Educational Psychology Review (2023) 35:75
https://doi.org/10.1007/s10648-023-09784-8

REVIEW ARTICLE

# Combining Retrieval Practice with Elaborative Encoding: Complementary or Redundant?

**Mark A. McDaniel**[1]

Accepted: 2 June 2023 / Published online: 11 July 2023


## Abstract
The benefits of retrieval practice (practice testing) are pervasive across various materials, learning conditions, and criterial tasks, and consequently researchers and educators have enthusiastically recommended retrieval practice for educational applications. Less research has been devoted to examining the effect of combining retrieval practice with other evidence-based learning strategies; this article focuses on an emerging literature that examines the outcomes of combining potent elaborative encoding methods with retrieval practice. Theoretically, several possibilities can be identified. Augmenting retrieval practice with effective encoding strategies could significantly improve learning relative to retrieval practice alone through complementary mechanisms of each or through effective encoding catalyzing retrieval practice effects. Alternatively, effective encoding combined with retrieval practice might not improve learning (relative to retrieval practice alone), because the processing produced by elaborative encoding strategies is overly redundant with those promoted by retrieval practice. The extant literature, which has focused on everyday learning tasks (e.g., name learning) and educationally relevant tasks ranging from learning of arbitrary associations (e.g., new vocabulary meanings) to learning from connected discourse, is reviewed, and helps inform these possibilities. The findings largely converge on the conclusion that incorporating elaborative encoding techniques with retrieval practice prior to, but not concurrent with, retrieval practice provides a boost for learning and retention.

**Keywords** Retrieval practice · Elaborative encoding · Elaboration prior to retrieval · Elaboration concurrent with retrieval · Test-enhanced learning

This article is part of the Topical Collection on *Test-Enhanced Learning and Testing in Education: Contemporary Perspectives and Insights*.

✉ Mark A. McDaniel
markmcdaniel@wustl.edu

[1] Department of Psychological and Brain Sciences and Center for Integrative Research on Cognition, Learning and Education, Washington University in St. Louis, St. Louis, MO 63130, USA

The strategy of using retrieval practice (e.g., practice testing) as a learning tool has been advocated in the psychological and educational literatures for over a hundred years (e.g., Gates, 1917; see Roediger & Karpicke, 2006a, for a review), and research on retrieval practice has increased substantially in the past two decades. The findings from laboratory and classroom studies demonstrate that retrieval practice supports robust learning and retention, as indicated by several recent reviews and metanalyses (Adesope et al., 2017; Agarwal et al., 2021; Pan & Rickard, 2018; Rowland, 2014; Trumbo et al., 2021). Indeed, the benefits of retrieval practice are so pervasive across various materials, learning conditions, and criterial tasks that one influential learning-strategy review rated the utility of retrieval practice as "high" for educational applications (Dunlosky et al., 2013).

Perhaps because retrieval practice is considered by some as "the best strategy we've got" (a sentiment presented by an instructor in a cognitive course), much less research has been devoted to examining the effect of combining retrieval practice with other evidence-based learning strategies. (One exception is an emerging literature examining benefits of combining retrieval practice with spacing relative to retrieval practice without spacing; Latimier et al., 2021, provide a review.) In this article, I focus on the potential outcomes of combining potent encoding methods with retrieval practice. Typically, retrieval practice effects have been evaluated when retrieval practice follows a learning session in which learners are either instructed to read the to-be-learned material or are simply told to study the material for a later test (on rare occasion retrieval practice is implemented in the absence of initial study; e.g., Butler & Roediger, 2008). The question posed herein is whether retrieval practice effects could be augmented, and perhaps substantially so, if learners were engaged in effective elaboration during the initial learning session. I also consider a related question drawing on paradigms in which elaboration is concurrently engaged during retrieval (rather than prior to retrieval).

To inform this issue, the typical experimental design has been to implement a control condition like that just described (no elaborative strategy, no retrieval practice), an elaboration-only condition (an elaboration strategy is instructed but there is no retrieval practice), a retrieval practice-only condition (no instructed elaboration strategy) and a combined condition with both elaborative encoding and retrieval practice. The question is theoretically interesting because the outcomes are uncertain. As developed next, several theoretical possibilities are plausible. There is also clear applied value in determining if, and to what extent, effective elaborative encoding methods increase the benefits of retrieval practice. Were that to be the case, then even more potent uses of retrieval practice could be implemented (cf. Dunlosky et al., 2013).

## Theoretical Approaches

Two general approaches assume that learning is improved when retrieval practice is augmented with effective encoding strategies (relative to retrieval practice alone). The most straightforward is that effective encoding strategies and retrieval practice serve complementary roles for learning and retention (the *complementary* view;

e.g., Cummings et al., 2023; Miyatsu & McDaniel, 2019; Morris et al., 2005; Roelle et al., 2022). Elaborative encoding strategies such as generating explanations and generating examples benefit learning and retention by activating prior knowledge and fostering connections between this prior knowledge and the to-be-learned material (McDaniel & Donnelly, 1996; Pressley et al., 1988; Roelle & Nückles, 2019; Rosenshine et al., 1996; Seifert, 1993). For more impoverished materials (word lists; paired associate lists) successful elaborative strategies can include imagery (as indicated by a robust literature, e.g., see Clark & Paivio, 1991) and creating mnemonic linkages (e.g., Bellezza, 1987; Levin & Levin, 1990). In short, a wide range of elaborative strategies can foster learning and retention through creating richer and more cohesive representations of the target material. By contrast, on some views, retrieval practice is thought to generally strengthen later retrieval (possibly by reinforcing retrieval routes, Bjork, 1995, or by narrowing retrieval constraints, Thomas & McDaniel, 2013, Pyc & Rawon, 2010).

Thus, the theoretical idea is that effective encoding strategies produce robust representations of the material to be learned and retrieval practice then serves to improve retrieval of that material or consolidation in memory (Roelle et al., 2022). In short, each technique (effective encoding, retrieval practice) is effective because of different mechanisms, with these mechanisms providing boosts to memory in different ways. Empirical support for the idea that the mnemonic mechanisms of retrieval practice do not rest on elaborative encoding has been reported by Karpicke and Smith (2012). Accordingly, combining effective encoding strategies with retrieval practice should produce superior learning and retention over retrieval practice alone. A more precise prediction, based on the assumption that the mechanisms of effective encoding and retrieval practice are different, is that combining the two should produce additive effects (Fritz et al., 2007).

The second approach is perhaps more provocative. This approach suggests that there are boundary conditions (but still educationally relevant contexts) that minimize or disfavor strong effects of retrieval practice. These conditions can include situations in which (a) learning requires the acquisition of arbitrary associations (e.g., new vocabulary and terminology, biology taxonomies, scientists' names and their contributions) or is based on material that is not overly cohesive (e.g., low-cohesion text), and (b) retrieval practice is limited (e.g., one or two retrieval practice trials), perhaps because of learners' or instructors' time constraints (e.g., see Miyatsu & McDaniel, 2019, for absence of retrieval practice effects in learning new vocabulary with limited retrieval practice, and Roelle & Nückles, 2019, Experiment 2, for just nominal retrieval practice effects after one practice trial in learning low-cohesion text). In these conditions, the amount of target material retrieved might be relatively low, thereby restricting retrieval benefits to a limited amount of material (Roelle et al., 2022; Rowland, 2014). Accordingly, even less effective restudy, which allows review of all the material, could produce final test performance equivalent to that of retrieval practice (e.g., consistent with the Kornell et al., 2011, theoretical account of retrieval practice). In these cases, this second approach suggests that using effective encoding strategies during initial study catalyzes subsequent retrieval (labeled the *catalytic* view, following Miyatsu & McDaniel, 2019). The idea is that rich, elaborative encodings are more easily retrieved, thereby supporting effective retrieval and the consequent benefit of retrieval practice. Put another way, the view is that in certain learning situations, as just outlined, implementing effective encoding strategies leverages retrieval-practice benefits that are disproportionately greater than when effective encoding strategies are not used.

A sharply different theoretical possibility is that elaborative encoding involves processing that overlaps with or is redundant with that stimulated by retrieval practice (labeled the *redundancy* view; Cummings et al., 2023; Miyatsu & McDaniel, 2019). In particular, several accounts of the retrieval practice effect converge on the idea that retrieval enhances memory by stimulating elaboration. For instance, the elaborative retrieval account (Carpenter, 2009, 2011) suggests that the process of cue-guided retrieval involves generating elaborative information that might provide additional retrieval routes to the target (see also McDaniel et al., 1989). In a similar vein, the encoding-variability account (McDaniel & Masson, 1985) holds that retrieval produces a richer, more variable encoding of the studied information, again providing multiple routes for subsequent retrieval (see also Kang, 2010; Roediger & Butler, 2011). These mechanisms echo those assumed to underlie the value of elaborative encoding: elaborative encoding produces a richer network of information that affords multiple routes for retrieval.

The upshot is that elaborative encoding and retrieval practice would activate very similar memory processes, thereby rendering elaborative encoding techniques redundant. Therefore, retrieval practice effects would not be expected to be enhanced when elaborative encoding is also engaged. (Findings demonstrate that combining two conditions that produce redundant processing will not be more effective than one of the methods alone (Walsh & Jenkins, 1973) and that adding a processing task that is redundant with processing already engaged in a control condition does not significantly improve memory performance (Einstein et al., 1990); for review, see McDaniel & Butler, 2011.) It is worth mentioning that a modified theoretical analysis might underpin "redundancy" effects (no benefit of elaborative encoding for retrieval practice effects), if found. Multiple retrieval practice trials are more effective than fewer trials (e.g., Roediger & Karpicke, 2006b), accordingly, if elaborative encoding recruits similar processing to that of retrieval practice, then elaborative encoding (at study) could potentially mimic an additional retrieval practice trial and improve performance[1]. A redundancy pattern could be observed, however, were it the case that retrieval practice effects hinge on stimulating elaboration that provides more potent mnemonic components than that of elaborative encoding per se (cf. Whiffin & Karpicke, 2017). For instance, the episodic context account of retrieval benefits suggests that retrieval practice reinstates the original encoding context, thereby enriching that context with contextual details present during retrieval, which in turn would aid subsequent retrieval (Karpicke et al., 2014). Elaborative encoding per se would not necessarily confer that context reinstatement advantage (Whiffin & Karpicke, 2017).

