Multiple-choice pretesting potentiates learning of related information

Jeri L. Little^{1,2,3} · Elizabeth Ligon Bjork³

Published online: 13 May 2016
© Psychonomic Society, Inc. 2016

Abstract Although the testing effect has received a substantial amount of empirical attention, such research has largely focused on the effects of tests given after study. The present research examines the effect of using tests prior to study (i.e., as pretests), focusing particularly on how pretesting influences the subsequent learning of information that is not itself pretested but that is related to the pretested information. In Experiment 1, we found that multiple-choice pretesting was better for the learning of such related information than was cued-recall pretesting or a pre-fact-study control condition. In Experiment 2, we found that the increased learning of non-pretested related information following multiple-choice testing could not be attributed to increased time allocated to that information during subsequent study. Last, in Experiment 3, we showed that the benefits of multiple-choice pretesting over cued-recall pretesting for the learning of related information persist over 48 hours, thus demonstrating the promise of multiple-choice pretesting to potentiate learning in educational contexts. A possible explanation for the observed benefits of multiple-choice pretesting for enhancing the effectiveness with which related nontested information is learned during subsequent study is discussed.

Keywords Pretesting · Testing effects · Learning · Multiple choice · Test-potentiated learning

That pretesting can potentiate the learning of pretested information during subsequent study has been demonstrated for a variety of materials and across a variety of methodologies (e.g., Arnold & McDermott, 2013; Kornell, Hays, & Bjork, 2009). This finding has clear applications for learning in educational contexts (Anderson & Biddle, 1975; Kane & Anderson, 1978; Little & Bjork, 2011; Pressley, Tanenbaum, McDaniel, & Wood, 1990; Richland, Kornell, & Kao, 2009; Rothkopf, 1966). Relatively little work, however, has examined whether such pretesting can also potentiate the learning of information that was not itself pretested, or instead, whether the learning of such information might be impaired by such pretesting. This question is particularly pertinent to conditions in which such non-pretested material is related to, and thereby potentially confusable or competitive with, the specifically pretested information.

The issue raised previously is important because students are often presented with large amounts of related and thus potentially confusable information (e.g., anatomy, geography, history courses), and they are sometimes pretested on subsets of that information (e.g., with “clicker” questions presented at the beginning of a lecture), followed by more comprehensive exams that would actually count toward their grades and that would include the testing of non-pretested related information as well. Of particular interest in the present research was whether, and if so how, pretesting might improve the learning of such related, and thus possibly competitive, non-pretested information. The present research

✉ Jeri L. Little
jerilittle@gmail.com

¹ Department of Psychology, Washington University in St. Louis, St. Louis, MI, USA

² Present address: Department of Psychology, Hillsdale College, 33 E. College St., Hillsdale, MI 49242, USA

³ Department of Psychology, University of California, Los Angeles, CA, USA

examines such possible effects for multiple-choice pretests in comparison to cued-recall pretests. We see a comparison of these two types of testing formats to be particularly critical, given the already extensive use of multiple-choice testing in large classrooms and its increasing use as a tool for learning with classroom response systems (e.g., “clickers”; Buhay, Best, & McGuire, 2010; Lantz & Stawiski, 2014; Mayer et al., 2009; Smith et al., 2009), with such tests often given before instruction (Glass, Brill, & Ingate, 2008).

Previously, Little and Bjork (2011) demonstrated that answering multiple-choice questions prior to study can improve one’s ability to learn such competitive non-pretested related information compared to no-pretest control conditions—when answers to related questions had appeared as incorrect alternatives on the pretest questions. In their Experiment 3, participants read passages (e.g., about Yellowstone National Park and the planet Saturn), with one preceded by multiple-choice questions (e.g., *What geyser is thought to be the oldest in the world? a. Old Faithful, b. Steamboat Geyser, c. Castle Geyser, d. Daisy Geyser*) and one preceded by the study of facts containing all of the information in the questions, including the incorrect alternatives (a fact-study control condition; e.g., *Castle Geyser is thought to be the oldest geyser in the world, not Old Faithful, Steamboat Geyser, or Daisy Geyser*). When later asked the name of the tallest (rather than the oldest) geyser, participants were better able to access the correct answer, *Steamboat Geyser*, when they had answered a multiple-choice pretest question for which *Castle Geyser* was the correct answer than when they had studied the comparable fact beforehand, demonstrating that the benefit depends upon more than simply preexposure to the correct answers during a prestudy activity. How multiple-choice pretest questions induce such enhanced learning of related information and whether their benefits would surpass those provided by cued-recall pretest questions are issues that we address in the present research.

