power program (Version 3.1.9.7; Faul et al., 2007). Setting α to .05 and β to .20, a sample of 26 subjects per condition was recommended to detect the meta-analytic effect size of the FT effect with prose materials ($g = 0.80$; Chan, Meissner, & Davis, 2018). Closely following this recommendation, a total of 112 students (mean age = 24.4 years; 85 female, 27 male) of

different German universities were enrolled to take part in Experiment 1, with 28 subjects in each of the four experimental conditions. All participants spoke German as their native language and received monetary reward or course credit for participation. Participants in all three experiments were tested individually in an online video conference with the experimenter, who was present for the entire experimental session and monitored compliance with instructions.

Materials

In Experiment 1, the study material comprised three prose passages that were directly related to one another, covered a common topic, and contained information that was intended to be integrated. The passages were consecutive sections of an expository text about the Earth`s internal structure (e.g., crust, mantle, core) and tectonic processes (e.g., continental drift, subduction, earthquakes), which was developed based on information described in Frisch and Meschede (2021) (for the English version of the text, see osf.io/bckyn). Each passage was presented in German language and approximately 180 words in length. Presentation order of the three passages was the same for all participants. Thus, passage 3 always served as the critical final passage, and passages 1 and 2 always served as the preceding passages.

The final test was composed of 10 short-answer questions referring to the three prose passages` content (for the English version of the questions, see osf.io/bckyn). All of the questions were bridging-inference questions that required subjects to integrate information stated in two or more sentences within or across the single passages to infer the correct answer. One example of an inference question is "*Where does the oceanic crust show the highest ages?*", with "*In the subduction zone*" being the correct answer. In order to answer the question, the following sentences from the text must be understood and linked together: "*Oceanic crust is constantly being formed. This occurs along the mid-ocean ridges […] In the process, the older crust is pushed away from the ridges on either side. Around the Pacific, oceanic crust*

submerges beneath the adjacent continental crust […] Due to the compression, mountain ranges emerge above the subduction zone." The Cronbach`s alpha estimate of internal consistency reached .66 for the 10 inference questions, reflecting an acceptable reliability (cf. Schmitt, 1996).

Design and Procedure

The experiment comprised four conditions which were manipulated between participants: restudy, retrieval practice, concept mapping, and passage-3 only (for an illustration of the experiment`s procedure and conditions, see Figure 1). In all four conditions, participants received a brief instruction about the nature of concept mapping prior to the start of the experiment, adopting the procedure outlined in previous research (Blunt & Karpicke, 2014; Karpicke & Blunt, 2011). Participants were told that concept maps are diagrams where the terms in a set of material are represented as nodes and relations among the terms as lines linking the nodes together. Participants were also shown an example of a concept map depicting properties of water (e.g., liquid, solid, gaseous) and stages of the water cycle (e.g., evaporation, condensation). The example concept map was chosen to represent scientific content of common knowledge without directly relating to information contained in the study text (for the English version of the example concept map, see osf.io/bckyn). After completion of this initial phase of instruction, all participants started to pass through the experiment depending on the experimental condition.

In the restudy, retrieval practice, and concept mapping conditions, participants studied three prose passages in anticipation of a final test. Prior to the study phase, participants were told to expect various activities that might follow the presentation of each of the three passages: restudy of a passage that they had just studied, free recall of a passage that they had just studied, or creating a concept map about a just studied passage. It was pretended that the activities following the passages were determined on a random basis when, in fact, the activities varied

between conditions. In particular, participants passed through the same activities after study of passages 1 and 2 within each experimental condition. Participants were also informed that, regardless of the activities that followed the presentation of the three passages, there would definitely be a final test on all the studied passages.

The three passages were presented one after another for 6 min each at the center of the computer screen. After presentation of each single passage, participants counted backward in steps of threes from a three-digit number for 30 s. The activities that followed the backward counting after passages 1 and 2 lasted a total of 12 min each and depended on the experimental condition. Participants were either asked to (i) restudy the preceding passage during two 5-min cycles of repeated presentation, separated by 2 min filled with simple arithmetic tasks (restudy condition), (ii) recall the preceding passage during two 5-min cycles by typing in all the information they could remember in an empty response box on the screen, separated by 2 min filled with simple arithmetic tasks (retrieval practice condition), or (iii) create a concept map about the preceding passage (concept mapping condition). In the concept mapping task, participants were given 12 min to construct a concept map on a blank sheet of paper including as much information as possible from the text passage. Participants were allowed to refer to the passage and the example concept map from the initial phase, which were both provided on the screen for the entire 12-min period. Once the time was up, participants were asked to turn their sheet of paper over so that the created map could no longer be seen.

After presentation of passage 3 and the backward counting task, participants in all three conditions were given 5 min to type in all the information they could remember from passage 3 in an empty response box on the screen. After completion of the passage-3 recall test and another 3 min filled with Tetris as distractor task, participants passed through the final test comprising 10 short-answer questions on passages 1–3. The questions were presented one at a time on the screen in random order, and participants were asked to type in their answers and

work through the questions in a self-paced manner. Participants were not allowed to go back to a previous question once they proceeded to the next question. All participants were able to complete the final test within a 15-min duration.

In the passage-3 only condition, participants were asked to just study passage 3 for an upcoming recall test, without prior study of passages 1 and 2. Presentation of passage 3, backward counting, and the passage-3 recall test were identical in procedure to the other three conditions. Recall in this condition was used as a baseline to measure whether prior study and practice of passages 1 and 2 improved or impaired passage-3 recall in the other three conditions. There was no final test on passages 1–3 in the passage-3 only condition.

Scoring

To assess performance during practice and on the passage-3 recall test, passages 1–3 were parsed into idea units, and the number of ideas that were either included in the recall protocols or on the concept maps was counted to calculate the participants` individual scores. Participants were given 1 point for correctly producing an entire idea unit, or 0.5 points for correctly producing an idea unit`s critical part. Credit was assigned for verbatim responses and for correct paraphrases of idea units. For example, at passage-1 practice, the response "*With growing depth, the pressure and temperature in the Earth's interior increase*" was scored to be worth 1 point, as the response correctly reproduced the idea unit "*With the increase in depth, the pressure and temperature in the Earth's interior rise*". However, the responses "*With the increase in depth, the pressure in the Earth's interior rises*", and "*The pressure and temperature in the Earth's interior rise*" were both scored to be worth 0.5 points only, because both responses failed to mention certain details of the idea unit but correctly incorporated the idea`s critical information. Passage-3 recall was of greatest interest, and two raters separately scored 56 of the passage-3 protocols. High interrater reliability was obtained, with a Pearson correlation of .99 between their scores. The remaining protocols on passages 1–3 and the

concept maps were scored by one rater. Recall protocols and concept maps were coded using the identical coding scheme.

To assess performance on the final test, the number of questions correctly answered was counted for each participant. All 10 questions were scored by awarding 1 point for completely correct answers and 0.5 points for partially correct answers, resulting in a maximum score of 10 possible points. Raters were provided with a coding scheme that identified the key information a response had to contain to be scored as completely or partially correct. For example, referring to the question "*Where does the oceanic crust show the highest ages?*", the responses "*In the subduction zone*" and "*Around the Pacific*" were both considered completely correct and assigned 1 point, whereas the response "*Furthest from the mid-ocean ridges*" was considered partially correct and assigned 0.5 points only. Two raters were blind to condition and separately scored all 84 final test protocols. High interrater reliability was obtained, with a Pearson correlation of .99 between their scores.

