Original Article

Combining Retrieval Practice and Generative Learning in Educational Contexts

Promises and Challenges

Julian Roelle[1], Judith Schweppe[2], Tino Endres[3], Andreas Lachner[4], Claudia von Aufschnaiter[5], Alexander Renkl[3], Alexander Eitel[6], Detlev Leutner[7], Ralf Rummer[8], Katharina Scheiter[9], and Andreas Vorholzer[10]

[1]Faculty of Philosophy and Educational Research, Ruhr University Bochum, Germany
[2]Faculty of Arts and Humanities, University of Passau, Germany
[3]Department of Psychology, University of Freiburg, Germany
[4]Faculty of Economics and Social Sciences, University of Tübingen, Germany
[5]Department of Physics, Justus Liebig University Giessen, Germany
[6]Department of Psychology, Justus Liebig University Giessen, Germany
[7]Faculty of Educational Sciences, University Duisburg-Essen, Germany
[8]Department of Psychology, University of Kassel, Germany
[9]Department of Educational Sciences, University of Potsdam, Germany
[10]School of Social Sciences and Technology, Technical University of Munich, Germany

Abstract: Engaging learners in practicing the retrieval of learned information fosters the consolidation of learners' mental representations and hence long-term retention. Retrieval practice research has enriched the instructional design literature by providing a wealth of evidence for these benefits of retrieval-based learning and thus emphasizing the value of means to consolidate knowledge. The present article makes the case that a fruitful next step could be to focus on the interplay between retrieval practice and generative activities. Rather than consolidating mental representations, generative activities should have as their main function the construction of coherent mental representations. Hence, from a theoretical perspective, generative activities and retrieval practice should functionally complement each other; hence, combinations of both activities might be particularly suitable to promote lasting learning. Given the challenge to beneficially combine these activities, we discuss open questions that could substantially advance both the retrieval practice and the generative learning field.

Keywords: retrieval practice, generative learning, learning activities, lasting learning

Die Kombination von Abrufübungen und generativem Lernen in Bildungskontexten. Chancen und Herausforderungen

Zusammenfassung: Das Üben des Abrufs von zuvor gelernten Informationen aus dem Gedächtnis fördert die Konsolidierung mentaler Repräsentationen und damit das langfristige Behalten. Durch eine Fülle an Belegen für diese Vorteile des abrufbasierten Lernens und die Betonung des Wertes von Maßnahmen zur Konsolidierung von Wissen, hat die Forschung zu Abrufübungen die Instruktionsdesignforschung deutlich bereichert. In dem vorliegenden Artikel argumentieren wir, dass ein sinnvoller nächster Schritt darin bestehen könnte, das Zusammenspiel zwischen Abrufübungen und generativen Lernaktivitäten zu untersuchen. Anstatt mentale Repräsentationen zu konsolidieren, besteht die Hauptfunktion von generativen Lernaktivitäten im Aufbau kohärenter mentaler Repräsentationen. Aus theoretischer Sicht sollten sich also generative Aktivitäten und Abrufübungen funktional ergänzen, so dass Kombinationen beider Arten von Aktivitäten besonders geeignet sein könnten, um nachhaltiges Lernen zu fördern. Angesichts der Herausforderung, diese Aktivitäten sinnvoll zu kombinieren, diskutieren wir offene Fragen, die sowohl die Forschung zu Abrufübungen als auch das Feld des generativen Lernens wesentlich voranbringen könnten.

Schlüsselwörter: Abrufübungen, generatives Lernen, Lernaktivitäten, nachhaltiges Lernen

Practicing the retrieval of previously learned information from memory fosters retention of this information compared to learning activities that do not require learners to recall information from memory, such as restudy or notetaking. The empirical foundation of this core notion of the learning activity of *retrieval practice* was laid over 100 years ago (e.g., Abott, 1909; Kühn, 1914; Witasek, 1907) but did not receive much attention by educational psychologists until Roediger and Karpicke's (2006) seminal study. In their study, the authors compared different follow-up activities to reading short text passages and found that retrieval practice (here: free recall) was inferior to restudy in terms of learning outcomes 5 minutes after the learning phase but superior to restudy 1 week later. Regarding publications in journals listed in the Web of Science, research on retrieval practice has since experienced exponential growth (see Karpicke, 2017; Yang et al., 2021). Although large parts of this research were conducted in the lab, the *retrieval practice research wave* has also spilled over to applied settings. This research suggests that retrieval practice can foster learning in authentic educational contexts. A recent meta-analysis by Yang et al. (2021) and a recent review by Agarwal and colleagues (2021) indicate that (individual) retrieval practice can be implemented to good effect in the classroom.

The extensive research on retrieval practice has highlighted the importance of *consolidation* as part of the learning process, which had only scarcely been considered in educational psychology-driven instructional-design research. Rather, based on generative learning theory (e.g., Craik & Lockhart, 1972; Fiorella & Mayer, 2016; Wittrock, 1974), the educational, psychology-driven, instructional-design research had largely focused on generative activities, which are crucial in *constructing* coherent mental representations of new knowledge and integrating them with learners' prior knowledge. This line of research yielded a wealth of well-established generative learning tasks that engage learners in the organization, elaboration, and generation of inferences, which contributes to quality mental representations and therefore to comprehension and transfer. For instance, tasks that require learners to self-explain previously read texts or provided examples (e.g., for recent overviews, see Bisra et al., 2018; Renkl, 2014) or journal writing tasks in which learners reflect on previously encountered content (for a recent overview, see Nückles et al., 2020) have been shown to contribute to quality mental representations. Similar results were found for learning-by-teaching tasks, in which learners explain previously studied content to a (fictitious) audience (for a recent overview, see Lachner et al., 2022), concept mapping tasks (for a recent overview, see Schroeder et al., 2018), or generative drawing tasks, which engage learners in constructing drawings that depict concepts and relations explained in expository texts (for recent overviews, see Leutner & Schmeck, 2021; Wu & Rau, 2019).