Cued-recall tests are generally heralded for their ability to serve as effective learning events (e.g., Foos & Fisher, 1988), and indeed, most extant research has shown cued-recall tests to be more effective than multiple-choice tests for improving later retention of the tested material—both when tests serve as post-tests and, critically for the present purposes, when they serve as pretests (e.g., Anderson & Biddle, 1975; Hamaker, 1986). In addition, smaller benefits for the retention of information related to tested information have also been shown, but there are some notable exceptions with regard to cued-recall testing (e.g., retrieval-induced forgetting observed with educational materials; Little, Storm, & Bjork, 2011) and, furthermore, the types of relationships that have been examined are limited. In Hamaker’s review, for example, in which he reported the finding of greater benefits from cued-recall testing versus multiple-choice testing on retention of related information, such

“related” items were almost exclusively of a noncompetitive nature—like rewording of questions from initial test to later test and definition questions on an initial test with concept questions on a later test. The type of competitive relationship between tested and nontested information explored in the present research—namely, the situation in which an answer to one question might compete with the answer to a related question—has yet to be systematically examined with different pretest formats (but see Little, Bjork, Bjork, & Angello, 2012, for an examination with posttest formats).

The relationship between pretested and non-pretested information and the format of the pretest questions with respect to this relationship is important because pretesting may not always be expected to potentiate learning of related information. To the extent, for example, that pretesting focuses attention toward the subsequent encoding of the specific pretested information, it might draw attention away from the encoding of other information that does not remind the learner of a pretest question (Fraser, 1967; Hamaker, 1986; Reynolds & Anderson, 1982; Reynolds, Standiford, & Anderson, 1979), and we postulate that such might be the case when pretested and non-pretested information is related in a competitive manner and the relationship is not established and/or brought to mind in the pretest question. By contrast, the multiple-choice format, we would argue, can serve to make the relationship between pretested and competitive non-pretested information explicit. When, for example, answers to competitive non-pretested questions serve as incorrect alternatives on a multiple-choice pretest question, learners may be alerted to the need to learn (or at least primed to learn) information pertaining to the question and all the alternatives. That is, in contrast to a cued-recall pretest question (on which competitive information is not presented), a multiple-choice pretest question (on which the answers to potential competitive questions are presented as incorrect alternatives) might direct attention more broadly—toward the processing, not only of the information explicitly addressed by the pretest question but also to the processing of information having to do with the incorrect alternatives, with a consequent increase in performance on questions about that competitive information should they appear on a later test. Following from the example presented previously (Little & Bjork, 2011), a multiple-choice question in which four different geysers serve as alternatives may increase the likelihood that learners attend to information about all of those geysers when later encountered, whereas a cued-recall question may direct attention just to the geyser that was addressed by the question (e.g., the oldest geyser).

Furthermore, a commonsensical hypothesis ensuing from this attention-direction notion would be that any improvements in performance would be associated with *increased* time spent attending to that information (both for pretested information and for information pertaining to the incorrect competitors), but this is not a foregone conclusion. First, with respect to the explicitly pretested information, learners may

not spend more time on such information when it is encountered during subsequent study, especially if feedback had been provided on the pretest. Learners will have had practice with such information—perhaps making its subsequent encoding seem more fluent—and thus may not feel it necessary to allocate additional time to it (Little & McDaniel, 2015). But what should the assumption be regarding non-pretested information? Unless by chance it is something already known to the participant, such information should not be learned or encoded during the taking of the pretest but would have to be acquired through subsequent reading of the passage. On the one hand, it seems reasonable to assume that any performance benefits occurring for such information would be associated with increased time spent processing that information when encountered during subsequent reading of the passage. On the other hand, and more interestingly, perhaps the nature of the multiple-choice pretesting format invokes a type of processing during the attempt to select an answer such that the pertinent non-tested information can then be learned more efficiently when encountered during subsequent study. We explored these possibilities in the present Experiment 2.

The present research

Our overarching goal in the present research was to compare the benefits afforded by multiple-choice pretests to those afforded by cued-recall pretests and to assess the extent to which any observed benefits might be afforded by allocation of time during study. Additionally, we hoped to replicate the finding of Little and Bjork (2011) that showed multiple-choice pretesting to be better than a pre-fact-study control condition for the learning of related information and to assess the possible contribution of study time allocation to this difference. Toward this goal, we conducted three experiments, and in all of these experiments we defined related information in terms of its competitiveness to the pretested material. In Experiment 1, we compared effects of multiple-choice pretesting to cued-recall pretesting (as well as pre-fact-study activities, serving as no-pretest control conditions, similar to Little & Bjork, 2011; see also, Richland et al., 2009). In Experiment 2, we assessed how learners allocate their study time during reading of a subsequent passage following pretests (multiple choice vs. cued recall) or the study of facts. In Experiment 3, we examined the effects of multiple-choice pretesting as compared to cued-recall pretesting over both a 5-min and 2-day delay. Additionally, Experiments 1 and 2 employed pretests with feedback, whereas Experiment 3 employed pretests without feedback. Feedback was necessary in Experiments 1 and 2 owing to our goal to directly compare effects of pretesting to fact study prior to study of the to-be-learned passages.