In the following sections, I review the available literature on combining elaborative encoding techniques with retrieval practice. The review is organized in terms of

[1] I thank John Dunlosky for raising this possibility.

**Table 1** Overview of studies combining elaboration with retrieval practice, specifying to-be-learned materials, the elaborative strategies, the kind and quantity of retrieval practice, retention interval, summative test format, and the major outcomes for each study

| Materials | Elaboration strategy | Retrieval practice | Retention interval | Final test | Results | Study |
|---|---|---|---|---|---|---|
| First-last names | Semantic: find meaning in parts of names | 3 cued-recall trials of last name, no feedback | 5 min | Cued recall of last name | Complementary: combined RP-elaboration effects (additive) | Morris et al., (2005, Exp. 1) |
| First-last names plus person's photo | Elaboration: meaning + like-name familiar person image | 3 cued-recall trials of full name cued with photo, feedback | 5 min | Cued recall of both names (cued with photograph) | Complementary: combined RP-elaboration effects (additive) | Morris et al., (2005, Exp. 2) |
| Foreign-language vocabulary | Keyword-encoding mnemonic | 4 cued-recall trials of German word cued with English, no feedback | Short delay, then 1-week delay test | Recall cued with English; recall cued with German word | Redundancy: no advantage of combined versus RP alone | Fritz et al., (2007, Exp. 3) |
| Foreign-language vocabulary | Keyword-encoding mnemonic | 2 cued-recall trials of English translation, feedback | 2 days | Cued recall of English translation | Catalytic: RP effects found only when combined with elaboration | Miyatsu & McDaniel (2009, Exp. 1) |
| Foreign-language vocabulary | Keyword-encoding mnemonic | 2 cued-recall trials of English translation, feedback | 1 week | Cued recall of English translation | Catalytic | Miyatsu & McDaniel (2009, Exp. 2) |
| Foreign-language vocabulary | Keyword-encoding mnemonic | 2 or 4 cued-recall trials of English translation, feedback | 1 week | Cued recall of English translation | Catalytic with 2 RP trials; complementary with 4 RP trials (approximately additive) | Miyatsu & McDaniel (2009, Exp. 3) |
| Unfamiliar English vocabulary | Keyword-encoding mnemonic | Cued recall of meaning to criterion (many trials), feedback with interspersed study periods | 1 week | Cued recall of meaning (cued with the vocabulary word) | Redundancy: no significant advantage of combined versus RP only | Karpicke & Smith (2012, Exp. 1 and 2) |
| Scientists' names and their contributions | Keyword-encoding mnemonic | 2 cued-recall trials of contributions, feedback | 2 days | Cued recall of contribution (cued with scientist's name) | Complementary (approximately additive) | Cummings et al., (2023, Exp. 1 and 2) |

**Table 1** (continued)

| Materials | Elaboration strategy | Retrieval practice | Retention interval | Final test | Results | Study |
|---|---|---|---|---|---|---|
| Lecture | Elaborate, relate to knowledge (concurrent with retrieval) | 1 free recall trial, no feedback | 1 week | Comprehension and fact questions (both short-answer) | Combined better than RP only for comprehension questions | Endres et al. (2017) |
| Didactic text: high cohesion and elaboration | Elaborate by highlighting connections of main content and giving examples (concurrent with retrieval when combined with RP) | 1 free recall trial, no feedback | 1 week | Comprehension and reproduce content (both short answer)<br>Recognition of statements | *Comprehension*: complementary (additive, but n.s.)<br>*Recognition*: no effects of RP, elaboration, or two combined<br>*Reproduction*: same as recognition but combined best (n.s.) | Roelle & Nückles (2019, Exp. 1) |
| Didactic text: low cohesion and elaboration | Same as for cohesive text (previous entry) | 1 free recall trial, no feedback | 1 week | Comprehension and reproduce content (both short-answer)<br>Transfer questions (solve problems based on content) | *Comprehension and transfer*: significant effect of elaboration alone but not RP alone or combined with elaboration<br>*Reproduction*: no effects of RP, elaboration, or two combined | Roelle & Nückles (2019, Exp. 2) |
| Short science text | Concept mapping (concurrent with retrieval when combined with RP) | 2 free recall trials; reread passage after 1st recall (i.e., feedback after 1st recall) | 1 week | Factual and inference questions (both short answer) | *Factual*: redundancy--effect of each alone but no advantage of combined<br>*Inference*: no effect of elaboration alone or combined | Blunt & Karpicke (2014, Exp. 2) |

**Table 1** (continued)

| Materials | Elaboration strategy | Retrieval practice | Retention interval | Final test | Results | Study |
|---|---|---|---|---|---|---|
| Medical-school lesson | Self-explanation (concurrent with retrieval when combined with RP) | 4 short-answer test trials, with feedback | 6 months | Essay for managing a patient based on a clinical scenario | Effects of RP and elaboration alone; nominal combined effect | Larsen et al. (2013) |

*RP* retrieval practice

the learning task (material) examined, starting with an everyday task and proceeding to educationally relevant learning tasks with increasingly complex materials. Though the number of available studies is somewhat limited, this literature is nevertheless informative when viewed from the lens of the theoretical possibilities just outlined. Table 1 provides an overview of these studies, highlighting the elaborative techniques implemented, the retrieval-practice parameters, and the major outcomes for each study.

## Name Learning

A common everyday learning task is learning and remembering people's names. Names are difficult to remember (Cohen & Burke, 1993; Valentine et al., 1996), and surveys have shown that many would like to better remember people's names (Higbee, 2001). Based on a repeated retrieval practice technique, Morris and Fritz (2000, 2002) developed a "name game" strategy to enhance name learning. The key question aligning with the scope of this article is whether combining elaborative encoding with repeated retrieval practice produces learning gains relative to that obtained with retrieval practice alone. A study by Morris et al. (2005) provided consistent and compelling results informing this question.

Morris et al. (2005) manipulated the absence/presence of retrieval practice and the absence/presence of a semantic elaboration technique that could be applied to name learning. The technique involves attempting to find meaning in parts of a name (e.g., for *Herrmann* one could think of airman) or in the whole name (e.g., *Baker*). Participants in the semantic elaboration conditions were instructed to think of meanings of names and given examples. For instance, to learn *Linda Fielding* the suggestion was to think of someone you know named *Linda* and imagine her fielding in a cricket game. In Experiment 1, college students attempted to learn first and last name pairings, such that they could provide the last name when cued with the first name. All four learning groups (reflecting the factorial combination of absence/presence of retrieval practice and absence/presence of instructions for semantic elaboration) received four repetitions of each name; the repetitions were presented in an expanding schedule (1, 5, and 9 items were interleaved between the 1st–2nd, 2nd–3rd, and 3rd–4th, repetitions respectively). The study list consisted of 20 first-last name pairings and 9 filler names to enable the expanding repetition schedule. In the two groups with retrieval practice (retrieval practice following uninstructed study; retrieval practice following semantic-elaboration study), one study opportunity for each name was followed by 3 retrieval practice trials in which the first name was presented, and last name recall was attempted; no feedback was provided after the retrieval attempts. Five minutes after the learning session, participants attempted to recall the last names when cued with the first names.

As can be seen in Table 2 (first row), retrieval practice alone (2nd column) and semantic elaboration alone (3rd column) produced substantial increases in the proportion of last names correctly recalled relative to the uninstructed (control) repeated study condition (1st column). Note that these effects reflected very large effect sizes (Cohen's $d$ = 3.1 and 2.6, respectively), with semantic elaboration

**Table 2** Mean proportion correct on the final name-recall test as a function of condition from Experiments 1 and 2 in Morris et al. (Morris et al. (2005). Strategies for learning proper names: Expanding retrieval practice, meaning and imagery. *Applied Cognitive Psychology, 19*(6), 779–798)

| | SSSS | SRRR | EEEE | ERRR |
|---|---|---|---|---|
| Experiment 1 | .17 | .45 | .42 | .70 |
| Experiment 2 | .16 | .38 | .28 | .53 |

Experiment 2 but not Experiment 1 provided correct answer feedback after each retrieval attempt and tested secondary school participants

*SSSS* repeated study control condition, *SRRR* retrieval practice condition (with three retrieval practice attempts for each item), *EEEE* elaborative encoding condition in which participants were instructed to generate semantic associations for the names, *ERRR* elaborative encoding plus three retrieval practice attempts

encoding producing gains approximately equivalent to that seen for retrieval practice ($t < 1$). Of central importance for present purposes, combining retrieval practice with semantic elaboration (4th column) produced enormous gains in learning relative to the repeated study control group and substantial gains relative to each technique alone. Further, the magnitude of the gain (relative to the repeated study control) when combining the techniques (.53) was virtually equivalent to the sum of the gain produced by retrieval practice alone and semantic elaboration alone (.54).

In Experiment 2, the name learning task was modified such that photographs of faces were paired with full names (first and last); at test participants were required to try to remember the full name when cued with the photograph. Secondary school students served as participants, and for the conditions with retrieval practice, after each recall attempt the full name was provided as feedback. The results paralleled those from Experiment 1. As can be seen in Table 2 (second row), combining retrieval practice with semantic-association encoding produced substantial gains on the final recall test (relative to the repeated study control). Further, this advantage again approximated an additive effect of each individual technique alone.