Results

For all experiments, we provided Bayes factors (B_{01}) – which reflect the odds for the null hypothesis over the alternative hypothesis – when the results did not reach conventional level of statistical significance (i.e., $\alpha = .05$) (Masson, 2011; Raftery, 1995). All data are available at osf.io/bckyn.

Practice Performance for Passages 1 and 2

Table 1 shows the percentage of idea units from passages 1 and 2 that were either recalled during the first and second cycle of retrieval practice or represented on the concept maps. Because presentation order of passages was not manipulated in this study, we did not carry out comparisons across the single passages, and practice performance for passages 1 and 2 was thus pooled for statistical analyses. Regarding performance in the retrieval practice

condition, pairwise comparisons between the first and second cycles of retrieval practice revealed that recall remained relatively stable across cycles (54.5% vs. 56.1%), $t(27) = 1.349$, $p = .189$, $d = 0.26$, 95% CI of the difference = [-3.9, 0.8], $B_{01} = 0.462$. Practice performance, however, differed reliably between the retrieval practice and concept mapping conditions, with the concept mapping condition producing more idea units in the presence of the to-be-practiced material than the retrieval practice condition in its absence (66.9% vs. 55.3%), $t(54) = 3.324$, $p = .002$, $d = 0.89$, 95% CI of the difference = [4.6, 18.6].

Passage-3 Recall

Correct Recall. Figure 2A shows the percentage of correctly recalled passage-3 idea units for each of the four experimental conditions. A one-way ANOVA of the three conditions in which all three passages were presented for study (restudy, retrieval practice, concept mapping) showed a significant effect of condition, $F(2, 81) = 14.341$, $MSE = 438.53$, $p < .001$, $\eta_p^2 = .26$. Pairwise comparisons revealed that, relative to the restudy condition, recall was higher in both the retrieval practice and concept mapping conditions (retrieval practice vs. restudy: 50.5% vs. 24.1%, $t(54) = 5.385$, $p < .001$, $d = 1.44$, 95% CI of the difference = [16.6, 36.3]; concept mapping vs. restudy: 49.5% vs. 24.1%, $t(54) = 4.328$, $p < .001$, $d = 1.16$, 95% CI of the difference = [13.7, 37.3]), thus demonstrating an FT effect in response to retrieval practice and an FT-like effect in response to concept mapping. There was no reliable difference in recall between the retrieval practice and concept mapping conditions (50.5% vs. 49.5%), $t(54) = 0.161$, $p = .873$, $d = 0.04$, 95% CI of the difference = [-11.0, 12.9], $B_{01} = 7.384$.

Additional comparisons to the passage-3 only condition revealed that recall in this condition was similar to recall in the restudy condition (23.3% vs. 24.1%), $t(54) = 0.147$, $p_{adj} = .884$, $d = 0.04$, 95% CI of the difference = [-10.5, 9.0], $B_{01} = 7.400$, but lower than in the retrieval practice and concept mapping conditions (passage-3 only vs. retrieval practice: 23.3% vs. 50.5%, $t(54) = 5.502$, $p_{adj} = .003$, $d = 1.47$, 95% CI of the difference = [-37.0, -17.2]; passage-

3 only vs. concept mapping: 23.3% vs. 49.5%, $t(54) = 4.433$, $p_{adj} = .002$, $d = 1.19$, 95% CI of the difference = [-38.0, -14.3]). For all three comparisons to the passage-3 only condition, familywise error rates were controlled by adjusting *p*-values according to the sequential Bonferroni procedure.

Intrusions. Intrusions were defined as idea units from passages 1 and 2 that participants erroneously produced on the passage-3 recall test. Table 2 shows the number of intrusions during passage-3 recall in the restudy, retrieval practice, and concept mapping conditions. A one-way ANOVA of the three conditions showed a significant effect of condition, $F(2, 81) = 6.426$, $MSE = 0.23$, $p = .003$, $\eta_p^2 = .14$. Pairwise comparisons revealed that, relative to the restudy condition, participants produced fewer intrusions in both the retrieval practice and concept mapping conditions (retrieval practice vs. restudy: 0.04 vs. 0.46, $t(54) = 2.953$, $p = .005$, $d = 0.79$, 95% CI of the difference = [-0.7, -0.1]; concept mapping vs. restudy: 0.11 vs. 0.46, $t(54) = 2.338$, $p = .023$, $d = 0.63$, 95% CI of the difference = [-0.7, -0.1]). There was no reliable difference in intrusions between the retrieval practice and concept mapping conditions (0.04 vs. 0.11), $t(54) = 1.029$, $p = .308$, $d = 0.28$, 95% CI of the difference = [-0.2, 0.1], $B_{01} = 4.359$. The results align with prior work showing interim retrieval practice to reduce intrusions (Kriechbaum & Bäuml, 2024; Szpunar et al., 2013), and extend on prior work by showing that interim concept mapping can reduce intrusions to a similar degree.

Final Test Performance

Figure 3A shows the percentage of correctly answered questions on the final test in the restudy, retrieval practice, and concept mapping conditions. A one-way ANOVA of the three conditions showed a significant effect of condition, $F(2, 81) = 14.301$, $MSE = 329.91$, $p < .001$, $\eta_p^2 = .26$. Pairwise comparisons revealed that both the retrieval practice and concept mapping conditions outperformed the restudy condition (retrieval practice vs. restudy: 57.3% vs. 39.8%, $t(54) = 3.377$, $p = .001$, $d = 0.90$, 95% CI of the difference = [7.1, 27.9]; concept mapping vs.

restudy: 65.2% vs. 39.8%, $t(54) = 5.453$, $p < .001$, $d = 1.46$, 95% CI of the difference = [16.0, 34.7]). There was no reliable difference in performance between the retrieval practice and concept mapping conditions (57.3% vs. 65.2%), $t(54) = 1.667$, $p = .101$, $d = 0.45$, 95% CI of the difference = [-17.3, 1.6], $B_{01} = 1.837$.

The goal of Experiment 2 was to replicate the results of Experiment 1 using other related passages as study material.

Experiment 2

Method

Participants

Another 112 students (mean age = 24.0 years; 69 female, 42 male, 1 diverse) were enrolled from different German universities to take part in Experiment 2, with 28 subjects in each of the four experimental conditions. Sample size was identical to Experiment 1. Considering the topic of the study material (i.e., retrieval competition in memory), psychology students were excluded from participation to avoid ceiling effects of prior knowledge. All participants spoke German as their native language and received monetary reward or course credit for participation.

Materials

Like in Experiment 1, three prose passages were employed as study material that were directly related to one another, covered a common topic, and contained information that was intended to be integrated. In Experiment 2, however, the topic of the study material was changed, and the passages were now consecutive sections of an expository text about mechanisms of episodic forgetting (for the English version, see osf.io/bckyn). Collectively, the three passages discussed the problem of retrieval competition between memory entries, and how this competition is assumed to be overcome by two cognitive mechanisms, blocking and

inhibition. The text was developed on the basis of a textbook chapter (Bäuml et al., 2020) and covered educationally relevant content in the field of cognitive psychology. Again, each passage was presented in German language and about 180 words in length. Like in Experiment 1, presentation order was the same for all participants, with passage 3 always serving as the critical final passage and passages 1 and 2 always serving as the preceding passages.