Regarding learning that goes beyond rote learning of factual knowledge but aims at long-lasting comprehension and the ability to apply the respective knowledge in new situations (in the following referred to as *meaningful learning*), it is reasonable to assume that both knowledge construction, which generative tasks can foster, and knowledge consolidation, which retrieval practice tasks can foster, are crucial (for an overview and further important future pathways in fostering lasting learning, see Richter, Berger et al., 2022, Ebersbach et al., 2022, and Richter, Nemeth et al., 2022). At this point, it is important to highlight that the different functions of retrieval practice and generative activities (consolidation vs. construction of mental representations) do not have completely different effects on learning outcomes. For example, the structural improvement of mental representations not only leads to deeper comprehension but can also reduce forgetting rates (e.g., Kintsch et al., 1990). Similarly, the consolidation of mental representations not only leads to reduced forgetting rates but can also increase the depth of understanding (Pan & Rickard, 2018; see also Heitmann et al., 2021). Crucially, however, the two types of tasks are nonetheless substantially different. For example, the literature indicates that retrieval practice leads to lower forgetting rates than generative tasks (e.g., Roelle & Berthold, 2017), and that generative tasks can lead to a deeper understanding than retrieval practice (e.g., Endres et al., 2017, domain: educational psychology; Larsen et al., 2013, domain: medicine; Roelle & Nückles, 2019, domain: educational psychology). Consequently, despite partly overlapping effects, the two types of learning tasks can be assumed to be distinctly different.

Except for a few studies that started exploring whether and how retrieval practice and generative learning tasks can be combined to good effects (e.g., Blunt & Karpicke, 2014; Endres et al., 2017; Larsen et al., 2013), to date the interplay between these different types of learning activities has largely been ignored. From a theoretical point of view, it can be assumed that the benefits of consolidating mental representations via retrieval practice should increase with increasing quality of the respective mental representations, which can be fostered via generative activities (see Roelle et al., 2022). Hence, exploring the interplay between these two types of learning activities might be promising in terms of optimizing meaningful learning. Furthermore, combining retrieval practice with generative learning might also be fruitful in terms of optimizing the degree to which practitioners take up retrieval-based learning activities. To date, retrieval practice tasks implemented in authentic contexts frequently engage learners in practicing retrieval of facts via multiple-choice quizzes (see Agarwal et al., 2021), which might be viewed as contributing relatively little to meaningful learning in educational practice. By contrast, generative learning tasks might appear to be better aligned to promote meaningful learning, as they are closely linked to comprehension and transfer (see Fiorella & Mayer, 2016). Consequently, it is reasonable to assume that, when educators and educational researchers think about optimizing *cognitive activation*, which is considered one of three core dimensions of high-quality instruction (e.g., Praetorius et al., 2018), engaging learners in generative activities would pop up in their minds more frequently than engaging learners in practicing retrieval of facts. Combining the more *popular* generative activities with retrieval practice could thus offer the chance for retrieval-based learning to be taken up in educational practice less reluctantly. A sensible next step in both retrieval practice and generative learning research could hence be designing and investigating settings in which both the construction and the consolidation of quality mental representations are fostered.

Retrieval Practice Tasks That Engage Learners in Generative Activities

When pursuing the goal of combining generative learning and retrieval practice, one first step could be providing learners with retrieval practice tasks that require them to engage in generative learning activities as well. The tasks might engage learners in both constructing and consolidating mental representations that contribute to comprehension. To date, many studies that investigated retrieval practice have implemented the integration of retrieval practice by providing learners with multiple-choice or short-answer quiz questions that tapped relatively simple factual knowledge but did not require learners to engage in any generative activities (see Agarwal et al., 2021). This retrieval practice appears to exert the greatest benefits when the same questions in the same format are provided in both the retrieval practice phase and the final test (Agarwal et al., 2021). That is, practicing retrieval with simple (mostly factual) multiple-choice or short-answer questions mainly fosters performance on the same simple, factual, multiple-choice or short-answer questions, which appears to be of relatively limited utility when it comes to meaningful learning. Hence, it could be a fruitful approach to provide learners with meaningful retrieval practice tasks that require explanations, problem-solving, or transfer of previously acquired knowledge to consolidate not only factual knowledge but also foster comprehension. However, one challenge in this endeavor could be that retrieval practice benefits might generally be harder to achieve with complex, meaningful tasks that require comprehension of the respective knowledge than simple, factual-knowledge questions, because both *direct* and *indirect* effects of retrieval practice are harder to realize under these conditions.