A less central result is that the elaborative encoding effect in Experiment 2, though significant, was not as robust as that in Experiment 1 or as the retrieval practice effect (see Table 2); however, the instructed semantic-association technique was not designed to promote associative connections between the photograph and target name (critical for the recall task of responding with the name when given the photograph; see Pressley et al., 1982, Experiments 4 and 5, for an analogue with vocabulary learning). An elaborative technique that included enriching those connections could be expected to produce stronger elaboration effects.

Taken together, the results of these two experiments are clear cut and important. For name learning, effective elaborative encoding clearly augmented the memory benefits that retrieval practice produced on its own. Two aspects of these findings merit emphasis. First, implementing the elaborative encoding technique was straightforward and did not demand more study time than the control study condition.

Second, regarding the theoretical implications, the pattern consistently aligned with the view that elaborative encoding techniques are complementary to retrieval practice, such that each provides independent and additive benefits to learning and memory.

## Vocabulary Learning

A challenging learning task present in educational settings, as well as for individuals attempting on their own to acquire a foreign language, is learning and retaining foreign-language vocabulary meanings. This task, similar to name learning, is an ecologically valid instance of a venerable associative learning task—paired associate learning, and it is extremely tractable for laboratory experimentation. Likely for these reasons, laboratory experiments examining the effects of retrieval practice on foreign vocabulary learning are common (e.g., Carpenter et al., 2008; Carrier & Pashler, 1992; Jönsson et al., 2014; Kang & Pashler, 2014; Karpicke, 2009; Keresztes et al., 2014; Miyastu & McDaniel, 2019; Toppino & Cohen, 2009; Vestergren & Nyberg, 2014). The standard finding is that repeated retrieval practice produces significantly better memory of the meanings of foreign vocabulary items than does repeated study (notable exceptions are addressed later in this section). A more modest but equally important literature has examined whether combining elaborative encoding techniques with retrieval practice provides significant gains in learning and retention of foreign vocabulary relative to retrieval practice alone.

In a classroom experiment embedded in the regular instruction for a German-language class (for English speaking students), Fritz et al. (2007) manipulated eighth-grade students' study methods for German vocabulary. Of interest was the utility of retrieval practice, a mnemonic encoding technique (keyword method), and the combination of the two for improving learning over a more standard (control) condition in which the English words and corresponding German vocabulary word were presented along with an elaboration (to potentially add interest and improve learning). In foreign language learning, both productive and receptive aspects are important; accordingly, both were tested—the productive test required recall of the German vocabulary word when given its English meaning and the receptive test required recall of the English translation of a given German vocabulary word. Additionally, both tests were administered immediately on completion of study and were repeated after a 1-week delay. On all four tests, retrieval practice alone and keyword encoding alone produced better performance than the control. The improvement in vocabulary learning produced by each technique reflected large effect sizes (except for the keyword advantage on the delayed productive test, which approached a medium size effect). Thus, both techniques on their own were quite effective (similar results for adults were obtained in Experiments 1 and 2, using a rote rehearsal control and an own-method control, respectively; a combined condition was not included in these experiments).

This brings us to the central question: Did combining retrieval practice with the effective keyword method provide even greater gains in learning and retention relative to retrieval practice alone? Based on the reasoning that the positive effects of retrieval practice and the keyword encoding mnemonic would each be supported by different mechanisms (echoing the *complementary* view in the previous section), Fritz et al. (2007) anticipated that combining retrieval practice with the keyword encoding would produce additive effects relative to the effects of each method alone (i.e., the advantage of each method relative to the control study condition). Perhaps surprisingly, this pattern did not emerge. Relative to the control, the magnitude of the combined-condition advantage was smaller than the sum of effects of each individual method—that is, it was not additive. Moreover, for both the productive and receptive tests (immediate and delayed), the combined condition generally did not produce significant gains over the individual methods (keyword, retrieval practice). Most telling, for every test, the combined condition produced only a very slight increase in performance relative to the retrieval practice condition; the effect size consistently equaled 0.1 (Cohen's *d*), which does not reach the conventional level of even a small effect (0.2). Essentially, there was little evidence that supplementing retrieval practice with an effective encoding mnemonic improved either productive or receptive foreign-vocabulary learning (for eighth graders). This pattern is consistent with the redundancy view outlined earlier: The effective processes prompted by the keyword technique presumably overlapped with those associated with retrieval practice.

However, several aspects of the design present difficulties for clear interpretation of the results of the combined condition. First, the study conditions were manipulated within-students, with the conditions sequenced in a fixed order (control, retrieval practice, keyword, combined). This was because the researchers wanted to avoid the possibility that early exposure to the keyword condition could prompt the students to carry over the use of that method to subsequent conditions. Fritz et al. (2007) noted that students might have been fatigued when they finally encountered the combined condition, thereby penalizing performance in that condition.

A second aspect that complicates interpretation is that the retrieval practice method was restricted to practice on production; for the four retrieval practice trials of the vocabulary set, the students had to attempt recall of the German word when given the English (one-word) translation. By contrast, the keyword method is designed for receptive learning. In this method, for English speakers, a familiar English sound alike word (the keyword) is identified from the given foreign word, and then an interactive image is constructed linking the keyword to the meaning of the foreign word. For instance, for the German word PLATZ, which means *town square*, the keyword was PLATES, and one could imagine white plates littered over the town square. For a reception test, when given PLATZ, the keyword PLATES is identified and is used to cue the image of plates littered on the (town) square to directly retrieve the translation (*town square*). For a production test, the utility of the keyword is not as straightforward: The English word *town square* would presumably cue the image of the square littered with PLATES, from which the unfamiliar German word PLATZ must be retrieved or constructed. Most commonly, the keyword method is proposed to enhance performance on a receptive test but not necessarily a productive test, and the literature generally supports that assumption (Pressley & Levin, 1981; Pressley et al., 1980). (In the present experiment the keyword method did enhance productive learning relative to the control, but with only a modest effect size after a delay.)

The upshot is that each technique was arguably most directly aligned with a different type of test. Consequently, one possible interpretation of the seemingly curious patterns of the combined condition is that the mnemonic processes associated with each technique were more prominently relied upon for the aligned test. The processes for the “misaligned” technique, though effective on its own, were less recruited when combined with the more aligned technique. Specifically, when given the production test, in the combined condition students could rely heavily on mediators stabilized through retrieval practice (see Pyc & Rawson, 2010) instead of the keyword provided from the keyword method. By contrast, when given the reception test, students might have attempted to rely on the keyword embedded in the given German vocabulary item to retrieve the English translation. Complicating matters further the reception test followed the production test, perhaps somewhat contaminating the results of the reception test.

A more straightforward set of experiments focused only on receptive vocabulary learning (recall the English meaning when given the foreign vocabulary word) and manipulated study condition in between-subjects fashion (Miyatsu & McDaniel, 2019). The presence/absence of keyword mnemonic encoding was factorially combined with the presence/absence of retrieval practice, yielding four groups. Each group studied a list of 40 Lithuanian-English pairs for three rounds. In the study only group (SSS), participants simply studied the vocabulary for the three rounds without keyword encoding instructions and without retrieval practice. In the keyword group ($K_wK_wK_w$) participants studied the vocabulary pairs with suggested keywords and images for all three rounds. In the study plus retrieval practice group (SRR), participants studied the vocabulary list in round one and then attempted to retrieve the English meanings given the Lithuanian words during the second and third rounds of the learning phase. In the combined keyword plus retrieval practice condition ($K_wRR$) participants studied the list in round one using the keyword method, and then they engaged in retrieval practice during the second and the third rounds.

It is worth emphasizing that a key objective of the Miyatsu and McDaniel (2019) study was to explore the benefits of combining elaborative (keyword) encoding with retrieval practice when retrieval practice is limited (two rounds per item in this study). Studies investigating retrieval practice to enhance vocabulary learning have used relatively high dosages of retrieval practice, with positive results (e.g., Kang & Pashler, 2014: four times per item; Keresztes et al., 2014: six times per item). However, in applied settings, students are faced with high demands on their study time from their various courses (e.g., students indicate running out of time to study for a particular course because they prioritize upcoming exam(s); Susser & McCabe, 2013; Rea et al., 2022), and they indicate that a major barrier to using effective strategies like retrieval practice is lack of time (Rea et al., 2022). Accordingly, it is plausible that in educational settings, students’ use of retrieval practice for a set of vocabulary items is not necessarily extensive; rather retrieval practice for an assigned vocabulary set might be limited to a session or two. Under these circumstances, there may be substantial pay-off

**Table 3** Mean proportion meanings recalled (given the foreign vocabulary word) on the final test as a function of condition from Experiments 1-3 in Miyatsu and McDaniel (Miyatsu & McDaniel (2019). Adding the keyword mnemonic to retrieval practice: A potent combination for foreign language vocabulary learning? *Memory & Cognition*, 47, 1328-1343)

| | SSS | SRR | KKK | KRR |
|---|---|---|---|---|
| Experiment 1 | .35 | .37 | .44 | .51 |
| Experiment 2 | .22 | .22 | .32 | .51 |
| Experiment 3 | .21 | .22 | .28 | .42 |

In Experiment 1, the final test was 2 days after the study phase; in Experiments 2 and 3, the final test was 1 week after the study phase

*SSS* repeated study control condition, *SRR* retrieval practice condition (with two retrieval practice sessions), *KKK* keyword encoding condition in which participants were provided keywords in each of the three study sessions, *KRR* keyword encoding plus two retrieval practice sessions

to engage in elaborative encoding on initial study before attempting retrieval practice. Indeed, modest retrieval practice after a single typical read/study session (without elaborative encoding techniques) may do little to enhance retention of foreign vocabulary. For instance, when Kang and Pashler (2014) tested participants twice after a single initial exposure to Swahili-English pairs, test-enhanced learning was not consistently observed across three experiments (unlike a four-time testing conditions that showed reliable test-enhanced learning).