Again, the final test was composed of 10 short-answer questions referring to the three prose passages` content (for the English version, see osf.io/bckyn). Five of the questions were bridging-inference questions, and the other five were application questions. Both types of questions required subjects to integrate information stated in two or more sentences within or across the passages to infer the answer. Application questions additionally required subjects to apply the information to a novel situation. One example of an application question is: "*Two students are studying the names of 50 snakes for an upcoming exam. One of them studies the names in random order. The other divides the snakes into 5 groups according to their origin and not only studies the snakes` names, but also whether they come from North America, South America, Africa, Europe, or Asia. Which of the students is likely to do better, and why?*" (Answer: "*The second, as fewer snakes share the same retrieval cue*") To answer the question, the following sentences from the text must be integrated and correctly applied to the question: "*Retrieval competition occurs when memories share certain features with each other. When a common feature is provided as a retrieval cue, all associated memories are reactivated and compete for conscious recall […] For example, if a group of children share the feature that they all belong to the same school class, this common feature can cause problems for the teacher´s memory. When asked for the names of the children in his class, the larger the number of children in the class, the more difficult it becomes to remember each single name.*" The Cronbach`s alpha estimate of internal consistency reached .62 for the 10 inference questions, again reflecting an acceptable reliability (cf. Schmitt, 1996).

Design, Procedure, and Scoring

Experimental design and procedure were identical to Experiment 1 (see Figure 1). For scoring purposes, passages 1–3 were divided into idea units, and performance during practice and on the passage-3 recall test was assessed following the same criteria that had been used for scoring in Experiment 1. For example, referring to the passage-3 idea unit "*Inhibition can be triggered by retrieval practice of the relevant information, but not by relearning it*", the response "*Retrieval practice, but not relearning can lead to inhibition*" was scored to be worth 1 point, as the response correctly reproduced the entire idea unit. However, the responses "*Retrieval practice can lead to inhibition*", and "*Relearning does not lead to inhibition*" were both scored to be worth 0.5 points only, as both responses correctly incorporated critical parts of the idea unit but failed to mention certain aspects of it. Again, two raters separately scored 56 of the passage-3 protocols, and a Pearson correlation of .99 between their scores was obtained. The remaining protocols on passages 1–3 and the concept maps were scored by one rater. Like in Experiment 1, recall protocols and concept maps were coded using the identical coding scheme.

To assess performance on the final test, all 10 questions were scored by awarding 1 point for completely correct answers and 0.5 points for partially correct answers, again resulting in a maximum score of 10 points. Like in Experiment 1, raters were provided with a coding scheme that identified the key information a response had to contain to be scored as completely or partially correct. For example, regarding the above-mentioned question on the two students studying snake names for an exam, the responses "*The second, as fewer snakes share the same retrieval cue*" and "*The second, as fewer snakes compete for conscious recall*" were both considered completely correct and assigned 1 point. However, the response "*The second*" was considered to be partially correct and assigned 0.5 points only. Again, two raters were blind to

condition and separately scored all 84 final test protocols. High interrater reliability was obtained, with a Pearson correlation of .99 between their scores.

Results

Practice Performance for Passages 1 and 2

Like in Experiment 1, we pooled practice performance for passages 1 and 2 because presentation order of passages was not manipulated. Regarding performance in the retrieval practice condition, pairwise comparisons between the first and second cycles of retrieval practice again revealed that recall remained relatively stable across cycles (42.4% vs. 41.4%), $t(27) = 1.082$, $p = .289$, $d = 0.21$, 95% CI of the difference = [-0.9, 2.8], $B_{01} = 0.339$. Like in Experiment 1, practice performance differed reliably between the retrieval practice and concept mapping conditions, with the concept mapping condition producing more idea units during practice than the retrieval practice condition (51.0% vs. 41.9%), $t(54) = 2.119$, $p = .039$, $d = 0.57$, 95% CI of the difference = [0.5, 17.7] (see Table 1).

Passage-3 Recall

Correct Recall. Figure 2B shows the percentage of correctly recalled passage-3 idea units for each of the four experimental conditions. Like in Experiment 1, a one-way ANOVA of the restudy, retrieval practice, and concept mapping conditions showed a significant effect of condition, $F(2, 81) = 5.209$, $MSE = 333.70$, $p = .007$, $\eta_p^2 = .11$. Pairwise comparisons revealed that, relative to the restudy condition, recall was higher in both the retrieval practice and concept mapping conditions (retrieval practice vs. restudy: 36.2% vs. 22.8%, $t(54) = 3.173$, $p = .002$, $d = 0.85$, 95% CI of the difference = [4.9, 21.8]; concept mapping vs. restudy: 36.7% vs. 22.8%, $t(54) = 2.820$, $p = .007$, $d = 0.75$, 95% CI of the difference = [4.0, 23.8]), again demonstrating an FT effect in response to retrieval practice and an FT-like effect in response to concept mapping. There was no reliable difference in recall between the retrieval practice and

concept mapping conditions (36.2% vs. 36.7%), $t(54) = 0.095$, $p = .925$, $d = 0.03$, 95% CI of the difference = [-11.4, 10.4], $B_{01} = 7.449$.

Additional comparisons to the passage-3 only condition revealed that recall in this condition was again similar to recall in the restudy condition (20.8% vs. 22.8%), $t(54) = 0.560$, $p_{adj} = .578$, $d = 0.15$, 95% CI of the difference = [-9.4, 5.3], $B_{01} = 6.365$, but lower than in the retrieval practice and concept mapping conditions (passage-3 only vs. retrieval practice: 20.8% vs. 36.2%, $t(54) = 3.585$, $p_{adj} = .003$, $d = 0.96$, 95% CI of the difference = [-24.1, -6.8]; passage-3 only vs. concept mapping: 20.8% vs. 36.7%, $t(54) = 3.187$, $p_{adj} = .004$, $d = 0.85$, 95% CI of the difference = [-26.0, -5.9]). Like in Experiment 1, the *p*-values for the three comparisons were adjusted by employing the sequential Bonferroni procedure.

Intrusions. Table 2 shows the number of intrusions during passage-3 recall in the restudy, retrieval practice, and concept mapping conditions. Like in Experiment 1, a one-way ANOVA of the three conditions showed a significant effect of condition, $F(2, 81) = 8.278$, $MSE = 0.40$, $p < .001$, $\eta_p^2 = .17$. Pairwise comparisons revealed that, relative to the restudy condition, participants again produced fewer intrusions in both the retrieval practice and concept mapping conditions (retrieval practice vs. restudy: 0.04 vs. 0.68, $t(54) = 3.278$, $p = .002$, $d = 0.88$, 95% CI of the difference = [-1.0, -0.3]; concept mapping vs. restudy: 0.14 vs. 0.68, $t(54) = 2.623$, $p = .011$, $d = 0.70$, 95% CI of the difference = [-0.9, -0.1]). There was no reliable difference in intrusions between the retrieval practice and concept mapping conditions (0.04 vs. 0.14), $t(54) = 1.406$, $p = .166$, $d = 0.38$, 95% CI of the difference = [-0.3, 0.1], $B_{01} = 2.732$.