Experiment 1

Method

Participants and design Forty-eight students at the University of California, Los Angeles, participated for partial credit in a psychology course. As shown in Fig. 1, all participants read three passages (A, B, and C)—one preceded by a pretest activity (e.g., Passage A); one preceded by a fact-study activity (e.g., Passage B), a condition controlling for exposure to the to-be-tested information; and one not preceded by any type of activity but receiving extended reading time (e.g., Passage C). Additionally, for the pretest activity, a random half of the participants received a multiple-choice pretest for one passage, and these participants' fact-study activity presented facts with competitors (described more fully below), and these conditions are illustrated on the left in Fig. 1. The remaining participants received a cued-recall pretest and a comparable fact-study activity (i.e., facts without competitors); these conditions are illustrated on the right in Fig. 1. Then, on a delayed final cued-recall test, all participants answered (a) pretested questions and non-pretested related questions about their pretested passage; (b) questions about previously studied facts and questions related to those studied facts about the passage preceded by their fact-study activity; and (c) non-pretested questions from their extended-reading passage. The order of the passages and the order of the conditions were counterbalanced across participants.

This design allowed us to compare (a) multiple-choice pretesting to cued-recall pretesting with a between-subjects manipulation and (b) multiple-choice pretesting to fact study (which controlled for exposure to the alternatives) with a within-subjects manipulation. These were our main comparisons.

Materials The passages were about the planet Saturn, Yellowstone National Park, and stimulant drugs, each containing about 700 words and with 10 pairs of facts for each passage. The two facts in each pair were semantically related in that they could be used to construct pairs of multiple-choice questions that tested the same topic (e.g., geysers) and had the same four alternatives (e.g., *Old Faithful*, *Steamboat Geyser*, *Castle Geyser*, and *Daisy Geyser*) but different correct answers (e.g., *What is the tallest geyser in Yellowstone National Park?* Answer: *Steamboat Geyser*; *What geyser is thought to be the oldest in the world?* Answer: *Castle Geyser*). The Yellowstone passage and multiple-choice questions are shown in the Appendix.

When a passage was to be given a multiple-choice pretest, one question from each pair appeared on the pretest, with the remaining question from each pair then appearing as a related non-pretested question on the delayed final cued-recall test, with assignment of questions to the pretest or as related

Fig. 1 Schematic of a sample procedure from Experiment 1 where MC = multiple-choice and CR = cued-recall tests, respectively. In the experiment, order of conditions was counterbalanced across participants

questions on the final cued-recall test counterbalanced across participants. When a passage was to be given a cued-recall pretest, the same procedure for creating the pretests and the delayed final cued-recall test was used, but the questions presented on the pretest appeared without the alternatives. When a passage was to be preceded by the study of facts with competitors, the questions were changed to statements that included the items that would have been presented as incorrect alternatives for the corresponding multiple-choice question (e.g., *Castle Geyser is thought to be the oldest geyser in the world, not Old Faithful, Steamboat Geyser, or Daisy Geyser*). Facts without competitors were the same except for the absence of the alternatives (e.g., *Castle Geyser is thought to be the oldest geyser in the world*).

Procedure All participants read three passages, with two of the passages immediately preceded by a pre-reading activity: either a pretest or a fact-study activity, as shown in Fig. 1. The pre-reading activities for the random half assigned to the multiple-choice condition were a 10-question multiple-choice pretest and the study of 10 facts, including competitors. The pre-reading activities for the half assigned to the cued-recall condition were a 10-question cued-recall pretest and the study of 10 facts without competitors.

For the passage that was pretested, participants were given 20 s to answer each question. After typing their answer, they continued to view the question until the 20 s elapsed. After each question, participants received feedback (i.e., correct answer presented below the question) for 4 s. Immediately after

the pretest, participants were given 6 min to read the pretested passage. For the passage preceded by fact study, participants were presented with 10 facts for 24 s each and told to spend the full 24 s thinking about the fact. Immediately thereafter, participants were given 6 min to read the passage. For the passage not preceded by a pretest or fact-study activity, participants received 10 min to read the passage, with the extra 4 min compensating for the time taken for the pretest or fact-study activities, and thus controlling for time on task. For each passage, participants were told that if they finished reading early, they should spend the remainder of the time studying, like they would for a class.