Two major sets of results (from Miyatsu & McDaniel, 2019) supported the possibilities just outlined. First, across three experiments, two rounds of retrieval practice failed to significantly enhance recall of the meanings of the Lithuanian words. As the first and second columns of Table 3 show, recall in the repeated study (SSS) and the corresponding retrieval practice group (SRR) was essentially equivalent on the final test administered 2 days (Experiment 1) and 1 week (Experiments 2 and 3) after the learning phase. To maximize power to detect a retrieval practice effect, final test performance for these groups was combined across experiments (160 total participants), and still there was no retrieval practice effect ($SSS_M = .26$; $SRR_M = .27$; $d = .06$). Experiment 3 also included conditions with four rounds of retrieval practice (SRRRR) or four additional rounds of study (SSSSS). With more extensive rounds of retrieval practice, converging with studies previously mentioned, an advantage of retrieval practice began to emerge ($SSSSS_M = .27$; $SRRRR_M = .37$; $d = .44$).

Second, preceding retrieval practice with elaborative (keyword) encoding of the vocabulary list produced retrieval practice effects. That is, as shown in the third and fourth columns of Table 3, retrieval practice improved performance when it followed the effective keyword encoding (relative to keyword encoding alone). Also, as anticipated by the catalytic view, a planned interaction test indicated that the testing effect in the combined testing-keyword condition was significantly more robust than the testing effect without the keyword method (in Experiments 2 and 3 in which final performance was tested 1 week after the learning sessions). Overall, these patterns align with view that effective elaborative encoding can catalyze the mnemonic benefits of retrieval practice (within the boundary conditions specified earlier in this article when outlining the catalytic view). The idea is that elaborative encoding (the keyword method in this case) can support relatively high levels of initial retrieval, with initial retrieval promoting subsequent retention (positive retrieval practice effects). Without elaborative encoding, initial retrievals (e.g., the two rounds) may be relatively low, thereby undermining benefits of retrieval practice relative to a restudy condition that re-exposes all the items (only the few items retrieved gain the mnemonic benefit of retrieval; Kornell et al., 2011).

The recall results from the retrieval practice trials reinforced this account. Specifically, in the second retrieval round after study alone (without keyword-mnemonic instruction), participants could recall only about one-third of the meanings of the vocabulary words (consistently across the 3 experiments). By contrast, the second round of retrieval practice after effective elaborative study (keyword-mnemonic) produced correct recall for about half of the vocabulary items. It is important to emphasize that these results were obtained even though correct answer feedback was provided in both retrieval practice rounds. Thus, even when the SRR group had the opportunity to learn correct answers from feedback, learning from feedback in the initial round of retrieval practice still did not allow learners to "catch-up" to the boost in retrieval provided by initial elaborative encoding.

In terms of applied importance, the combination of effective elaborative (keyword) encoding and limited retrieval practice for learning foreign vocabulary meanings was extremely potent. Referring again to Table 3, with no more total time spent in the acquisition phase (relative to the total study time for the SSS group) the combined group's final recall was substantially greater than that supported by restudy alone. Indeed, recall was doubled (Experiment 3) or more than doubled (Experiment 2) by combining elaborative encoding with retrieval practice (relative to restudy alone). Further, combining the two techniques added recall gains relative to using either technique alone.

Another index of the potency of preceding retrieval practice with effective encoding is the savings in total time spent during acquisition in the combined condition relative to using more extensive retrieval practice alone. In Experiment 3, some groups received four total retrieval practice trials after being exposed to the vocabulary list. Yet, this extended retrieval practice group (SRRRR) group did not achieve the level of final recall performance (M = .37) that was reached by the group using keyword encoding and only two retrieval practice trials (M = .42). (Again, this result emerged even though the extended retrieval practice was accompanied by correct-answer feedback in every retrieval practice round.) For students faced with heavy demands on their study time, reducing additional retrieval practice repetitions (needed to achieve a particular performance level) by initially employing effective elaborative encoding would seem highly attractive.

I now turn to an examination of whether elaborative encoding benefits retrieval practice for vocabulary learning in a quite different acquisition situation. When implementing retrieval practice for learning meanings of new vocabulary, rather than set a fixed amount of time for acquisition and rather than study always preceding retrieval practice, another approach could be to persist in alternating periods of study and retrieval practice until all meanings have been successfully recalled at least once. Within this kind of learning paradigm, Karpicke and Smith (2012) examined learning of 30 unfamiliar English vocabulary items (e.g., *antiar*, which means *poison*). Of interest for present purposes, one manipulation was whether participants received no instructions on how to study or were instructed with the keyword elaborative encoding technique. Because the paradigm was somewhat complex, below I outline only the features most pertinent to the present issue.

All participants received multiple study and test periods. Each study period preceded each test period, and each period included several study trials and several test trials, respectively. After the first study and first test period, there were variations in how the following study and test periods were configured. The key condition I discuss here is the repeated retrieval condition. In this condition, once an item's meaning had been successfully recalled, it was dropped from the study list but was retained in the list for two more retrieval periods (each of which had several retrieval trials). (In other conditions, once an item was successfully recalled it was dropped from the following retrieval periods.) The elaborative (keyword) study manipulation was implemented differently across two experiments. In Experiment 1, after the first successful recall of an item, the keyword mnemonic was presented (for that item), and participants were instructed to use that mnemonic for all further study/test trials. In Experiment 2, the keyword mnemonic was given on every study trial (but remember that in the repeated retrieval condition, after study period 1, an item was dropped from study if previously successfully recalled).

One week after the acquisition session, participants were tested for their memory of the meaning of each vocabulary item when cued with the item. The result of interest is whether in the repeated retrieval condition, elaborative encoding enhanced the performance relative to no (provided) elaborative encoding. In both experiments, the elaborative encoding group recalled numerically more meanings than the no-elaborative encoding group (Figs. 2 and 4, last set of bars, presented in Karpicke & Smith, 2012), but these differences were not statistically significant. This paradigm, however, may have obviated the need for elaborative encoding because study (and retrieval periods) continued until an item was successfully recalled. That is, even in the no-elaborative encoding condition, study was repeated until the item could be recalled, and then repeated retrieval practice continued for that item. Moreover, retrieval practice was relatively extensive: Retrieval practice (there was a minimum of a first retrieval period with at least several retrieval trials) always occurred for an item along with study, and this was repeated, until the item was recalled once; then retrieval practice for recalled items continued to be repeated in subsequent retrieval periods.

Accordingly, the implications for application (which was not the purpose of the study) may be limited because such extensive retrieval practice would be unusual in educational settings (based on literature in which the testing effect has been applied in classrooms; see reviews by Agarwal et al., 2021; Trumbo et al., 2021). Yet, even when retrieval practice is implemented within a learning-to-criterion paradigm, there still may be a reliable, albeit small, boost to retrieval practice from elaborative encoding (consistent across Experiments 1 and 2 in Karpicke & Smith, 2012). Unfortunately, the statistical power in Karpicke and Smith (2012) to detect such an effect was not adequate (16 participants in each condition, providing negligible power [.28] to detect a medium-size effect and low power [.59] to detect even a large-size effect for the comparison). More studies are clearly warranted concerning the issue of whether elaborative encoding augments retrieval practice benefits produced by extensive retrieval practice.

## Learning Scientists' Names and Contributions

Another common and challenging task in educational settings is to learn to pair proper names with facts, events, and accomplishments (Jones & Hall, 1982). Elaborative encoding techniques such as the keyword method have been shown to significantly enhance performance on this type of task (Jones & Hall in a classroom setting; Shriberg et al., 1982). Regarding retrieval practice, based on experiments with like materials (learning to pair items with arbitrary associations, such as vocabulary learning) almost certainly sufficient retrieval practice would also enhance learning of proper name—fact/accomplishment pairings. Would combining the two enhance learning over each alone?

Cummings et al. (2023) addressed this key question for the task of learning the contributions of scientists in a particular discipline. They assembled materials from a university cultural anthropology course in which the summative assessment included having to recall the key contribution of each of the anthropologists covered in the course. In the experiment, each anthropologist's name (e.g., Marvin Harris) was accompanied by a brief explanation of their contribution to the field of anthropology (e.g., "The perspective that the physical world affects and puts constraints on human behavior"). Four learning conditions were included based on the factorial combination of the presence/absence of keyword encoding and the presence/absence of retrieval practice. Following Miyatsu and McDaniel (2019), the SSS (no keyword, no retrieval practice) control condition involved three study sessions in which each anthropologist and their contribution were presented. In the KwKwKw (keyword, no retrieval practice) condition, in each study session participants were presented with the anthropologist, their contribution and a keyword link between the two (e.g., "Think of a heavy hair blanket (HARris) covering a person and constraining them"). In the SRR condition after the first study session, participants were given two sessions of retrieval practice (correct answer feedback included). In the KwRR condition, the first study session included the keyword link followed by two retrieval practice sessions. The final test, in which participants attempted recall of the contribution when cued with the anthropologist, was administered 2 days after the learning session.

In Experiment 1, with a modest list length (18 items), retrieval practice (SRR) produced significantly better retention than did repeated study (SSS). In Experiment 2, with a more challenging list length (30 items), there was a nonsignificant advantage of retrieval practice relative to repeated study. To contextualize these effects, the retrieval-practice effects were medium size in both experiments; by contrast, the significant advantages of the elaborative encoding (keyword) technique (KwKwKw vs. SSS) for final recall were large-size effects in both experiments (Table 4 provides means). Most importantly, combining retrieval practice

**Table 4** Mean proportion recalled of anthropologists' contributions (given the anthropologist's name) on the final test as a function of condition from Experiments 1 and 2 in Cummings et al. (Cummings et al. (2023). Do not forget the keyword method: Learning educational content with arbitrary associations. *Journal of Applied Research in Memory and Cognition, 12, 70-81*)

| | SSS | SRR | KKK | KRR |
|---|---|---|---|---|
| Experiment 1 | .28 | .41 | .46 | .55 |
| Experiment 2 | .24 | .38 | .43 | .59 |

*SSS* repeated study control condition, *SRR* retrieval practice condition (with two retrieval practice sessions), *KKK* keyword encoding condition in which participants were provided keywords in each of the three study sessions, *KRR* keyword encoding plus two retrieval practice sessions

with effective elaborative encoding (KwRR) produced substantial learning benefits. Relative to study alone the combined condition increased final test performance by nearly 90% in Experiment 1 and more than 145% in Experiment 2. Further, the combined condition produced better performance than either technique alone. This pattern is consistent with the complementary view outlined earlier: Effective encoding strategies and retrieval practice serve complementary roles for learning and retention, thereby producing approximately additive effects.