Final Test Performance

Figure 3B shows the percentage of correctly answered questions on the final test in the restudy, retrieval practice, and concept mapping conditions. Like in Experiment 1, a one-way ANOVA of the three conditions showed a significant effect of condition, $F(2, 81) = 5.510$, $MSE = 313.47$, $p = .006$, $\eta_p^2 = .12$. Pairwise comparisons revealed that both the retrieval practice and

concept mapping conditions again outperformed the restudy condition (retrieval practice vs. restudy: 52.0% vs. 37.9%, $t(54) = 2.916$, $p = .005$, $d = 0.78$, 95% CI of the difference = [4.4, 23.8]; concept mapping vs. restudy: 50.9% vs. 37.9%, $t(54) = 2.984$, $p = .004$, $d = 0.80$, 95% CI of the difference = [4.3, 21.8]). There was no reliable difference in performance between the retrieval practice and concept mapping conditions (52.0% vs. 50.9%), $t(54) = 0.216$, $p = .830$, $d = 0.06$, 95% CI of the difference = [-8.9, 11.0], $B_{01} = 7.305$.

Experiment 2 replicated the results of Experiment 1 with a new set of related passages. The goal of Experiment 3 was to examine whether the results of Experiments 1 and 2 generalize to unrelated passages.

Experiment 3

Method

Participants

Another 112 students (mean age = 22.5 years; 79 female, 33 male) were enrolled from different German universities to take part in Experiment 3, with 28 subjects in each of the four experimental conditions. Sample size followed Experiments 1 and 2. All participants spoke German as their native language and received monetary reward or course credit for participation.

Materials

Like in Experiments 1 and 2, the study material comprised three prose passages. In Experiment 3, however, the three passages were unrelated to one another and covered entirely different topics: passage 1 was an expository text about former U.S. President Woodrow Wilson, passage 2 was an expository text about the chronic wasting disease, and passage 3 was an expository text about the history and uses of garlic (for the English versions, see osf.io/bckyn). The text material was adapted from prior studies (passage 1: Kliegl et al., 2024;

passage 2: Divis & Benjamin, 2014; passage 3: Fritz & Morris, 2015), translated into German, and shortened to approximately 180 words per passage. Like in Experiments 1 and 2, passage 3 was always presented as the critical final passage and passages 1 and 2 as the preceding passages.

Design and Procedure

Design and procedure were identical to Experiments 1 and 2, with the only exception of the final test. In Experiment 3, the final test on passages 1–3 was not carried out at the end of the experiment, because integration of information across passages should have been rare due to the unrelated nature of the three study texts.

Scoring

Again, passages 1–3 were parsed into idea units, and scoring of performance during practice and on the passage-3 recall test followed the same criteria as in Experiments 1 and 2. For example, regarding the passage-3 idea unit "*The Crusaders are believed to have brought garlic to Europe and Britain*", the response "*The Crusaders are thought to have taken garlic to Europe and Britain*" was scored with 1 point, whereas the responses "*The Crusaders are thought to have taken garlic to Europe*" and "*The Crusaders are thought to have taken garlic to Britain*" were both scored with 0.5 points only. Again, two raters separately scored 56 of the passage-3 protocols, and a Pearson correlation of .99 between their scores was obtained. The remaining protocols on passages 1–3 and the concept maps were scored by one rater. Like in Experiments 1 and 2, recall protocols and concept maps were coded using the identical coding scheme.

Results

Practice Performance for Passages 1 and 2

Like in Experiments 1 and 2, practice performance was pooled for passages 1 and 2 because presentation order of passages was not manipulated. Regarding performance in the retrieval practice condition, pairwise comparisons between the first and second cycles of retrieval practice again revealed that recall remained relatively stable across cycles (51.3% vs. 51.2%), $t(27) = 0.055$, $p = .956$, $d = 0.01$, 95% CI of the difference = [-2.2, 2.4], $B_{01} = 0.187$. Like in Experiments 1 and 2, practice performance differed reliably between the retrieval practice and concept mapping conditions, with the concept mapping condition producing more idea units during practice than the retrieval practice condition (69.6% vs. 51.2%), $t(54) = 4.230$, $p < .001$, $d = 1.13$, 95% CI of the difference = [9.7, 27.1] (see Table 1).

Passage-3 Recall

Correct Recall. Figure 4 shows the percentage of correctly recalled passage-3 idea units for each of the four experimental conditions. A one-way ANOVA of the restudy, retrieval practice, and concept mapping conditions showed a significant effect of condition, $F(2, 81) = 8.696$, $MSE = 324.77$, $p < .001$, $\eta_p^2 = .18$. Critically, pairwise comparisons revealed that while recall in the retrieval practice condition was higher than in the restudy condition (50.2% vs. 33.2%), $t(54) = 3.325$, $p = .002$, $d = 0.89$, 95% CI of the difference = [6.8, 27.3], recall in the concept mapping condition was similar to recall in the restudy condition (32.5% vs. 33.2%), $t(54) = 0.191$, $p = .849$, $d = 0.05$, 95% CI of the difference = [-8.2, 6.7], $B_{01} = 7.343$. Thus, there was an FT effect in the retrieval practice condition, but there was no FT-like effect in the concept mapping condition. The difference in recall between the retrieval practice and concept mapping conditions was significant (50.2% vs. 32.5%), $t(54) = 3.263$, $p = .002$, $d = 0.87$, 95% CI of the difference = [6.8, 28.6].

Additional comparisons to the passage-3 only condition revealed that recall in this condition was similar to recall in the retrieval practice condition (48.4% vs. 50.2%), $t(54) = 0.324$, $p_{adj} = .747$, $d = 0.09$, 95% CI of the difference = [-12.9, 9.3], $B_{01} = 7.087$, but higher than in the restudy and concept mapping conditions (passage-3 only vs. restudy: 48.4% vs. 33.2%, $t(54) = 3.950$, $p_{adj} = .003$, $d = 1.06$, 95% CI of the difference = [7.5, 23.0]; passage-3 only vs. concept mapping: 48.4% vs. 32.5%, $t(54) = 3.738$, $p_{adj} = .002$, $d = 1.00$, 95% CI of the difference = [7.4, 24.5]). For all three comparisons to the passage-3 only condition, *p*-values were again adjusted by using the sequential Bonferroni procedure.

Intrusions. Not a single participant had an intrusion from a preceding passage, irrespective of condition (see Table 2). The results mimic those reported in a recent study, which did not find intrusions to emerge with unrelated passages either (Kriechbaum & Bäuml, 2024).

General Discussion

This study replicates three lines of findings from prior work. First, retrieval practice can produce an FT effect with both related prose passages (Experiments 1 and 2) and unrelated prose passages (Experiment 3) (Kriechbaum & Bäuml, 2024; Wissman et al., 2011). Second, with related but not with unrelated passages, the FT effect creates higher recall levels for the final critical passage than a condition in which only the final passage has been studied (Experiments 1–3), indicating that only for related prose materials, interim retrieval practice induces proactive facilitation (Kriechbaum & Bäuml, 2024; Wissman et al., 2011). Third, interim retrieval practice seems to aid the integration of knowledge from across a study session (Jing et al., 2016), as is reflected by the finding that retrieval practice between study of related passages improved performance on a final inference test (Experiments 1 and 2). Going beyond the prior work, this study demonstrates parallel effects on learning for related passages when, instead of interim retrieval practice, interim elaborative studying with concept mapping is applied.

Concept mapping between study of related passages did not only improve performance on the final inference test, but also induced an FT-like effect and proactive facilitation (Experiments 1 and 2). In contrast, no FT-like effect was observed in response to interim concept mapping for unrelated passages (Experiment 3).