Direct effects means that the act of engaging in retrieval itself is beneficial (e.g., Karpicke, 2017). These benefits, however, seem to consistently occur only when learner performance on the respective retrieval practice tasks exceeds approx. 75 % (see Karpicke, 2017; Rowland, 2014). For several reasons, learners might show such high performance less frequently on meaningful retrieval practice tasks that require comprehension than on simple factual quiz questions. For instance, in meaningful tasks, learners usually need not only retrieve certain idea units or schemata but also apply the retrieved knowledge to solve new problems or generate new explanations. In such tasks, not only the retrieval of the respective knowledge but also its application can be prone to errors, which should reduce performance relative to factual knowledge questions. Furthermore, failing to retrieve only one of several crucial and related idea units in meaningful learning tasks can hinder learners from solving the entire task. Hence, compared to simple and often isolated factual knowledge questions, the detriments of failed retrieval are potentiated regarding performance on the respective task. That is, a learner who could successfully retrieve five of six idea units and hence correctly answer approx. 83 % of factual knowledge questions concerning a specific topic might score much lower on a task that requires combining the six idea units, because the one missing idea unit could hinder her from performing most of the required solutions steps or explanations (see Roelle & Berthold, 2017). For more complex retrieval practice tasks to be beneficial, it might therefore be necessary to implement means designed to improve initial retrieval success.

One way to ensure sufficient performance in meaningful retrieval practice tasks may be to provide learners with more extensive initial study phases and instructional support than is typically assigned in retrieval practice studies before learners are engaged in the respective retrieval practice tasks (see Rummer & Schweppe, 2022). Given sufficient time and instructional support before retrieval practice, learners might reach the required level of comprehension to perform sufficiently well even on more complex retrieval practice tasks. An alternative instructional means preventing learners from solving retrieval practice tasks that tap comprehension largely incorrectly could be providing the respective tasks in an *open-book format* (see also Rummer, 2021; Rummer & Schweppe, 2022). Other than in typical closed-book retrieval practice tasks, an open-book format entails access to the learning material (e.g., textbooks and notes) when responding to the respective tasks. This access to the learning material can help learners compensate for failed retrieval and hence foster performance compared to closed-book retrieval practice tasks (e.g., Agarwal et al., 2008; Waldeyer et al., 2020; Wenzel et al., 2022). However, one could also argue that open-book retrieval practice benefits less from direct effects because learners are likely to engage in less retrieval practice when they have access to the learning materials (Rummer et al., 2019). To date, however, the questions remain open whether prolonging as well as instructionally supporting the initial study phase(s) and providing retrieval practice tasks (at least partly) in an open-book format increase the direct effects of meaningful retrieval practice tasks.

Indirect effects refers to the influences that retrieval practice can have on metacognition, motivation, or the activation of prior knowledge, which can affect remedial learning activities (e.g., Arnold & McDermott, 2013; Endres et al. 2020, Endres & Renkl, 2022). The most prominent of those indirect effects likely is that retrieval practice tasks help learners notice what they do not know or cannot retrieve from memory and thus guide subsequent restudying. The combination of generative learning and retrieval practice tasks may influence this metacognitive monitoring effect. In contrast to retrieval practice tasks that consist of relatively simple short-answer or multiple-choice questions, it is unclear in which direction the metacognitive effects of a combined task would develop. On the one hand, combining retrieval practice and generative learning could improve metacognitive monitoring. As shown in text comprehension research (Prinz et al. 2020), in judging one's level of comprehension, the monitoring of comprehensive cues is important when complex content is to be learned. When engaging in a combined generative and retrieval practice task, learners could get metacognitive feedback not only about the correct recall of the content but also about their level of understanding. On the other hand, at the same time, monitoring might get more complex as not only recall performance but also performance on the generative part of the task must be monitored. This, in turn, could overwhelm learners. For instance, there is evidence that learners have substantial difficulties in accurately self-evaluating their performance in tasks that require not only retrieval from memory but also the generation of own examples (see Zamary et al., 2016, domain: cognitive psychology). In pure retrieval practice tasks, providing learners with standards (i.e., external representations of a correct answer to a task assignment) can enhance monitoring accuracy (e.g., Lipko et al., 2009; Rawson & Dunlosky, 2007). In tasks that require generative activities as well, however, the effects of providing standards on judgment accuracy are mixed (e.g., Zamary et al., 2016; see also Froese & Roelle, 2022, Waldeyer & Roelle, 2021). Hence, research on how learners could be supported to accurately monitor their performance in tasks that require both retrieval practice and generative activities might be needed to pave the way for monitoring-driven indirect benefits of such tasks.

Control or regulation activities informed by learners' metacognitive monitoring might get more difficult as well. In pure retrieval practice tasks that merely tap certain vocabulary or isolated facts, simply presenting the correct answer is sufficient to elicit beneficial regulation activities on the part of the learners. By contrast, in meaningful learning tasks, simply showing the correct answer is not beneficial (see Pan & Rickard 2018). To achieve beneficial regulation effects in meaningful tasks, one must provide more extensive elaborative relearning opportunities. In generative drawing, an activity that is mainly elicited in STEM domains, for instance, providing learners with instructor-generated drawings after learners have constructed their own drawings and providing them with prompts that are designed to elicit the crucial comparison processes seems to be a promising feedback procedure (e.g., Van Meter, 2001; Van Meter et al., 2006). However, in self-explanation tasks, a similar procedure was found to yield even detrimental effects because it seduces learners to reduce the effort they invest into generating accurate task solutions on their own (see Schworm & Renkl, 2006, domain: educational psychology; see also Roelle, Rahimkhani-Sagvand et al., 2017, domain: chemistry). Hence, the underlying mechanisms and boundary conditions of the outlined procedure are not yet fully understood.