Finally, after a 5-min retention interval during which they played Tetris, participants received a 60-item final cued-recall test (20 items each for the pretest, fact-study, and extended-reading conditions), with questions presented one at a time on the computer screen. Of the 20 final-test questions about the pretested passage, half were identical to the 10 questions on the initial pretest (except now appearing as cued-recall questions in the multiple-choice pretest condition), while the other half consisted of the corresponding related questions for that passage that had not appeared on the pretest. Of the 20 final-test questions about the passage preceded by fact study, 10 items tested information central to each fact and 10 items tested information related to each fact. (That is, questions for a given passage were the same on the final cued-recall test whether that passage had appeared in one of the pretest conditions or the pre-fact-study condition.) Of the 20 final-test questions about the extended-reading passage, all came from the set of 10-paired questions constructed for that passage, and none had been previously exposed.

Because we considered assessing possible effects on the retention of non-pretested related information to be our most important research goal, non-pretested questions from the pretest conditions and questions pertaining to the competitive alternatives presented in the fact-study conditions were included in the first half of the final test along with the comparison questions from the extended-reading condition. This ordering would reduce output interference from the stronger pretested or prestudied information, whose corresponding questions were tested in the second half of the final test, along with the corresponding extended-reading questions (see Roediger & Schmidt, 1980; Smith, 1971). The order in which passages were presented, the condition (pretest, fact study, or extended reading) to which each passage was assigned, and which 10-item question set and 10-item fact set were presented during the pretest or fact-study conditions were counterbalanced across participants.

Results and discussion

Pretest performance Overall, participants chose the correct answer to 28% ($SD = 6\%$) of the multiple-choice questions on

the multiple-choice pretests, which was not significantly different from chance (25%), $t(23) = 1.03$, $p = .31$, and they produced the correct answer to 7% ($SD = 10\%$) of the cued-recall questions on the pretest.

Final-test performance for related items We had predicted that multiple-choice pretesting would enhance the learning of related information as compared to cued-recall pretesting and fact study. The performance related to this prediction appears in Fig. 2, where we show overall correct percentages on the final cued-recall test for information related to questions on the pretests or to prestudied facts. As indicated there, performance on related items appears to be better when those items are related to multiple-choice pretest questions than when they are related to cued-recall pretest questions or to prestudied facts.

A 2 (pre-reading activity: multiple-choice/facts with alternatives, cued-recall/facts without alternatives) $\times$ 3 (item type: pretest related, fact-study related, extended reading) ANOVA examining performance for related information revealed an interaction between pre-reading activity and item type, $F(2, 45) = 6.27$, $p < .01$, $\eta^2 = .22$, as well as a main effect of item type, $F(2, 45) = 8.79$, $p < .001$, $\eta^2 = .28$. With respect to our main question of interest as to whether multiple-choice pretesting would improve learning of related information more than would cued-recall pretesting, a planned independent-samples t test confirmed that items related to multiple-choice pretest questions ($M = 50\%$, $SE = 4\%$) were recalled at a higher rate than were items related to cued-recall

pretest questions ($M = 37\%$, $SE = 4\%$), $t(46) = 2.29$, $p = .03$, $d = 0.66$. Additionally, with respect to our second main question of interest as to whether our findings would replicate those of Little and Bjork (2011), showing that multiple-choice pretests improve learning as compared to a control condition in which statements with alternatives were presented for prestudy, a planned pairwise t test also confirmed that higher performance was obtained in the multiple-choice pretest condition than in the comparable fact-study condition ($M = 36\%$, $SE = 4\%$), $t(23) = 3.40$, $p < .01$, $d = 0.69$. As in that previous study, this result demonstrates that simple exposure of alternatives—including the correct answer to what will be a related question on the final test—is not what leads to the increased learning of related information following competitive multiple-choice pretests. Finally, performance was marginally higher for items related to multiple-choice questions than for questions from the extended-reading condition ($M = 42\%$, $SE = 4\%$), $t(23) = 1.99$, $p = .06$.

In contrast, participants' performance on questions related to cued-recall pretest questions did not differ from their performance on questions related to facts ($M = 33\%$, $SE = 4\%$), $t(23) = 1.31$, $p = .20$, and was worse than performance for corresponding items in the extended-reading condition ($M = 50\%$, $SE = 5\%$), $t(23) = 3.05$, $p < .001$.