## Learning from Connected Discourse

Of the learning tasks considered herein, by far the most ubiquitous for educational settings is learning from connected discourse—students are routinely required to learn content delivered in lectures and assigned readings. Experimental work has established that retrieval practice can promote learning and retention for content from expository text (e.g., Glover, 1989; Kang et al., 2007; Roediger & Karpicke, 2006b). Indeed, retrieval practice is a keystone of the effective read-recite-review study strategy for text learning (Martin et al., 2016; McDaniel et al., 2009). The central question here is whether combining elaborative learning strategies with retrieval practice augments these positive effects of retrieval practice. Four studies with connected discourse have begun to address this question, but with a paradigm that diverges from that implemented in the experiments reviewed thus far focusing on associative learning tasks. For the associative learning tasks, elaborative encoding was introduced prior to retrieval practice. By contrast, the paradigms with connected discourse implement an elaborative strategy *concurrently* with retrieval practice.

In Endres et al. (2017), participants viewed a 30-min lecture for the purpose of learning its contents for a posttest. Immediately after the lecture, the retrieval practice conditions were manipulated. In one condition, participants attempted to recall the contents of the lecture (standard free recall). In the other condition, termed here *elaborative free recall*, participants were additionally instructed to "refer to examples from your own life, which illustrate the learning material, are consistent with it, or stand in conflict with it" (Endres et al., p. 15). This elaborative prompt was designed to stimulate learners to draw connections among concepts and relate those concepts to prior knowledge, elaborative processes that support learning and comprehension. As indicated by the recall protocols, the elaborative free recall condition did engage in substantially more elaboration than did the standard recall condition; the standard recall condition showed very little elaboration. The conditions were equivalent, however, in the amount of lecture content recalled.

One week later, participants completed a final test composed of fact questions (e.g., "What is the central executive?") and of comprehension questions (e.g., "Describe the three kinds of cognitive load in their relation to each other. Provide your own example which illustrates the load types in relation to each other."). Elaborative free recall did not produce greater fact learning than did standard free recall (standard retrieval practice). Critically, however, engaging in elaboration when attempting recall significantly improved performance on the comprehension test ($M = 7.13$ points) relative to simply attempting recall ($M = 5.56$ points). These results indicate that retrieval-practice effects with complex materials can be enhanced when recall is enriched with an elaboration strategy.

Roelle and Nückles (2019) conducted a similar study on learning from expository text using an expanded design that factorially manipulated the presence/absence of the elaborative strategy and the presence/absence of retrieval practice. In the initial study phase, participants first read a didactic text on academic self-concept, with the instruction to prepare themselves for subsequent questions about the content. Then, participants in the no elaboration/no retrieval practice group were instructed to reread the text for 30 more minutes. In the no elaboration/retrieval practice group, instead of rereading for 30 min, participants were given 30 min to free recall everything they remembered from the text. In the elaboration/no retrieval practice group, with the text in view, participants spent 30 min elaborating by responding to prompts to highlight the main content items and their connections and to illustrate the main content by giving their own examples. In the combined elaboration/retrieval practice group, participants responded to the elaboration prompts for 30 min without the text in view (which required retrieval). After a 1-week delay, participants completed a post-test with three sets of questions. The first set required reproduction of the text's content (open-ended questions like "What is the big-fish-little-pond effect and what does basking-in-reflected-glory mean?"); the second set assessed comprehension with open ended questions that required explanation of concrete scenarios; and the third set was multiple-choice questions testing recognition of statements from the text.

In one experiment, the text was highly cohesive and elaborated[2], a text that the authors anticipated would produce retrieval practice effects on text learning and retention. Confirming this expectation, as can be seen in Table 5 the retrieval practice (free recall) alone group performed better than the reread group on the comprehension

[2] Using accepted guidelines for preparing texts with high cohesion, the text (on academic self-concept and achievement) was written "such that (a) within each of the four subtopics only one main line of argumentation was followed at a time (i.e., there was no jumping back and forth between different lines of argumentation), and that (b) interrelations between the four subtopics were explicated at the end and/or at the beginning of the respective paragraphs" (Roelle & Nückles, 2019, p. 1345). To ensure high elaboration, main ideas were enriched by illustrative examples.

**Table 5** Mean proportion correct for the comprehension questions and reproduction questions as a function of condition from Experiment 1 in Roelle and Nückles (Roelle & Nückles (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)

| | Reread | Read-retrieve | Elaborate | Elaborate-retrieve |
|---|---|---|---|---|
| Comprehension | .41 | .51 | .45 | .55 |
| Reproduction | .40 | .43 | .41 | .49 |

Comprehension was assessed with open ended questions that required explanation of concrete scenarios. Reproduction of the text's content was assessed with open-ended questions (e.g., "What is the big-fish-little-pond effect and what does basking-in-reflected-glory mean?"). These questions were administered 1 week after the learning session

questions (compare 2nd vs. 1st columns). Further, there was only a slight and non-significant advantage for the elaboration only group relative to reread (3rd vs. 1st column). Most importantly, diverging from Endres et al. (2017; with a lecture), requiring elaboration during retrieval did not significantly improve comprehension performance relative to retrieval practice alone. It is worth emphasizing that for this experiment the elaborative activity was not effective in improving final test performance (relative to reread), and consequently would not be seen as a desirable elaborative strategy for augmenting retrieval practice effects for these materials. However, from a theoretical perspective, the pattern is quite interesting in as much as the comprehension benefit for the combined group (.14 relative to reread) is exactly the sum of the nominal benefit of elaboration alone (.04) and retrieval alone (.10). These additive effects are completely in line with the idea that elaborative strategies and retrieval strategies provide complementary benefits for learning and retention.

For the reproduction questions, there were no significant effects (see row 2 of Table 5); thus, retrieval practice did not significantly increase recall (relative to reread or in general as a main effect). Still, the combined group produced the best performance relative to the other three groups, perhaps hinting that engaging in elaboration during retrieval practice provides unique advantages to either alone. Performance on the multiple-choice questions was high and varied little across the groups (ranging from .81 to .85).

Blunt and Karpicke (2014) reported yet another pattern in an experiment that examined the effects of combining an organizational elaborative task—concept mapping—with retrieval practice. In a concept-mapping activity, students generate diagrams composed of nodes that are linked together; the nodes represent concepts, and the links represent relations among the concepts, with the nature of the relationship typically represented by labeling the links. In a critical experiment (Experiment 2), the presence/absence of the elaborative concept mapping activity was factorially combined with the presence/absence of retrieval practice. The learning session intermixed repeated reading with the assigned learning activity; participants first read a relatively short science text (about 280 words), then performed the learning activity, reread the text, and then performed the learning activity again. For the control (no concept mapping; no retrieval practice), the learning activity involved copying the paragraphs of the text. For the concept mapping condition without retrieval practice, participants generated a concept map while viewing the text. For the retrieval practice condition without concept mapping, participants had to recall as much of the text as they could remember. For the combined condition (concept mapping + retrieval practice), participants had to generate a concept map without having access to the text (i.e., while retrieving). One week later participants completed a short-answer test that included factual (verbatim) questions ("what do proteins lose at high temperature?") and inference questions ("What happens to catalytic activity if temperature decreases?").

For factual questions, concept mapping was effective; when completed while viewing the text (no retrieval) it produced learning and retention gains relative to the control (M's = .43. vs. .33). Retrieval practice alone was also effective (M = .49). However, engaging in concept mapping while retrieving (M = .46) had no additional benefit over retrieval practice alone and only a slight benefit if any over concept mapping alone. Performance on inference questions could not inform the issue of whether effective elaboration boosts retrieval-practice effects, because concept mapping without retrieval (M = .32) did not benefit performance relative to the no concept-mapping control (M = .30) (somewhat surprisingly as noted by the authors).

The pattern for the factual questions converges on the conclusion that the organizational or relational processing presumed to be promoted by concept mapping "may be redundant with the processing people already engage in when practicing retrieval" (Blunt & Karpicke, 2014, p. 857). This interpretation is clouded by the following consideration, though. The benefits of requiring elaboration when performed concurrently with retrieval could be depressed because requiring retrieval to perform the elaborative activity can reduce the amount of content from the text that is elaborated. Indeed, when engaging in concept mapping alone (in the presence of the text) participants on average included over half of the idea units from the text (53%); whereas when concept mapping was performed with retrieval, fewer than a third of the idea units were included in the maps (32%) (averaged across the two learning-activity sessions).[3] A study that required concept mapping during study and followed by retrieval practice would be informative for evaluating the redundancy interpretation.

Larsen et al. (2013) also factorially manipulated the presence/absence of an elaborative strategy (self-explanations) and the presence/absence of retrieval practice in an experiment in which 1st-year medical students learned clinical neurology topics from a classroom teaching session (using an interactive didactic format). The elaboration and retrieval practice manipulations were within-subjects, such that each student experienced each of the four conditions (each condition was assigned to one of four topics taught). Following the initial lesson (in which all four topics were studied), students completed each of the four conditions (repeated study, retrieval,

[3] The paradigm also blended direct benefits of successful retrieval with indirect benefits of potentiating further learning on subsequent study (the first retrieval session was followed by a second reading of the text) (Arnold & McDermott, 2013; Little & McDaniel, 2015). Accordingly, it is possible that the processing learners engaged in while restudying after a retrieval attempt (the processing associated with test-potentiated learning) is the processing that was primarily redundant with organizational or relational processing presumed to be promoted by concept mapping

self-explanation, retrieval + self-explanation), with a particular condition assigned to a particular topic (counterbalanced across students). Subsequent sessions of study or retrieval practice were repeated four times, spaced over a 3-week period after the lesson was taught. For the repeated study condition, students studied (4 times) a review sheet of 26-30 items included in the lesson for the topic. For the retrieval practice alone condition, students had to complete a short answer written test (4 times), and then they scored their responses from an answer key. For the elaboration condition, students were provided with the review sheet and instructed to generate explanations for why each piece of information was important and how it related to existing knowledge. The retrieval plus elaboration condition required students to respond to the short-answer questions and generate explanations for the information they recalled (and then scored their responses).