Sequences and Switches Between Generative Learning and Retrieval Practice

In a certain sense, the outlined difficulties regarding retrieval practice tasks which also require learners to engage in generative activities all go back to the notion that learners need to have reached a certain level of comprehension before they perform the respective tasks (for similar arguments, see Roelle & Nückles, 2019; Rummer & Schweppe, 2022). From this perspective, consolidation via retrieval practice should be elicited only after learners have formed the respective quality mental representations (e.g., through engagement in generative learning tasks). However, this *construction-before-consolidation* or *generative-learning-before-retrieval-practice* assumption, reflected in part in the above-mentioned idea of extending the initial study phase before learners engage in retrieval, does not seem to apply in each setting. A recent study by Roelle et al. (2022) suggests that, in acquiring new concepts (here: from the domain of cognitive psychology), it can be more beneficial to engage learners in practicing the retrieval of the concepts before engaging them in elaborating on the respective concepts by generating own illustrative examples. Although this study arguably raises more questions than it can answer because it is one of the first studies on this topic, it indicates that retrieval practice can also play an important role in the early phases of knowledge acquisition and mental representation formation.

The basic idea why early retrieval practice that consolidates factual knowledge might be helpful is that learners can build on this consolidated factual knowledge in subsequent (generative) learning activities (see Rawson et al., 2015). However, this notion does not entail any specific predictions on how and why the consolidated knowledge facilitates subsequent learning. One specific mechanism through which early retrieval practice could contribute to forming quality mental representations is that it facilitates subsequent generative activities (see Roelle et al., 2022). Generative learning tasks are often implemented in an open-book format in which learners are not required to retrieve the idea units needed to perform the respective generative task from memory. Nevertheless, consolidating basic factual knowledge before engaging in generative learning tasks could reduce cognitive load because it frees learners from effortful rereading and holding the respective idea units in working memory while performing the generative activities. This effect, in turn, could foster the quality of the performed generative activities and ultimately enhance the quality of learners' mental representations. In addition, it could increase the probability that learners (try to) retrieve the information from memory even when they can look up the respective information (i.e., during open-book generative tasks) and thus result in even more retrieval practice. An alternative explanation for the superiority of engaging in retrieval practice first in the study by Roelle and colleagues (2022) could be that the generative tasks were too difficult and hence learners' performance on the generative tasks when performed first was too low (for findings that partly point to a similar argument, see O'Day & Karpicke, 2021). Future research needs to address these potential moderating factors to understand the role of early retrieval practice in forming quality mental representations.

At first glance, the question of which sequence of retrieval practice and generative activities is more effective may appear like the questions asked in preparation for future learning research. In these approaches, learners are required to retrieve knowledge from memory (e.g., in the context of solving new types of problems) before engaging in generative learning activities as well. However, at second glance, there is an important difference between these approaches and the above-mentioned sequences of generative and retrieval practice tasks. The above-mentioned line of argumentation refers to the case that both learning tasks are used as *follow-up tasks* to an initial study phase in which learners studied new content. By contrast, in preparation for future learning approaches such as productive failure (e.g., Kapur, 2016), problem-solving before instruction (e.g., Weaver et al., 2018), or inventing to prepare future learning (e.g., Schwartz et al., 2011), the tasks that require learners to retrieve prior knowledge while solving new problems are provided *before* learners receive initial instruction on the respective to-be-learned principles and concepts. In these settings, the retrieval does not mainly serve the function of consolidating existing knowledge but of making learners aware of errors and knowledge gaps to focus them on the main learning content and increase their curiosity in the subsequent learning phase, in which they are to engage in generative learning activities.

A further interesting and open question that relates to forming and consolidating comprehension through retrieval practice concerns the effects of mixtures of open- and closed-book formats of generative learning tasks (see Roelle & Nückles, 2022). When a generative learning task designed to help learners construct quality mental representations and hence reach a deep understanding of a certain topic is implemented in a closed-book format, it can be argued that – in theory – it should serve a consolidation function as well, because learners need to retrieve the respective idea units that are needed to perform the respective generative task from memory. As outlined above, when learners still need to form quality mental representations, a pure closed-book format of generative tasks likely would not be very effective because the direct effects of retrieval practice (because of low performance/low retrieval success) would be hard to realize and because the *generative part* of the task would likely suffer from the consequences of failed retrieval as well. However, compared to a purely open-book format, which is the established format for many generative learning tasks such as generative drawing, self-explanation, or journal writing tasks, a mixture of closed- and open-book formats might entail certain benefits. For instance, when learners are instructed to first perform a generative task as well as they can in a closed-book format and then correct and extend their generated products in a subsequent open-book format, at least the factual knowledge that learners retrieve to perform subsequent generative activities would be consolidated simultaneously to the construction of quality mental representations without the risk of failed construction. In the subsequent open-book phase, learners likely could realize the same quality of the generative part of the task as in a pure open-book format. When learners would be prompted to engage in this sequence twice (e.g., in generative drawing, learners could be prompted to generate the respective drawings in two consecutive cycles), in the second closed-book phase (i.e., the closed-book phase of the second cycle), the retrieval practice might even contribute to consolidating the level of comprehension reached in the first cycle. By this means, a repeated closed-then-open-book format might potentially contribute to both forming and consolidating quality mental representations and hence to lasting comprehension.