The findings are consistent with the idea that, when faced with the study of related prose materials, the learning of later parts of such materials can be enhanced by interspersing study with learning activities that promote one`s knowledge on earlier parts of the material. Prior work provided evidence that retrieval practice and elaborative studying with concept mapping can both promote knowledge on practiced material (see Karpicke, 2017; Schroeder et al., 2018). Because background knowledge is assumed to be a precondition to deal with comprehension gaps when reading complex texts (e.g., Kintsch, 1994, 1998), practice-induced knowledge on earlier parts of such materials may enable subjects to fill in the information necessary for comprehending later parts of the materials, which may facilitate learning. In line with this suggestion, the present results for related passages revealed that active practice on the preceding passages – be it by virtue of retrieval practice or concept mapping – did not only promote performance on the final inference test, which is thought to measure knowledge integration (e.g., Graesser et al., 1994), but did also enhance recall of the final passage relative to the passage-3 only condition, indicating that practice-induced knowledge on the preceding passages indeed facilitated learning of the final passage. The findings thus align with research showing that prior knowledge is a driving factor in learning (e.g., Boscolo & Mason, 2003; Bransford & Johnson, 1972; Buchin & Mulligan, 2023; McNamara et al., 1996).

While, with related passages, FT(-like) effects were present in response to both retrieval practice and elaborative studying with concept mapping, retrieval practice only was found to induce such an effect with unrelated passages, suggesting that the FT effect with the two types of materials may have been caused by different cognitive mechanisms. Indeed, in contrast to

related passages, prior study and practice of unrelated passages should hardly have afforded facilitated comprehension of the final passage, considering that knowledge effects on learning typically arise for domain-related material only (e.g., Hunt & Rawson, 2011; Witherby & Carpenter, 2022). Consistently, concept mapping did no longer induce an FT-like effect with unrelated passages, nor did retrieval practice induce proactive facilitation. Kriechbaum and Bäuml (2024) recently provided evidence that the FT effect with unrelated passages may be caused by an enhancement in "list" discrimination, showing that, for this type of material, the effect generalized to retrieval of (semantic) information unrelated to the study material. Such semantic generation is assumed to enhance list discrimination similar to retrieval practice (Divis & Benjamin, 2014; Jang & Huber, 2008; Pastötter et al., 2011), and to induce release from proactive interference. Indeed, the effect was found to enhance recall to a level similar but not superior to the level in a final-passage-only condition, a finding that we replicated in present Experiment 3. [1]

The present results do not easily align with metacognitive accounts of the FT effect, which assume that retrieval practice between study allows subjects to optimize their strategies for subsequent learning (see Chan, Meissner, & Davis, 2018). As a first reason, it is not clear why interim concept mapping should be equally effective in enhancing subsequent learning for

[1] Whereas, with unrelated prose passages, release from proactive interference may be due to enhanced list discrimination, with related passages, proactive facilitation may result from promoting integration of studied information (Wahlheim, 2015). If so, the present results for related passages may indicate that while proactive interference turned to proactive facilitation in the retrieval practice and concept mapping conditions, in the restudy condition knowledge integration was only partially effective, resulting in a neutral effect on recall of the final passage relative to the passage-3 only condition.

related passages, given that retrieval practice should be clearly more effective in alerting subjects to retrieval failures and the requirements of the critical (free-recall) test, both of which have been proposed to drive strategy changes (e.g., Finley & Benjamin, 2012; Soderstrom & Bjork, 2014). And as a second reason, if one assumes that concept mapping can also induce strategy change by alerting subjects to their (poor) prior learning, one may expect interim concept mapping to also enhance subsequent learning for unrelated passages, which was not the case. Metacognition can critically influence the use of knowledge during learning, which reflects an active and effortful process that learners sometimes skip (e.g., Hinze et al., 2013; McNamara et al., 1996; Nguyen & McDaniel, 2016). Asking subjects to monitor their current state of understanding or asking them for self-explanation have both been shown to effectively guide subjects toward deeper-level processing of a to-be-studied text. Metacognition thus may also play a modulating role for whether FT(-like) effects caused by facilitated comprehension come to full fruition, a prediction that may be addressed in future research.

Although there is an extensive research base on concept mapping (see Nesbit & Adesope, 2006; Schroeder et al., 2018), few studies have directly compared the effects of concept mapping and retrieval practice on learning, and, to the best of our knowledge, those studies were focused on learning of the practiced material itself. In most cases, elaborative studying with concept mapping was found to be less effective than retrieval practice in promoting long-term retention (e.g., Karpicke & Blunt, 2011; Lechuga et al., 2015). Here, we provide first evidence that, with related prose materials, concept mapping can promote the learning of subsequently studied, new material to a similar degree to how retrieval practice does, as is reflected in the observed equivalent inferential question performance. Given the widespread popularity of concept mapping in educational settings (e.g., Hay et al., 2008; Quinn et al., 2003), these findings may have important practical implications, suggesting that one

effective way to implement elaborative studying with concept mapping is to apply it in-between the study of related prose materials.

The present study focused on immediate learning outcomes, leaving it open whether FT(-like) effects as induced by retrieval practice and concept mapping change in size over longer retention intervals. While the classic testing effect has been shown to be more pronounced when the criterion test is delayed (e.g., Kornell et al., 2011; Roediger & Karpicke, 2006), the influence of retention interval on the FT effect is largely unclear to date, in particular for more complex study materials (for preliminary results, see Wissman & Rawson, 2015; for results on word lists, see Chan, Manley et al., 2018; Kliegl & Bäuml, 2021). However, because differences in encoding typically result in long-lasting differences in memory performance (e.g., Wixted, 2022; Slamecka & McElree, 1983), one may speculate that FT(-like) effects caused by facilitated comprehension persist with delay. This holds while the exact pattern of results may critically depend on the cognitive processes underlying such facilitated comprehension, and the cognitive processes mediating the effects of retrieval practice (e.g., increased accessibility; Wahlheim, 2015) and concept mapping (e.g., relational processing; Novak & Cañas, 2008) be differently effective in promoting knowledge integration with different experimental settings. Future research may address this important issue.

Conclusions

The present findings provide clues for students on how to learn with lengthy textbook chapters. When faced with a chapter comprising highly related material, the learning of later parts can be aided by interspersing study of the chapter with retrieval practice or concept mapping, both of which can promote the students` knowledge on the earlier parts of the chapter and thus facilitate later learning. In contrast, when faced with a chapter comprising unrelated material, the learning of later parts can be aided by interim retrieval practice but not by interim elaborative studying with concept mapping, suggesting that practice-induced knowledge on

earlier unrelated parts may not be critical in driving learning of a chapter`s later material. Educators are encouraged to employ both types of practice during courses dealing with related contents, whereas only practice by virtue of retrieval is recommended during courses dealing with unrelated contents.

References

Agarwal, P. K., Nunes, L. D., & Blunt, J. R. (2021). Retrieval practice consistently benefits student learning: A systematic review of applied research in schools and classrooms. *Educational Psychology Review, 33*(4), 1409–1453.

Bäuml, K.-H. T., Abel, M., & Kliegl, O. (2020). Inhibitory processes in episodic memory. In M. W. Eysenck & D. Groome (Eds.), *Forgetting: Explaining Memory Failure* (pp. 125–146). Sage Publishing Ltd.