Another novel feature of this experiment was that final testing was administered after a long (6-month) interval. The test provided a short clinical scenario for a patient (demographic information and the chief complaint). Students were instructed to write an essay outlining their approach to managing the patient, with the essay to include all the information provided in the initial teaching session (and further practiced). One notable finding was that both repeated self-explanation (M = .29) and repeated retrieval practice (M = .36) produced significantly more information correctly recalled than did repeated study (M = .20). This finding reinforces the benefits of effective elaboration (in this case self-explanation) and of retrieval practice for very long term retention (and application) of course material.

For present purposes, the central finding was that combining retrieval practice with self-explanation (M = .40) nominally improved performance relative to retrieval practice alone, but the difference was not statistically significant. Moreover, the combined effects of self-explanation and retrieval relative to repeated study (an increase of .20) was somewhat less than the sum of the individual effect of self-explanation (.09) and individual effect of retrieval practice (.16). However, it must be noted that self-explanations were successfully completed for 96% of the key lecture information in the self-explanation alone condition (the review sheet from which the explanations were generated included all the key pieces of information), whereas for the retrieval plus self-explanation condition, self-explanations were completed for 71% of the information recalled; and only about 75% of the information was recalled. Thus, the upshot is that students generated self-explanations for much less of the lecture when self-explanations were attempted during retrieval than when attempted without retrieval (review sheets present) (similar to the Blunt & Karpicke, 2014, findings). Nevertheless, the authors concluded that "educators should seek to incorporate both of these powerful learning techniques [repeated retrieval and self-explanation] into their curricula in order to enhance long-term retention and the application of learned material" (p. 682).

The results from the limited literature taken in concert (one study with learning from a lecture, one with a classroom lesson, and two with expository texts) mildly suggest that engaging in elaboration *during* retrieval practice might augment the benefits of retrieval practice (except Blunt & Karpicke, 2014). To the extent that requiring retrieval while elaborating can impair the quantity and quality of elaboration because of the hurdle of having to retrieve the idea units for which elaboration is engaged (Larsen et al., 2013; Roelle & Nückles, 2019), a more fruitful and successful approach to combining elaboration with retrieval practice could be to implement elaboration and retrieval practice successively (as was done in all of the studies that investigated combining elaborative encoding and retrieval practice for associative learning tasks). Specifically, elaborative tasks would be implemented with the initial study/reading session. A few of the many examples in the text-learning literature include inserting embedded questions to prompt self-explanation (McDaniel & Donnelly, 1996; Seifert, 1993), training self-explanation reading strategies (McNamara, 2004), instructing readers to pose and answer deep questions (Rosenshine et al., 1996), and instructing students to create outlines when reading (Einstein et al., 1990) or to enhance notetaking (Bui & McDaniel, 2015). Retrieval practice would then follow a study/reading session in which an elaborative technique was explicitly prompted or supported. To the extent that these initial elaborative techniques promote text learning (which they generally do), then more idea units should be retrieved during retrieval practice, thereby augmenting the benefits of retrieval practice on final retention tests (see Roelle et al., 2022). This avenue seems like an especially promising direction for further research with connected discourse on the combined effects of elaborative encoding and retrieval practice.

## Discussion

An emerging literature has begun to examine the effects of combining potent elaborative encoding methods with retrieval practice. Inspection of Table 1 indicates that one factor that appears to be important in the pattern of these effects is whether the elaborative encoding is engaged prior to retrieval practice or concurrent with retrieval practice. I first discuss the results when learners engage in effective elaborative encoding during study followed by retrieval practice. In this case, the results are extremely promising: Final test performance is enhanced over retrieval practice alone (the Fritz et al., 2007, and Karpicke & Smith, 2012, exceptions are discussed in the next section). This pattern has important educational implications for the use of retrieval practice to promote learning and retention. For learning tasks that require the acquisition of associations that are relatively arbitrary (i.e., do not have clear semantic connections)—in education these include learning of foreign vocabulary meanings and learning the contributions associated with individual scientists—providing learners with an effective elaborative encoding strategy to use during initial study substantially increases the benefit of retrieval practice alone (Cummings et al., 2023, Experiments 1–2; Miyatsu & McDaniel, 2019, Experiments 1–3; Morris et al., 2005, Experiments 1–2).

These results are especially noteworthy because combining elaborative encoding with retrieval practice did not penalize efficiency (study time). Learners were given no more time to implement the instructed elaborative-study strategy for initial study relative to learners that were not instructed to use an elaborative strategy. In fact, one experiment demonstrated that augmenting retrieval practice with elaborative encoding increased the efficiency of obtaining particular levels of performance relative to that required from retrieval practice alone. Specifically, learners who engaged an elaborative study technique (keyword method) followed by only 2 sessions of retrieval practice recalled more foreign-vocabulary meanings than did learners not instructed in an elaborative study technique who completed 4 sessions of retrieval practice (.42 vs. .37, respectively; Miyatsu & McDaniel, 2019, Experiment 3). Clearly, for this kind of learning task (given the experimental parameters of list length, study time per item, etc.), retrieval practice alone does not provide optimal efficiency. Students challenged with carving out study time for the substantial amount of material to-be-learned in their courses would seem to be better served by preceding retrieval practice with effective elaborative encoding, rather than relying on retrieval practice alone.

Perhaps somewhat surprisingly, the positive effects just summarized when effective elaboration precedes retrieval practice are observed even though correct answer feedback was provided on retrieval practice trials. Researchers have noted that direct effects of retrieval practice (no feedback given)—benefits of initial retrieval on retention for to-be-learned material relative to restudy—occur primarily when initial retrieval is relatively successful (estimates range from 50%-75% success; Karpicke, 2017; Roelle et al., 2022; Rowland, 2014). Accordingly, when initial retrieval can be expected to be somewhat low (e.g., after studying lists of items with arbitrary associations like foreign vocabulary or scientists' contributions; reading low-cohesion texts, Roelle & Nückles, 2019), correct-answer feedback may be used to overcome limited direct effects of retrieval. Learning is facilitated from processing feedback after failed retrieval attempts (Butler & Roediger, 2008; McDaniel & Fischer, 1991; Pashler et al., 2005), one of the potent indirect effects of retrieval (Roediger et al., 2011). Indeed, a strong argument for the effectiveness of providing correct-answer feedback is the theoretical claim that the *test-memory* responsible for the testing effect is equally robust regardless of whether the test response is retrieved from memory or provided through feedback (Gupta et al., 2022).

Thus, it might be expected that providing feedback would obviate possible dampened testing effects when initial retrieval success is not overly high. That was not the case, however, in the experiments with learning meanings of foreign language vocabulary materials and learning scientists' contributions. In these cases, the low initial retrieval success ranging from 15% to 22% (across four independent groups of participants after a study session in which participants were not instructed to use elaborative encoding; Miyatsu & McDaniel, 2019), even with feedback on the initial and second retrieval attempts, did not produce a testing effect (see also Cummings et al., 2023, Experiment 2). Effective elaborative encoding, however, increased initial retrieval (retrieval practice levels) by about 15–20% (for the foreign vocabulary and scientists' contributions materials), and this advantage remained stable even across 4 retrieval attempts with feedback (Miyatsu & McDaniel, 2019, Experiment 3).

The importance of these patterns is that they highlight the apparently unique and efficient advantage of incorporating effective elaborative encoding to support initial successful retrievals (for augmenting testing effects), rather than relying solely on corrective feedback across retrieval attempts to produce higher levels of retrieval (and higher final test performance). This pattern reinforces the complementary view that construction of rich mental representations, achieved through effective elaborative and generative study processes, are important components of learning that provide somewhat distinct benefits over that of retrieval practice alone (Fritz et al., 2007; Miyatsu & McDaniel, 2019; Morris et al., 2005; Roelle et al., 2022; see also Karpicke & Smith, 2012). Therefore, combining the two techniques should have great utility in educational practice. One limitation of the existing studies, however, is that final performance levels were below what is typically required for passing courses and demonstrating competence in many educational settings (with the possible exception of Karpicke & Smith, 2012) (Rawson & Dunlosky, 2022, have noted this limitation for many memory studies in general). Research along these lines with real-world learning objectives and in applied settings would be extremely valuable in extending the current literature, which has been limited primarily to relatively low performance levels in laboratory experiments.

Sharply contrasting results (see Table 1) have generally been obtained when learners are prompted to elaborate when attempting to retrieve the contents of the to-be-learned material (Blunt & Karpicke, 2014; Larsen et al., 2013; Roelle & Nückles, 2019). This paradigm is a significant departure from the proposition that elaborative encoding engaged during study could augment benefits of retrieval practice conducted after study. With elaboration performed concurrently with retrieval, the demands of retrieval (in terms of less content being retrieved for elaboration [e.g., Blunt & Karpicke, 2014; Larsen et al., 2013] or considerable amount of time/effort required to retrieve content for elaboration) apparently attenuate the benefits of elaboration. For instance, in an experiment with a low-cohesion text that was also relatively unelaborated (thereby challenging retrieval), the expected benefits of elaboration alone (when performed with the text present) were virtually eliminated when elaboration was required during retrieval (Roelle & Nückles, 2019, Experiment 2). Similarly, observed benefits of an elaborative concept-mapping study task did not emerge when concept mapping was required during retrieval (Blunt & Karpicke, 2014, Experiment 2). In this case, as well as Larsen et al. (2013), many fewer idea units were elaborated when the elaborative activity (concept mapping) had to be completed while retrieving the content than when the elaborative activity was performed with the content in view. It seems quite plausible that if effective elaboration were engaged during reading/studying of didactic texts and other connected discourse (preceding retrieval practice), then retrieval practice effects for retention of the text's or lesson's contents could be enhanced (over retrieval practice following reading alone).