To prevent low retrieval rates and hence low direct effects of retrieval practice in the first closed-book phase of such (repeated) closed-then-open-book generative learning tasks, it might be useful to engage learners in deep processing the crucial idea units in the initial study phase. For instance, in the domain of chemistry, Hiller and colleagues (2020) found that prompts requiring learners to explain crucial idea units before they engaged in a closed-book generative task (here: a self-explanation task) resulted in generative activities of similar quality in comparison to a pure open-book format. Furthermore, it might be useful to inform learners about the purpose of the closed- and open-book phases in advance. Research on *informed training*, which basically entails that learners are informed about the reasons why certain procedures and activities would be beneficial, shows that learners' engagement in the respective activities and compliance with the respective instructions can be effectively increased through the provision of such metastrategic knowledge (e.g., Paris et al., 1982; Simon et al., 1987; see also Hübner et al., 2010). In closed-open-book formats of generative tasks, such informed training could help learners see the rationale behind making the task more difficult initially by performing it from memory (i.e., closed-book format), which could convince them to invest the required effort and hence overcome the difficulty.

Informing learners about the rationale of both closed- and open-book phases could also pave the way for a format in which learners can flexibly switch between closed- and open-book phases. Waldeyer and colleagues (2020, Exp. 2, domain: social psychology) showed that, in comparison to both a pure closed-book and a pure open-book format, providing learners with the opportunity to flexibly switch between both formats (starting with a closed-book format) fostered learning outcomes 1 week later. However, whether such a flexible-closed-open-switch format would be superior to a static closed-then-open-book format is an open question. For the flexible format to be effective, it is likely crucial to inform learners, not only about the rationale behind the closed-book phases, but also about the rationale behind the open-book phases. Recently, Roelle and Renkl (2020, domain: chemistry) found that, when learners were simply told to perform as much as possible of a task in a closed-book format and switch to the open-book format only when they could not retrieve a required idea unit or wanted to check their generated products (here: self-explanations), the flexible-closed-open-switch format was superior to a closed-book format only for learners with relatively low academic self-concepts – and even detrimental for learners with relatively high academic self-concepts. Although the underlying mechanisms of this moderation effect still need to be illuminated, Roelle and Renkl's (2020) findings suggest that a flexible-closed-open-switch format entails certain risks. In comparison to a static closed-then-open-book format, the flexible format might have the advantage that learners could potentially perform more parts of the respective generative tasks in a *retrieval mode* because, when only single idea units are missing, they can quickly look them up and switch back to working in a retrieval mode. By contrast, in the static format, learners would be forced to complete the respective task in an open-book format once they needed to look up idea units they could not retrieve. An advantage of the static format, however, could be that it potentially poses lower demands on learners' self-regulation skills. Furthermore, when retrieval success is already high because of extensive and instructionally supported initial study phases (see above), the need to proceed (too) early to an open-book phase might be relatively low, making the option to flexibly switch between open- and closed-book formats redundant.

Conclusion

Because not only consolidating mental representations but also constructing quality mental representations matters when it comes to meaningful learning, exploring the interplay between retrieval practice and generative learning might be a fruitful endeavor. Specifically, it is reasonable to assume that analyzing and optimizing the interplay of generative learning activities and retrieval practice could yield important implications regarding how to effectively support the consolidation of previously formed quality mental representations as well as regarding how to foster the formation of such representations. Concerning consolidating and forming quality mental representations, it is crucial to ensure that learners perform well on the tasks. Concerning consolidating comprehension, our main hypotheses that need to be tested in future research are that increasing instructional support in the initial learning phases or increasing the share of open-book retrieval practice are effective measures to improve performance during retrieval and hence optimize consolidation. In forming quality mental representations, our main predictions are that early practice of retrieval of factual knowledge and, compared with the established open-book format of generative learning tasks, a higher share of closed-book generative tasks and switches between closed- and open-book formats could be helpful.

At first glance, we seem to propose opposing hypotheses in arguing that both decreasing and increasing the share of a closed-book format could be beneficial. These hypotheses, however, differ in the baseline to which they refer. In consolidating (quality) mental representations, the established format provides learners with closed-book retrieval practice tasks. To enhance learner performance on these tasks and hence the direct effects of retrieval practice, including open-book phases or providing learners with the opportunity to switch to open-book phases on demand could be useful, at least in the initial phases of comprehension consolidation. In forming quality representations via generative learning activities, the established format is open-book. We argue that incorporating closed-book phases could potentially contribute to a certain degree of consolidation of the respective knowledge without interfering with the targeted knowledge construction.

Finally, it is important to highlight that our paper is not the first to call for research on the crossroads of retrieval practice and generative learning. For example, in his recent comprehensive review of the retrieval practice literature, Karpicke (2017, p. 23) declared "[...] a pressing need to integrate retrieval practice into existing educational activities [...]." Likewise, in their recent review of the generative drawing literature, Fiorella and Zhang (2018) called for research investigating the benefits of drawing as a retrieval activity. It will, however, likely be not trivial to combine the two types of tasks and activities in a way that ensures positive effects. Overcoming this difficulty, however, would be highly desirable and substantially advance both the retrieval practice and the generative learning field.

References

Abott, E. E. (1909). On the analysis of the factor of recall in the learning process. *The Psychological Review: Monograph Supplements, 11*(1), 159 – 177. https://doi.org/10.1037/h0093018

Agarwal, P. K., Karpicke, J. D., Kang, S. H. K., Roediger, H. L., & McDermott, K. B. (2008). Examining the testing effect with open- and closed-book tests. *Applied Cognitive Psychology, 22*(7), 861 – 876. https://doi.org/10.1002/acp.1391

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. https://doi.org/10.1007/s10648-021-09595-9

Arnold, K. M., & McDermott, K. B. (2013). Test-potentiated learning: Distinguishing between direct and indirect effects of tests. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*(3), 940 – 945. https://doi.org/10.1037/a0029199