Blunt, J. R., & Karpicke, J. D. (2014). Learning with retrieval-based concept mapping. *Journal of Educational Psychology, 106*(3), 849–858.

Boscolo, P., & Mason, L. (2003). Topic knowledge, text coherence, and interest: How they interact in learning from instructional texts. *Journal of Experimental Education, 71*(2), 126–148.

Bransford, J. D., & Johnson, M. K. (1972). Contextual prerequisites for understanding: Some investigations of comprehension and recall. *Journal of Verbal Learning & Verbal Behavior, 11*(6), 717–726.

Buchin, Z. L., & Mulligan, N. W. (2023). Retrieval-based learning and prior knowledge. *Journal of Educational Psychology, 115*(1), 22–35.

Butler, A. C. (2010). Repeated testing produces superior transfer of learning relative to repeated studying. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36*(5), 1118–1133.

Chan, J. C. K., Manley, K. D., Davis, S. D., & Szpunar, K. K. (2018). Testing potentiates new learning across a retention interval and a lag: a strategy change perspective. *Journal of Memory and Language, 102*, 83–96.

Chan, J. C. K., Meissner, C. A., & Davis, S. D. (2018). Retrieval potentiates new learning: A theoretical and meta-analytic review. *Psychological Bulletin, 144*(11), 1111–1146.

Divis, K. M., & Benjamin, A. S. (2014). Retrieval speeds context fluctuation: Why semantic generation enhances later learning, but hinders prior learning. *Memory & Cognition, 42*(7), 1049–1062.

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*(1)*,* 4–58.

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*(2), 175–191.

Finley, J. R., & Benjamin, A. S. (2012). Adaptive and qualitative changes in encoding strategy with experience: Evidence from the test-expectancy paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 38*(3), 632–652.

Frisch, W., & Meschede, M. (2021). *Plattentektonik: Kontinentverschiebung und Gebirgsbildung* [Plate tectonics: continental drift and mountain formation] (6th ed.). Wissenschaftliche Buchgesellschaft.

Fritz, C. O., & Morris, P. E. (2015). Partset cuing of texts, scenes, and matrices. *British Journal of Psychology, 106*(1), 1–21.

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

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.

Hinze, S. R., Wiley, J., & Pellegrino, J. W. (2013). The importance of constructive comprehension processes in learning from tests. *Journal of Memory and Language, 69*(2), 151–164.

Hunt, R. R., & Rawson, K. A. (2011). Knowledge affords distinctive processing in memory. *Journal of Memory and Language, 65*(4), 390–405.

Jang, Y., & Huber, D. E. (2008). Context retrieval and context change in free recall: recalling from long-term memory drives list isolation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*(1), 112–127.

Jing, H. G., Szpunar, K. K., & Schacter, D. L. (2016). Interpolated testing influences focused attention and improves integration of information during a video-recorded lecture. *Journal of Experimental Psychology: Applied, 22*(3), 305–318.

Karpicke, J. D. (2017). Retrieval-based learning: A decade of progress. In J. H. Byrne (Ed.), *Cognitive psychology of memory, Vol. 2 of learning and memory: A comprehensive reference* (pp. 1–26). Elsevier.

Karpicke, J. D., & Blunt, J. R. (2011). Retrieval practice produces more learning than elaborative studying with concept mapping. *Science, 331*(6018), 772–775.

Kintsch, W. (1994). Text comprehension, memory, and learning. *American Psychologist, 49*(4), 294–303.

Kintsch, W. (1998). *Comprehension: A paradigm for cognition.* Cambridge University Press.

Kliegl, O., Bartl, J., & Bäuml, K.-H. T. (2024). The pretesting effect comes to full fruition after prolonged retention interval. *Journal of Applied Research in Memory and Cognition, 13*(1), 63–70.

Kliegl, O., & Bäuml, K.-H. T. (2021). When retrieval practice promotes new learning – The

critical role of study material. *Journal of Memory and Language, 120,* Article 104253.

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*(2), 85–97.

Kriechbaum, V. M., & Bäuml, K.-H. T. (2024). Retrieval practice can promote new learning with both related and unrelated prose materials. *Journal of Applied Research in Memory and Cognition, 13*(3), 319–328.

Kriechbaum, V. M., & Bäuml, K.-H. T. (2025, February 28). Retrieval Practice "versus" Elaborative Studying With Concept Mapping – Both Promote New Learning With Related Prose Materials. Retrieved from osf.io/bckyn

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.

Lindquist, S. I., & McLean, J. P. (2011). Daydreaming and its correlates in an educational environment. *Learning and Individual Differences, 21*(2), 158–167.

Masson, M. E. (2011). A tutorial on a practical Bayesian alternative to null-hypothesis significance testing. *Behavior Research Methods, 43*(3), 679–690.

McNamara, D. S., Kintsch, E., Songer, N. B., & Kintsch, W. (1996). Are good texts always better? Interactions of text coherence, background knowledge, and levels of understanding in learning from text. *Cognition and Instruction*, *14*(1), 1–43.

Nesbit, J. C., & Adesope, O. O. (2006). Learning with concept and knowledge maps: A meta-analysis. *Review of Educational Research, 76*(3), 413–448.

Nguyen, K., & McDaniel, M. A. (2016). The JOIs of text comprehension: Supplementing retrieval practice to enhance inference performance. *Journal of Experimental*

Psychology: Applied, 22(1), 59–71.

Novak, J. D., & Cañas, A. J. (2008). *The theory underlying concept maps and how to construct and use them.* Technical Report IHMC CmapTools 2006–01 Rev 01–2008, Florida Institute for Human and Machine Cognition. Retrieved on July 21, 2024 from http://cmap.ihmc.us/Publications/ResearchPapers/TheoryUnderlyingConceptMaps.pdf

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

Pastötter, B., & Bäuml, K.-H. T. (2014). Retrieval practice enhances new learning: the forward effect of testing. *Frontiers in Psychology, 5,* 286.

Pastötter, B., Schicker, S., Niedernhuber, J., & Bäuml, K.-H. T. (2011). Retrieval during learning facilitates subsequent memory encoding. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37*(2)*,* 287–297.

Quinn, H. J., Mintzes, J. J., & Laws, R. A. (2003). Successive concept mapping: Assessing understanding in college science classes. *Journal of College Science Teaching, 33*(3)*,* 12–16.

Raftery, A. E. (1995). Bayesian model selection in social research. *Sociological Methodology, 25*, 111–163.

Risko, E. F., Anderson, N., Sarwal, A., Engelhardt, M., & Kingstone, A. (2012). Everyday attention: variation in mind wandering and memory in a lecture. *Applied Cognitive Psychology, 26,* 234–242.

Roediger, H. L. III, & Karpicke, J. D. (2006). Test-enhanced learning: Taking memory tests improves long-term retention. *Psychological Science, 17*(3), 249–255.

Schmitt, N. (1996). Uses and abuses of coefficient alpha. *Psychological Assessment, 8*(4)*,* 350–

353.

Schroeder, N. L., Nesbit, J. C., Anguiano, C. J., & Adesope, O. O. (2018). Studying and constructing concept maps: A meta-analysis. *Educational Psychology Review, 30*(2), 431–455.

Slamecka, N. J., & McElree, B. (1983). Normal forgetting of verbal lists as a function of their degree of learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 9*(3), 384–397.

Soderstrom, N. C., & Bjork, R. A. (2014). Testing facilitates the regulation of subsequent study time. *Journal of Memory and Language, 73,* 99–115.