This possibility would be fruitful to examine in further research, not only for practical reasons but to better understand the possible boundary conditions of when elaborative activity will augment retrieval practice effects. Table 1 also shows that the studies implementing elaboration concurrently with retrieval exclusively examined connected discourse (lectures, didactic texts), whereas the studies implementing elaboration during study and followed by retrieval examined more impoverished materials (e.g., paired-associate lists). Consequently, it is possible in principle that learning from connected discourse constrains the benefit of combining elaboration with retrieval practice. That said, elaborative encoding (performed at retrieval) for a lecture did significantly enhance comprehension performance (after a week delay) relative to retrieval alone (Endres et al., 2017).

## Other Potential Limitations of Combining Elaborative Encoding with Retrieval Practice

Another possible limitation to enhancing retrieval practice effects through effective elaboration at study would be situations in which initial retrieval of studied material is relatively high. This could occur in a number of learning paradigms, some of which are not uncommon. The presence of extensive study sessions is one example. Extensive study can promote high levels of initial retrieval, thereby potentially mitigating the advantage of study with elaboration (because elaborative study helps promote higher levels of initial retrieval with few study trials). For instance, in Toppino and Cohen (2009) after 8 study repetitions of a list of foreign vocabulary-English translation pairs, performance on the retrieval practice trial approached 90% correct; and in Roediger and Karpicke (2006b) extensive study of a short text (about 260 words; studied in a 5–7-min study session, depending on the experiment) produced levels of recall at about 70% during initial retrieval practice. With such extensive study, incorporating elaboration would not necessarily improve the already high levels of initial retrieval (and consequently not enhance retrieval practice). Another instance where initial retrieval might be high is after study of a limited amount of material; if so, adding elaboration would not be necessary to improve initial retrieval. For example, in Fritz et al. (2007) in one experiment participants studied six-item lists, and initial retrieval practice levels could have been high for such lists (unfortunately, the initial retrieval-practice rates were not reported; however, elaborative encoding did not enhance the retrieval practice effects). A final instance worth mention is a learning context that includes multiple study-retrieval sessions. In Karpicke and Smith (2012) sessions involving several study repetitions and several retrieval practice repetitions were repeated such that all vocabulary items were retrieved at least once. With high retrieval levels ensured during learning, elaborative encoding (keyword method for new vocabulary) did not significantly augment retrieval practice effects.

Though this possible boundary condition is worth noting for a comprehensive consideration of the effects of combining elaborative/enriched initial encoding with retrieval practice, the applied educational implications remain: Extensive study, retrieval practice, or both might obviate the need for initial elaborative encoding, but its inefficiency is likely prohibitive for students in terms of the additional time and effort required (e.g., Maurer & Shipp, 2021). Further, the task cost to the student of such extensive study/retrieval practice could be expected to negatively impact motivation (e.g., Kim et al., 2022), further reducing its desirability. It merits emphasizing that the significant benefits of elaborative encoding for enhancing retrieval practice (after one study session) can persist even after fairly extensive retrieval practice (4 retrieval practice trials in Miyatsu & McDaniel, 2019, Experiment 3); moreover, combining elaborative (keyword) encoding with retrieval practice can dampen or eliminate the forgetting (comparing across experiments) after a 7-day delay (rows 2 and 3 in Table 3) relative to a 2-day delay (row 1) observed for repeated study, retrieval practice alone, and keyword encoding conditions.

## Theoretical Recapitulation

The reviewed findings largely contradict the view that effective elaborative encoding is not a useful supplement to retrieval practice, especially when elaborative encoding is engaged prior to retrieval practice. Exceptions are the experiments by Fritz et al. (2007), for which elaborative encoding did not enhance retrieval practice effects for middle school students, and Blunt and Karpicke (2014) who implemented elaborative encoding concurrently with retrieval (also Karpicke & Smith, 2012, but see previous discussion concerning nominal combined elaboration-retrieval effects with low power to detect significance that were nevertheless reliable across two experiments). However, as detailed earlier the interpretation of these results is clouded by several factors. In sum, there is little support for the view that elaborative encoding is redundant with retrieval practice and thus not necessary when retrieval practice techniques are implemented.

Instead, the preponderance of the evidence strongly suggests that effective elaborative encoding provides mnemonic benefits that are complementary to those produced by retrieval practice. As noted throughout the review, the clearest evidence of complementary effects is that the magnitude of the increased retention resulting from the combination of elaborative encoding with retrieval practice (relative to a study-only condition without instructions to elaborate) often reflects an approximately additive effect of the retention gains produced by each technique alone. This additive effect is not so prominent when elaborative processing is performed during retrieval of connected discourse, though even in this case—where elaboration necessarily is restricted to retrieved contents (as opposed to the elaboration-only condition, in which all the target content is available for elaboration)—it is interesting that additivity can be seen numerically (Roelle & Nückles, 2019, Experiment 1). It may be worth mention that the nature of the complementary processes activated by elaborative encoding on the one hand and retrieval practice on the other hand have been described by different researchers along various theoretical dimensions (e.g., Fritz et al., 2007; Karpicke & Smith, 2012; Miyatsu & McDaniel, 2019; Roelle et al., 2022).

Another notable pattern emerges for learning tasks in which limited retrieval practice does not significantly enhance retention (relative to study alone) (e.g., Miyatsu & McDaniel, 2019). In this case, effective elaborative encoding catalyzes the effectiveness of retrieval practice (for the set of items as a whole), by boosting initial retrieval levels such that enough material is retrieved to produce testing effects (retrieval practice effects) for the set of items (Kornell et al., 2011). Here, the combined effect of elaborative encoding with retrieval practice is in a sense super-additive, because the effects of retrieval-practice alone are negligible for the target content but become manifest when elaborative encoding is included. Clearly, work remains to understand the dynamics and consequences of incorporating elaborative encoding techniques with retrieval practice for a range of learning materials and learning contexts. At this juncture, however, the evidence is that incorporating elaborative encoding techniques with retrieval practice (prior to retrieval practice) generally enhances learning and retention over retrieval practice alone.

**Acknowledgements** I am grateful to Julie Bugg and Gilles Einstein for comments on an earlier version of this paper.

## References

Adesope, O. O., Trevisan, D. A., & Sundararajan, N. (2017). Rethinking the use of tests: A meta-analysis of practice testing. *Review of Educational Research, 87*(3), 659–701.

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*, 861–876.

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.

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*, 940–945.

Bellezza, F. S. (1987). Mnemonic devices and memory schemas. In M. A. McDaniel & M. Pressley (Eds.), *Imagery and related mnemonic processes* (pp. 34–55). Springer-Verlag.

Bjork, R. A. (1975). Retrieval as a memory modifier: An interpretation of negative recency and related phenomena. In R. L. Solso (Ed.), *Information processing and cognition. The Loyola symposium* (pp. 123–144). Erlbaum.

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

Bui, D., & McDaniel, M. A. (2015). Enhancing learning during lecture note taking using outlines and illustrative diagrams. *Journal of Applied Research in Memory and Cognition, 4*, 129–135.

Butler, A. C., & Roediger, H. L. (2008). Feedback enhances the positive effects and reduces the negative effects of multiple-choice testing. *Memory & Cognition, 36*, 604–616.

Carpenter, S. K. (2009). Cue strength as a moderator of the testing effect: The benefits of elaborative retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35*, 1563–1569.

Carpenter, S. K. (2011). Semantic information activated during retrieval contributes to later retention: Support for the mediator effectiveness hypothesis of the testing effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37*, 1547–1552.

Carpenter, S. K., Pashler, H., Wixted, J. T., & Vul, E. (2008). The effects of tests on learning and forgetting. *Memory & Cognition, 36*(2), 438–448.

Carrier, M., & Pashler, H. (1992). The influence of retrieval on retention. *Memory & Cognition, 20*(6), 633–642.

Clark, J. M., & Paivio, A. (1991). Dual coding theory and education. *Educational Psychology Review, 3*(3), 149–210.

Cohen, G., & Burke, D. M. (Eds.). (1993). *Memory for proper names*. Erlbaum.

Cummings, E. L., Reeb, A., & McDaniel, M. A. (2023). Do not forget the keyword method: Learning educational content with arbitrary associations. *Journal of Applied Research in Memory and Cognition, 12*, 70–81.

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.

Einstein, G. O., McDaniel, M. A., Owen, P. D., & Cote, N. C. (1990). Encoding and recall of texts: The importance of material appropriate processing. *Journal of Memory and Language, 29*, 566–581.

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.

Fritz, C. O., Morris, P. E., Acton, M., Voelkel, A. R., & Etkind, R. (2007). Comparing and combining retrieval practice and the keyword mnemonic for foreign vocabulary learning. *Applied Cognitive Psychology, 21*(4), 499–526.

Gates, A. I. (1917). Recitation as a factor in memorizing. *Archives of Psychology*, 6(40).

Glover, J. A. (1989). The '"testing"' phenomenon: Not gone but nearly forgotten. *Journal of Educational Psychology, 81*, 392–399.

Gupta, M. W., Pan, S. C., & Rickard, T. C. (2022). Prior episodic learning and the efficacy of retrieval practice. *Memory & Cognition, 50*(4), 722–735.

Higbee, K. L. (2001). *Your memory: How it works and how to improve it* (2nd ed.). Marlow.

Jones, B. F., & Hall, J. W. (1982). School applications of the mnemonic keyword method as a study strategy by eight graders. *Journal of Educational Psychology, 74*(2), 230–237.