Bisra, K., Liu, Q., Nesbit, J. C., Salimi, F., & Winne, P. H. (2018). Inducing self-explanation: A meta-analysis. *Educational Psychology Review, 30*(3), 703 – 725. https://doi.org/10.1007/s10648-018-9434-x

Blunt, J. R., & Karpicke, J. D. (2014). Learning with retrieval-based concept mapping. *Journal of Educational Psychology, 106*(3), 849 – 858. https://doi.org/10.1037/a0035934

Craik, F. I. & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, *11*(6), 671 – 684. https://doi.org/10.1016/S0022-5371(72)80001-X

Ebersbach, M., Lachner, A., Scheiter, K., & Richter, T. (2022). Using spacing to foster lasting learning in educational contexts: Promises and challenges. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie/German Journal of Developmental Psychology and Educational Psychology, 54*(4), 151 – 163. https://doi.org/10.1026/0049-8637/a000259

Endres, T., Carpenter, S., Martin, A., & Renkl, A. (2017). Enhancing learning by retrieval: Enriching free recall with elaborative prompting. *Learning and Instruction, 49*, 13 – 20. http://doi.org/10.1016/j.learninstruc.2016.11.010

Endres, T., Kranzdorf, L., Schneider, V., & Renkl, A. (2020). It matters how to recall: Task differences in retrieval practice. *Instructional Science, 48*(6), 699 – 728. https://doi.org/10.1007/s11251-020-09526-1

Endres, T., & Renkl, A. (2022). Indirekte Effekte von Abrufübungen: Intuitiv und doch häufig unterschätzt [Indirect effects of retrieval practice–Intuitive yet often underestimated]. *Unterrichtswissenschaft, 50, 75* – 98. https://doi.org/10.1007/s42010-021-00140-9

Fiorella, L., & Mayer, R. E. (2016). Eight ways to promote generative learning. *Educational Psychology Review*, *28*(4), 717 – 741. https://doi.org/10.1007/s10648-015-9348-9

Fiorella, L., & Zhang, Q. (2018). Drawing boundary conditions for learning by drawing. *Educational Psychology Review*, *30*(3), 1115 – 1137. https://doi.org/10.1007/s10648-018-9444-8

Froese, L., & Roelle, J. (2022). Expert example standards but not idea unit standards help learners accurately evaluate the quality of self-generated examples. *Metacognition and Learning, 17*, 565 – 588. https://doi.org/10.1007/s11409-022-09293-z

Heitmann, S., Obergassel, N., Fries, S., Grund, A., Berthold, K., & Roelle, J. (2021). Adaptive practice quizzing in a university lecture: A pre-registered field experiment. *Journal of Applied Research in Memory and Cognition, 10*(4), 603 – 620. https://doi.org/10.1016/j.jarmac.2021.07.008

Hiller, S., Rumann, S., Berthold, K., & Roelle, J. (2020). Example-based learning: Should learners receive closed-book or open-book self-explanation prompts? *Instructional Science, 48,* 623 – 649. https://doi.org/10.1007/s11251-020-09523-4

Hübner, S., Nückles, M., & Renkl, A. (2010). Writing learning journals: Instructional support to overcome learning-strategy deficits. *Learning and Instruction, 20*(1), 18 – 29. https://doi.org/10.1016/j.learninstruc.2008.12.001

Kapur, M. (2016). Examining productive failure, productive success, unproductive failure, and unproductive success in learning. *Educational Psychologist, 51*(2), 289 – 299. https://doi.org/10.1080/00461520.2016.1155457

Karpicke, J. D. (2017). Retrieval-based learning: A decade of progress. In J. H. Byrne (Ed.), *Learning and memory: A comprehensive reference* (2nd ed., pp. 487 – 514). Academic Press. https://doi.org/10.1016/B978-0-12-809324-5.21055-9

Kintsch, W., Welsch, D., Schmalhofer, F., & Zimny, S. (1990). Sentence memory: A theoretical analysis. *Journal of Memory and Language, 29*(2), 133 – 159. https://doi.org/10.1016/0749-596X(90)90069-C

Kühn, A. (1914). Über Einprägung durch Lesen und durch Rezitieren [About memorization through reading and recitation]. *Zeitschrift für Psychologie, 68*, 396 – 481.

Lachner, A., Hoogerheide, V., van Gog, T., & Renkl, A. (2022). Learning-by-teaching without audience presence or interaction: When and why does it work? *Educational Psychology Review, 34*(2), 575 – 607. https://doi.org/10.1007/s10648-021-09643-4

Larsen, D. P., Butler, A. C., & Roediger, H. L., III (2013). Comparative effects of test-enhanced learning and self-explanation on long-term retention. *Medical Education, 47*(7), 674 – 682. https://doi.org/10.1111/medu.12141

Leutner, D., & Schmeck, A. (2021). The drawing principle in multimedia learning. In R. Mayer & L. Fiorella (Eds.), *The Cambridge handbook of multimedia learning* (pp. 360 – 369). Cambridge University Press. https://doi.org/10.1017/9781108894333.038

Lipko, A. R., Dunlosky, J., Hartwig, M. K., Rawson, K. A., Swan, K., & Cook, D. (2009). Using standards to improve middle school students' accuracy at evaluating the quality of their recall. *Journal of Experimental Psychology: Applied*, *15*(4), 307 – 318. https://doi.org/10.1037/a0017599

Nückles, M., Roelle, J., Glogger-Frey, I., Waldeyer, J., & Renkl, A. (2020). The self-regulation-view in writing-to-learn: Using journal writing to optimize cognitive load in self-regulated learning. *Educational Psychology Review, 32*(4), 1089 – 1126. https://doi.org/10.1007/s10648-020-09541-1