Szpunar, K. K., Khan, N. Y., & Schacter, D. L. (2013). Interpolated memory tests reduce mind wandering and improve learning of online lectures. *Proceedings of the National Academy of Sciences of the United States of America, 110*(16), 6313–6317.

Szpunar, K. K., McDermott, K. B., & Roediger, H. L. (2008). Testing during study insulates against the buildup of proactive interference. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*(6), 1392–1399.

Wahlheim, C. N. (2015). Testing can counteract proactive interference by integrating competing information. *Memory & Cognition, 43*(1), 27–38.

Wissman, K. T., & Rawson, K. A. (2015). Grain size of recall practice for lengthy text material: Fragile and mysterious effects on memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 41*(2), 439–455.

Wissman, K. T., Rawson, K. A., & Pyc, M. A. (2011). The interim test effect: testing prior material can facilitate the learning of new material. *Psychonomic Bulletin & Review, 18*(6), 1140–1147.

Witherby, A. E., & Carpenter, S. K. (2022). The rich-get-richer effect: Prior knowledge predicts new learning of domain-relevant information. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 48*(4), 483–498.

Wixted, J. T. (2022). Absolute versus relative forgetting. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 48*(12), 1775–1786.

Yang, C., Potts, R., & Shanks, D. R. (2018). Enhancing learning and retrieval of new information: A review of the forward testing effect. *npj Science of Learning, 3*, 8.

Figure Captions

Figure 1. Procedure and conditions employed in Experiments 1 and 2. The experimental session started with an introduction in concept mapping, irrespective of condition. Afterwards, in three conditions, three prose passages were studied in succession. Each passage was followed by a short distractor task. After study of passages 1 and 2, participants either engaged in two restudy cycles on the preceding passage (restudy condition), in two retrieval cycles on the preceding passage (retrieval practice condition), or constructed a concept map about the preceding passage (concept mapping condition). Passage 3 was recalled immediately after study. Participants in all three conditions then took a final test on the three passages. In an additional passage-3 only condition, passage 3 was studied and recalled without prior study of passages 1 and 2.

Figure 2. Percentage of correctly recalled passage-3 idea units as a function of condition (restudy, retrieval practice, concept mapping) in Experiment 1 (A) and Experiment 2 (B). In both graphs, the dashed line represents mean recall in the passage-3 only condition. Error bars represent standard errors.

Figure 3. Percentage of correctly answered inference questions on the final test as a function of condition (restudy, retrieval practice, concept mapping) in Experiment 1 (A) and Experiment 2 (B). Error bars represent standard errors.

Figure 4. Percentage of correctly recalled passage-3 idea units as a function of condition (restudy, retrieval practice, concept mapping) in Experiment 3. The dashed line represents mean recall in the passage-3 only condition. Error bars represent standard errors.

Figure 1

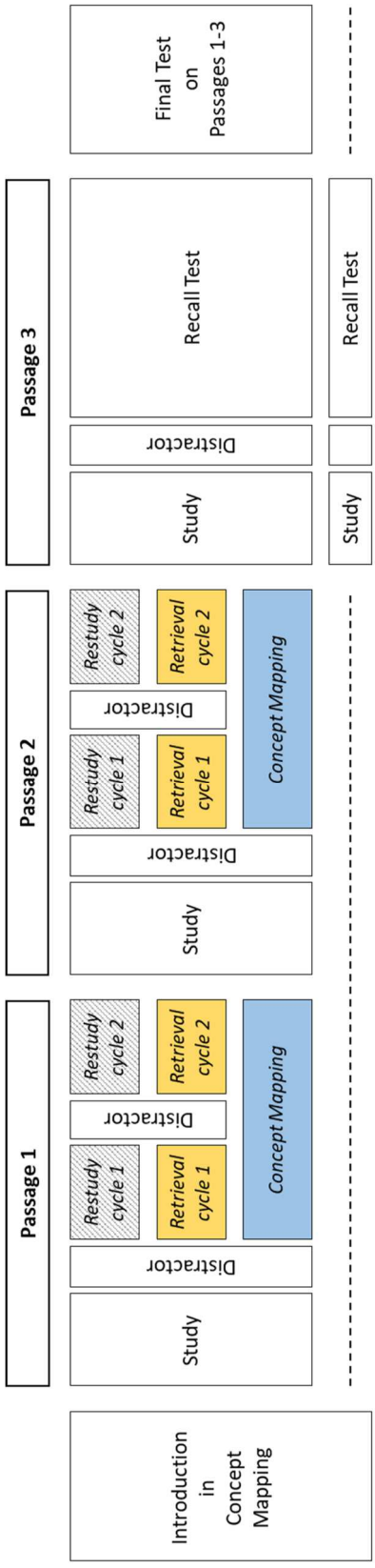
Introduction in Concept Mapping
Passage 1
Study
Distractor
Restudy cycle 1
Retrieval cycle 1
Restudy cycle 2
Retrieval cycle 2
Concept Mapping
Passage 2
Study
Distractor
Restudy cycle 1
Retrieval cycle 1
Distractor
Restudy cycle 2
Retrieval cycle 2
Concept Mapping
Passage 3
Study
Distractor
Recall Test
Study
Recall Test
Final Test on Passages 1-3

Figure 2

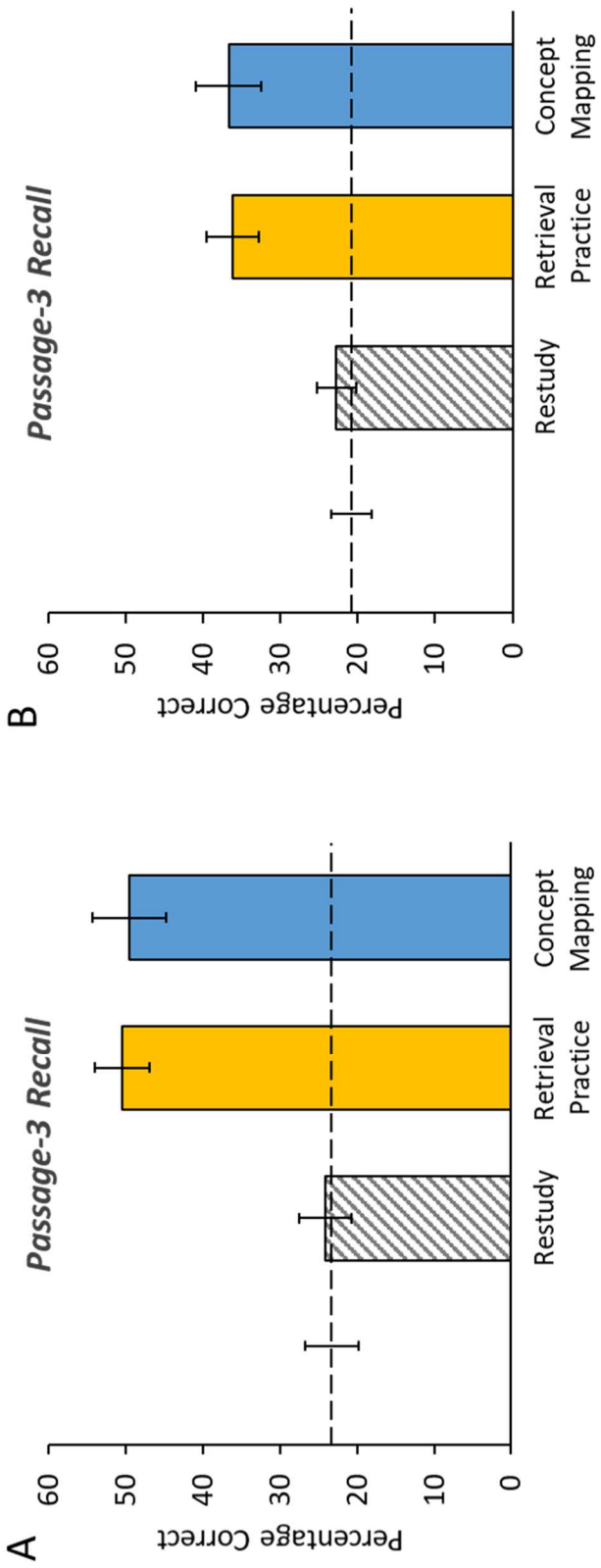
A
Passage-3 Recall
Percentage Correct
60
50
40
30
20
10
0
Restudy
Retrieval Practice
Concept Mapping
B
Passage-3 Recall
Percentage Correct
60
50
40
30
20
10
0
Restudy
Retrieval Practice
Concept Mapping