Jönsson, F. U., Kubik, V., Sundqvist, M. L., Todorov, I., & Jonsson, B. (2014). How crucial is the response format for the testing effect? *Psychological Research Psychologische Forschung, 78*(5), 623–633.

Kang, S. H. (2010). Enhancing visuospatial learning: The benefit of retrieval practice. *Memory & Cognition, 38*, 1009–1017.

Kang, S. H. K., McDermott, K. B., Roediger, H. L., & III. (2007). Test format and corrective feedback modify the effect of testing on long-term retention. *European Journal of Cognitive Psychology, 19*(4–5), 528–558.

Kang, S. H., & Pashler, H. (2014). Is the benefit of retrieval practice modulated by motivation? *Journal of Applied Research in Memory and Cognition, 3*(3), 183–188.

Karpicke, J. D. (2009). Metacognitive control and strategy selection: Deciding to practice retrieval during learning. *Journal of Experimental Psychology: General, 138*(4), 469–486.

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.

Karpicke, J. D., Lehman, & Aue, W. R. (2014). Retrieval-based learning: An episodic context account. In B. H. Ross (Ed.), *Psychology of learning and motivation* (Vol. 61, pp. 237–284). Elsevier.

Karpicke, J. D., & Smith, M. A. (2012). Separate mnemonic effects of retrieval practice and elaborative encoding. *Journal of Memory and Language, 67*(1), 17–29.

Keresztes, A., Kaiser, D., Kovács, G., & Racsmány, M. (2014). Testing promotes long-term learning via stabilizing activation patterns in a large network of brain areas. *Cerebral Cortex, 24*(11), 3025–3035.

Kim, Y.-E., Yu, S. L., Koenka, A. C., Lee, H., & Heckler, A. F. (2022). Can self-efficacy and task values buffer perceived costs? Exploring introductory- and upper-level physics courses. *The Journal of Experimental Education, 90*(4), 839–861.

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.

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*, 674–682.

Latimier, A., Peyre, H., & Ramus, F. (2021). A meta-analytic review of the benefit of spacing out retrieval practice episodes on retention. *Educational Psychology Review, 33*, 959–987.

Levin, M. E., & Levin, J. R. (1990). Scientific mnemonomies: Methods for maximizing more than memory. *American Educational Research Journal, 27*(2), 301–321.

Little, J., & McDaniel, M. A. (2015). Metamemory monitoring and control following retrieval for text. *Memory & Cognition, 43*, 85–98.

Martin, N. D., Nguyen, K., & McDaniel, M. A. (2016). Structure building differences influence learning from educational text: Effects on encoding, retention, and metacognitive control. *Contemporary Educational Psychology., 46*, 52–60.

Maurer, T. W., & Shipp, C. (2021). Challenges of shaping student study strategies for success. *Teaching & Learning Inquiry, 9*(1), 241–257.

McDaniel, M. A., & Butler, A. C. (2011). A contextual framework for understanding when difficulties are desirable. In A. Benjamin (Ed.), *Successful Remembering and Successful Forgetting: A Festschrift in Honor of Robert A. Bjork* (pp. 175–199). Taylor & Francis.

McDaniel, M. A., & Donnelly, C. M. (1996). Learning with analogy and elaborative interrogation. *Journal of Educational Psychology, 88*, 508–519.
McDaniel, M. A., Howard, D. C., & Einstein, G. O. (2009). The read-recite-review study strategy: Effective and portable. *Psychological Science, 20*, 516–522.
McDaniel, M. A., Kowitz, M. D., & Dunay, P. K. (1989). Altering memory through recall: The effects of cue-guided retrieval processing. *Memory & Cognition, 17*, 423–434.
McDaniel, M. A., & Masson, M. E. J. (1985). Altering memory representations through retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 11*, 371–385.
McDaniel, M. A., & Fisher, R. P. (1991). Tests and test feedback as learning sources. *Contemporary Educational Psychology, 16*, 192–201.
McNamara, D. S. (2004). SERT: Self-explanation reading training. *Discourse Processes, 38*, 130.
Miyatsu, T., & McDaniel, M. A. (2019). Adding the keyword mnemonic to retrieval practice: A potent combination for foreign language vocabulary learning? *Memory & Cognition, 47*, 1328–1343.
Morris, P. E., & Fritz, C. O. (2000). The name game: Using retrieval practice to improve the learning of names. *Journal of Experimental Psychology: Applied, 6*, 124–129.
Morris, P. E., & Fritz, C. O. (2002). The improved name game: Better use of expanding retrieval practice. *Memory, 10*, 259–266.
Morris, P. E., Fritz, C. O., Jackson, L., Nichol, E., & Roberts, E. (2005). Strategies for learning proper names: Expanding retrieval practice, meaning and imagery. *Applied Cognitive Psychology, 19*(6), 779–798.
Pan, S. C., & Rickard, T. C. (2018). Transfer of test-enhanced learning: Meta-analytic review and synthesis. *Psychological Bulletin, 144*(7), 710–756.
Pashler, H., Cepeda, N. J., Wixted, J. T., & Rohrer, D. (2005). When does feedback facilitate learning of words? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*(1), 3–8.
Pressley, M., & Levin, J. R. (1981). The keyword method and recall of vocabulary words from definitions. *Journal of Experimental Psychology: Human Learning and Memory, 7*(1), 72–76.
Pressley, M., Levin, J. R., Hall, J. W., Miller, G. E., & Berry, J. K. (1980). The keyword method and foreign word acquisition. *Journal of Experimental Psychology: Human Learning and Memory, 6*(2), 163–173.
Pressley, M., Levin, J. R., Kuiper, N. A., Bryant, S. L., & Michener, S. (1982). Mnemonic versus nonmnemonic vocabulary-learning strategies: Additional comparisons. *Journal of Educational Psychology, 74*(5), 693–707.
Pressley, M., Symons, S., McDaniel, M. A., Snyder, B. L., & Turnure, J. E. (1988). Elaborative interrogation facilitates acquisition of confusing facts. *Journal of Educational Psychology, 80*, 268–278.
Pyc, M. A., & Rawson, K. A. (2010). Why testing improves memory: Mediator effectiveness hypothesis. *Science, 330*(6002), 335. https://doi.org/10.1126/science.1191465
Rawson, K. A., & Dunlosky, J. (2022). Successive relearning: An underexplored but potent technique for obtaining and maintaining knowledge. *Current Directions in Psychological Science, 31*, 362–368.
Rea, S. D., Wang, L., Muenks, K., & Yan, V. X. (2022). Students can (mostly) recognize effective learning, so why do they not do it? *Journal of Intelligence, 10*, 127.
Roediger, H. L., & Butler, A. C. (2011). The critical role of retrieval practice in long-term retention. *Trends in Cognitive Sciences, 15*, 20–27.
Roediger, H. L., & Karpicke, J. D. (2006a). The power of testing memory: Basic research and implications for educational practice. *Perspectives on Psychological Science, 1*, 181–210.
Roediger, H. L., III., & Karpicke, J. D. (2006b). Test-enhanced learning: Taking memory tests improves long-term retention. *Psychological Science, 17*, 249–255.
Roediger, H. L., III., Putnam, A. L., & Smith, M. A. (2011). Ten benefits of testing and their applications to educational practice. In J. P. Mestre & B. H. Ross (Eds.), *Psychology of learning and motivation: Cognition in education* (Vol. 55, pp. 1–36). Elsevier.
Roelle, J., & Nückles, M. (2019). Generative learning versus retrieval practice in learning from text: The cohesion and elaborationof the text matters. *Journal of Educational Psychology, 111*(8), 1341–1361.
Roelle, J., Schweppe, J., Endres, T., Luchner, A., von Aufschnaiter, C., Renkl, A., ... & Vorholzer, A. (2022). Combining retrieval practice and generative learning in educational contexts: Promises and challenges. Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie, 54 (4):142–150.
Rosenshine, B., Meister, C., & Chapman, S. (1996). Teaching students to generate questions: A review of the intervention studies. *Review of Educational Research, 66*, 181–221.

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.

Seifert, T. L. (1993). Effects of elaborative interrogation with prose passages. *Journal of Educational Psychology, 85*, 642–651.

Shriberg, L. K., Levin, J. R., McCormick, C. B., & Pressley, M. (1982). Learning about “famous” people via the keyword method. *Journal of Educational Psychology, 74*(2), 238–247.

Susser, J. A., & McCabe, J. (2013). From the lab to the dorm room: Metacognitive awareness and use of spaced study. *Instructional Science, 41*(2), 345–363.

Thomas, R. C., & McDaniel, M. A. (2013). Testing and feedback effects on front-end control over later retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 39*, 437–450.

Toppino, T. C., & Cohen, M. S. (2009). The testing effect and the retention interval: Questions and answers. *Experimental Psychology, 56*(4), 252–257.

Trumbo, M. C. S., McDaniel, M. A., Hodge, G. K., Jones, A. P., Matzen, L. E., Kittinger, R. S., Kittinger, L. I., & Clark, V. P. (2021). Is the testing effect ready to be put to work?: Evidence from the laboratory to the classroom. *Translational Issues in Psychological Science, 7*(3), 332–355.

Valentine, T., Brennen, T., & Bredart, S. (1996). *The cognitive psychology of proper names*. Routledge.

Vestergren, P., & Nyberg, L. (2014). Testing alters brain activity during subsequent restudy: Evidence for test-potentiated encoding. *Trends in Neuroscience and Education, 3*(2), 69–80.

Walsh, D. A., & Jenkins, J. J. (1973). Effects of orienting tasks on free recall in incidental learning: “Difficulty”, “ effort”, and “process” explanations. *Journal of Verbal Learning and Verbal Behavior, 12*, 481–488.

Whiffin, J. W., & Karpicke, J. D. (2017). The role of episodic context in retrieval practice effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 43*, 1036–1046.