O'Day, G. M., & Karpicke, J. D. (2021). Comparing and combining retrieval practice and concept mapping. *Journal of Educational Psychology, 113*(5), 986 – 997. https://doi.org/10.1037/edu0000486

Pan, S. C., & Rickard, T. C. (2018). Transfer of test-enhanced learning: Meta-analytic review and synthesis. *Psychological Bulletin*, *144*(7), 710 – 756. https://doi.org/10.1037/bul0000151

Paris, S. G., Newman, R. S., & McVey, K. A. (1982). Learning the functional significance of mnemonic actions: A microgenetic study of strategy acquisition. *Journal of Experimental Child Psychology, 34*(3), 490 – 509. https://doi.org/10.1016/0022-0965(82)90073-X

Praetorius, A.-K., Klieme, E., Herbert, B., & Pinger, P. (2018). Generic dimensions of teaching quality: The German framework of Three Basic Dimensions. *ZDM*, *50*(3), 407 – 426. https://doi.org/10.1007/s11858-018-0918-4

Prinz, A., Golke, S., & Wittwer, J. (2020). To what extent do situation-model-approach interventions improve relative metacomprehension accuracy? Meta-analytic insights. *Educational Psychology Review, 32*(4), 917 – 949. https://doi.org/10.1007/s10648-020-09558-6

Rawson, K. A., & Dunlosky, J. (2007). Improving students' self-evaluation of learning for key concepts in textbook materials. *European Journal of Cognitive Psychology*, *19*(4 – 5), 559 – 579. https://doi.org/10.1080/09541440701326022

Rawson, K. A., Thomas, R. C., & Jacoby, L. L. (2015). The power of examples: Illustrative examples enhance conceptual learning of declarative concepts. *Educational Psychology Review, 27*(3), 483 – 504. https://doi.org/10.1007/s10648-014-9273-3

Renkl, A. (2014). Towards an instructionally-oriented theory of example-based learning. *Cognitive Science, 38*(1), 1 – 37. https://doi.org/10.1111/cogs.12086

Richter, T., Berger, R., Ebersbach, M., Eitel, A., Endres, T., Borromeo Ferri, R., Hänze, M., Lachner, A., Leutner, D., Lipowsky, F., Nemeth, L., Renkl, A., Roelle, J., Rummer, R., Scheiter, K., Schweppe, J., von Aufschnaiter, C., & Vorholzer, A. (2022). How to promote lasting learning in schools: Theoretical approaches and an agenda for research. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie/German Journal of Developmental Psychology and Educational Psychology, 54*(4), 135 – 141. https://doi.org/10.1026/0049-8637/a000258

Richter, T., Nemeth, L., Berger, R., Borromeo Ferri, R., Hänze, M. & Lipowsky, F. (2022). Using interleaving to foster inductive learning in educational contexts: Promises and challenges. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie/German Journal of Developmental Psychology and Educational Psychology, 54*(4), 164 – 175. https://doi.org/10.1026/0049-8637/a000260

Roediger III, H. L., & Karpicke, J. D. (2006). Test-enhanced learning: Taking memory tests improves long-term retention. *Psychological Science, 17*(3), 249 – 255. https://doi.org/10.1111/j.1467-9280.2006.01693.x

Roelle, J., & Berthold, K. (2017). Effects of incorporating retrieval into learning tasks: The complexity of the tasks matters. *Learning and Instruction, 49*, 142 – 156. https://doi.org/10.1016/j.learninstruc.2017.01.008

Roelle, J., Froese, L., Krebs, R., Obergassel, O., & Waldeyer, J. (2022). Sequence matters! Retrieval practice before generative learning is more effective than the reverse order. *Learning and Instruction, 80*, Article 101634. https://doi.org/10.1016/j.learninstruc.2022.101634

Roelle, J., & Nückles, M. (2019). Generative learning versus retrieval practice in learning from text: The cohesion and elaboration of the text matters. *Journal of Educational Psychology, 111*(8), 1341 – 1361. https://doi.org/10.1037/edu0000345

Roelle, J., & Nückles, M. (2022). Zwei auf einen Streich? Der Nutzen der Integration von Abrufübung in Aufgaben zur Verständnisförderung [Two in one go? The benefits of integrating retrieval practice into generative learning tasks]. *Unterrichtswissenschaft*, *50*, 53 – 73. https://doi.org/10.1007/s42010-021-00134-7

Roelle, J., Rahimkhani-Sagvand, N., & Berthold, K. (2017). Detrimental effects of immediate explanation feedback. *European Journal of Psychology of Education, 32*(3), 367 – 384. https://doi.org/10.1007/s10212-016-0317-6

Roelle, J., & Renkl, A. (2020). Does an option to review instructional explanations enhance example-based learning? It depends on learners' academic self-concept. *Journal of Educational Psychology, 112*(1), 131 – 147. https://doi.org/10.1037/edu0000365

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. https://doi.org/10.1037/a0037559

Rummer, R. (2021). Der Testungseffekt beim Lernen mit Texten: Ein Beispiel für das schwierige Verhältnis zwischen Grundlagenforschung und Anwendung. *Psychologische Rundschau, 72*(4), 259 – 272. https://doi.org/10.1026/0033-3042/a000518

Rummer, R., & Schweppe, J. (2022). Komplexität und der Testungseffekt: Die mögliche Bedeutung der Verständnissicherung für den Nutzen von Abrufübung bei komplexem Lernmaterial. *Unterrichtswissenschaft, 50*, 37 – 52. https://doi.org/10.1007/s42010-021-00137-4