Figure 3

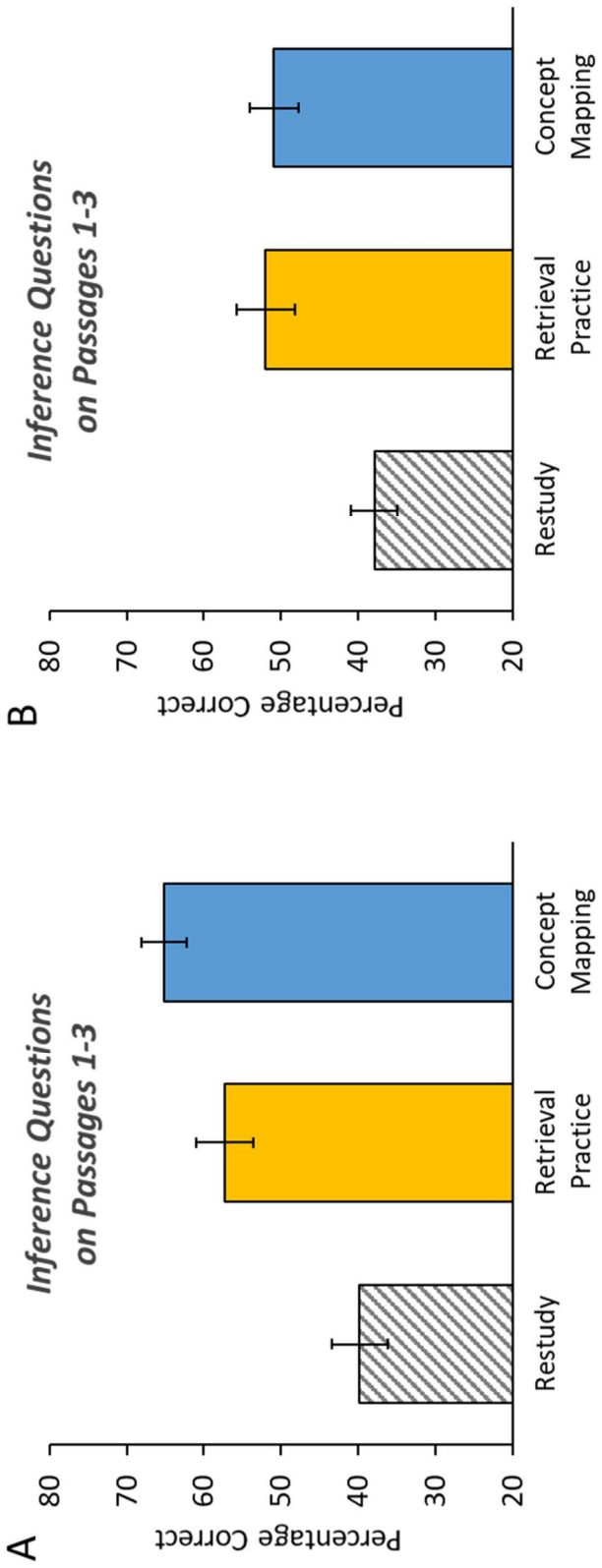
A
Inference Questions on Passages 1-3
Percentage Correct
80
70
60
50
40
30
20
Restudy
Retrieval Practice
Concept Mapping
B
Inference Questions on Passages 1-3
Percentage Correct
80
70
60
50
40
30
20
Restudy
Retrieval Practice
Concept Mapping

Figure 4

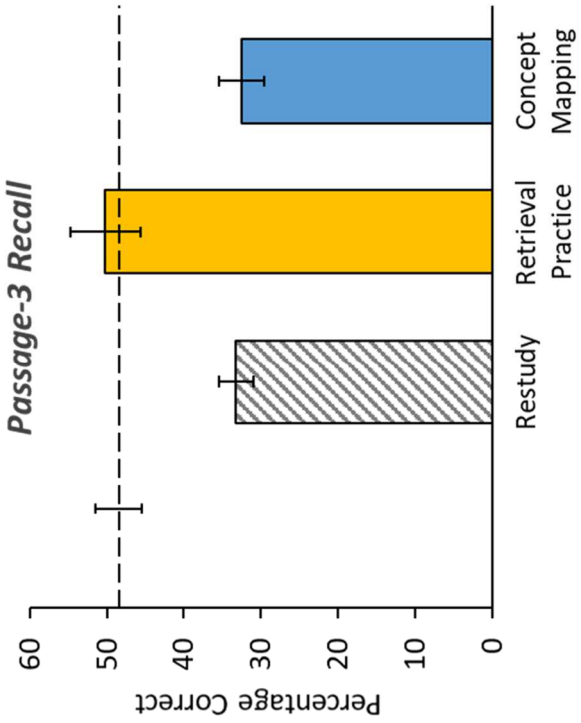
Passage-3 Recall
Percentage Correct
60
50
40
30
20
10
0
Restudy
Retrieval Practice
Concept Mapping

Table 1

Percentage of Correctly Produced Idea Units During Practice of Passages 1 and 2 as a Function of Condition for Experiments 1–3 (Standard Errors in Parenthesis)

	Retrieval practice		Concept mapping
	Cycle 1	Cycle 2	
Experiment 1			
Passage 1	58.7 (2.86)	59.3 (2.97)	70.7 (2.40)
Passage 2	50.3 (3.21)	52.9 (2.94)	63.2 (2.79)
Experiment 2			
Passage 1	39.9 (2.82)	39.7 (3.03)	50.3 (3.63)
Passage 2	44.9 (3.43)	43.2 (2.98)	51.8 (3.50)
Experiment 3			
Passage 1	50.9 (4.10)	52.3 (4.37)	64.7 (3.16)
Passage 2	51.7 (3.35)	50.1 (3.23)	74.6 (2.78)

Table 2

Mean Number of Intrusions From Passages 1 and 2 During Passage-3 Recall as a Function of Condition for Experiments 1–3 (Standard Errors in Parenthesis)

	Restudy	Retrieval practice	Concept mapping
Experiment 1	0.46 (0.14)	0.04 (0.04)	0.11 (0.06)
Experiment 2	0.68 (0.19)	0.04 (0.04)	0.14 (0.07)
Experiment 3	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)