Rummer, R., Schweppe, J., & Schwede, A. (2019). Open-book versus closed-book tests in university classes: A field experiment. *Frontiers in Psychology*, *10*, 463. https://doi.org/10.3389/fpsyg.2019.00463

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. https://doi.org/10.1007/s10648-017-9403-9

Schwartz, D. L., Chase, C. C., Oppezzo, M. A., & Chin, D. B. (2011). Practicing versus inventing with contrasting cases: The effects of telling first on learning and transfer. *Journal of Educational Psychology*, *103*(4), 759 – 775. https://doi.org/10.1037/a0025140

Schworm, S., & Renkl, A. (2006). Computer-supported example-based learning: When instructional explanations reduce self-explanations. *Computers & Education, 46*(4), 426 – 445. https://doi.org/10.1016/j.compedu.2004.08.011

Simon, T., McShane, J., & Radley, S. (1987). Learning with microcomputers: Training primary school children on a problem-solving program. *Applied Cognitive Psychology*, *1*(1), 35 – 44. https://doi.org/10.1002/acp.2350010106

Van Meter, P. (2001). Drawing construction as a strategy for learning from text. *Journal of Educational Psychology*, *93*(1), 129 – 140. https://doi.org/10.1037/0022-0663.93.1.129

Van Meter, P., Aleksic, M., Schwartz, A., & Garner, J. (2006). Learner-generated drawing as a strategy for learning from content area text. *Contemporary Educational Psychology*, *31*(2), 142 – 166. https://doi.org/10.1016/j.cedpsych.2005.04.001

Waldeyer, J., Heitmann, S., Moning, J., & Roelle, J. (2020). Can generative learning tasks be optimized by incorporation of retrieval practice? *Journal of Applied Research in Memory and Cognition, 9*(3), 355 – 369. https://doi.org/10.1016/j.jarmac.2020.05.001

Waldeyer, J., & Roelle, J. (2021). The keyword effect: A conceptual replication, effects on bias, and an optimization. *Metacognition and Learning, 16,* 37 – 56. https://doi.org/10.1007/s11409-020-09235-7

Weaver, J. P., Raymond, J. C., DeCaro, D. A., & DeCaro, M. S. (2018). Reverse the routine: Problem solving before instruction improves conceptual knowledge in undergraduate physics. *Contemporary Educational Psychology, 52*, 36 – 47. https://doi.org/10.1016/j.cedpsych.2017.12.003

Wenzel, K., Schweppe, J., & Rummer, R. (2022). Are open-book tests still as effective as closed-book tests even after a delay of 2 weeks? *Applied Cognitive Psychology, 36*(3), 688 – 707. https://doi.org/10.1002/acp.3943

Witasek, S. (1907). Über Lesen und Rezitieren in ihren Beziehungen zum Gedächtnis [About reading and reciting in their relations to memory]. *Zeitschrift für Psychologie, 44*, 161 – 185.

Wittrock, M. C. (1974). Learning as a generative process. *Educational Psychologist*, *11*(2), 87 – 95. https://doi.org/10.1080/00461527409529129

Wu, S. P. W., & Rau, M. A. (2019). How students learn content in science, technology, engineering, and mathematics (STEM) through drawing activities. *Educational Psychology Review, 31*(1), 87 – 120. https://doi.org/10.1007/s10648-019-09467-3

Yang, C., Luo, L., Vadillo, M. A., Yu, R., & Shanks, D. R. (2021). Testing (quizzing) boosts classroom learning: A systematic and meta-analytic review. *Psychological Bulletin, 147*(4), 399 – 435. https://doi.org/10.1037/bul0000309

Zamary, A., Rawson, K. A., & Dunlosky, J. (2016). How accurately can students evaluate the quality of self-generated examples of declarative concepts? Not well, and feedback does not help. *Learning and Instruction*, *46*, 12 – 20. https://doi.org/10.1016/j.learninstruc.2016.08.002

Funding
This article is the product of a collaboration of the authors within the research unit “Lasting Learning: Cognitive Mechanisms and Effective Instructional Implementation” (FOR 5254), funded by the German Research Foundation (DFG) (project no. 450142163; grants AU 155/12 – 1, LA 4009/2 – 1, RE 1040/20 – 1, RO 5012/7 – 1, RU 891/8 – 1, SCHW 1514/1 – 1).
Open access publication enabled by the Ruhr University Bochum.

ORCID
Julian Roelle
https://orcid.org/0000-0001-9581-9120
Judith Schweppe
https://orcid.org/0000-0002-8661-3072
Tino Endres
https://orcid.org/0000-0003-3334-4064
Andreas Lachner
https://orcid.org/0000-0001-5866-7164
Claudia von Aufschnaiter
https://orcid.org/0000-0002-6522-7084
Alexander Renkl
https://orcid.org/0000-0001-7971-0619
Alexander Eitel
https://orcid.org/0000-0001-6403-3467
Detlev Leutner
https://orcid.org/0000-0001-5308-9094
Ralf Rummer
https://orcid.org/0000-0002-2568-0925
Katharina Scheiter
https://orcid.org/0000-0002-9397-7544
Andreas Vorholzer
https://orcid.org/0000-0002-0915-2516

Prof. Dr. Julian Roelle
Faculty of Philosophy and Educational Research
Ruhr University Bochum
Universitätsstraße 150
44780 Bochum
Germany
julian.roelle@ruhr-uni-bochum.de