Final-test performance for pretested items and previously studied facts We had predicted that multiple-choice pretesting would be as good as cued-recall pretesting and fact study for the learning of pretested information. Relevant to this prediction, the overall performance on the final cued-recall test for information that was either pretested or studied as facts prior to reading the passage is presented in Fig. 3. As indicated there, correct performance for questions pretested as multiple-choice questions or prestudied as facts with competitors appears to be comparable to that for questions pretested as cued-recall questions or prestudied as facts without competitors, respectively. A 2 (pre-reading activity: multiple-choice/facts with alternatives, cued-recall/facts without alternatives) $\times$ 3 (item type: pretested, fact-study control, extended-reading) ANOVA did not reveal an interaction between pre-reading activity and item type on performance for previously tested information, $F(2, 45) = 0.08$, $p = .93$. A main effect of item type, $F(1, 45) = 42.34$, $p < .001$, $\eta^2 = .65$, however, was observed.

In assessing these effects with planned pairwise t tests, we found that—in contrast to the general assumption of cued-recall testing being more beneficial than multiple-choice testing for retention—correct performance for items pretested as cued-recall questions ($M = 75\%$, $SE = 3\%$) did not significantly differ from that for items pretested as multiple-choice questions ($M = 73\%$, $SE = 5\%$), $t(46) = 0.30$, $p = .76$, but correct performance on either was better than performance on corresponding items in the extended-reading condition ($M = 43\%$,

Fig. 2 Correct performance percentages on the final cued-recall test in Experiment 1 for information related to pretested questions or related to information studied as facts prior to reading the passage, as a function of the type of pretest (multiple-choice or cued-recall, both with feedback) and corresponding fact-study condition. Facts corresponding to the multiple-choice pretest condition contained competitors, one of which served as the correct answer to the related question; facts corresponding to the cued-recall pretest condition did not contain competitors. White bars show performance for questions from the passage that received extended reading and that were tested in the first half of the final test. Error bars represent ± 1 SEM

Fig. 3 Correct performance percentages on the final cued-recall test in Experiment 1 for information that was pretested or studied as facts prior to reading a passage, as a function of the type of pretest (multiple choice or cued recall, both with feedback) and corresponding fact-study condition. Facts corresponding to the multiple-choice pretest condition contained competitors; facts corresponding to the cued-recall pretest condition did not contain competitors. White bars show performance for questions from the passage that received extended reading and that were tested in the second half of the final test. Error bars represent ± 1 SEM

$SE = 4\%$, $t(23) = 5.74$, $p < .01$, and $M = 42\%$, $SE = 4\%$, $t(23) = 6.92$, $p < .01$, for multiple-choice and cued-recall conditions, respectively. Overall, pretesting ($M = 74\%$, $SE = 3\%$) improved performance more than fact study ($M = 67\%$, $SE = 3\%$), $t(47) = 2.05$, $p < .05$.

The overall pattern of results obtained in Experiment 1 suggests that cued-recall and multiple-choice pretests differentially affect later recall of competitive related information. Specifically, as indicated by performance on related items in the two types of pretest conditions, competitive related information appears to be learned less well as a consequence of taking a cued-recall pretest than as a consequence of taking a multiple-choice pretest with competitive incorrect alternatives, even when the final test utilizes a cued-recall format. This result suggests that being exposed to incorrect alternatives associated with related information in the context of a multiple-choice question induces processing that improves later learning of that related information beyond that provided by a cued-recall question. It is not exposure of the alternatives alone, however, that induces deep processing leading to improved performance: Studying facts with the alternatives provided does not provide the same benefit, a finding first shown by Little and Bjork (2011) and replicated in the present experiment.

What processing differences during subsequent reading of passages preceded by a multiple-choice pretest versus a cued-recall pretest or the study of facts might produce these differences in final-exam performance? One possibility is that learners spend more time studying such competitive information following a multiple-choice pretest than following a cued-recall pretest or the study of facts. A more interesting

possibility, however, is that the enhanced performance for these items on the later test occurs—not because learners spend more time on that critical information but rather because they are able to process such information more efficiently. Perhaps these competitive incorrect alternatives receive deep processing during pretesting, which then facilitates the learning of information pertaining to them when it is encountered during subsequent reading of the passage.

To gain some leverage on this issue, we examined how long learners spent reading sentences containing information pertaining to the pretested (or prestudied) information and information related to that pretested (or prestudied) information in Experiment 2. The basic design employed was the same as that of Experiment 1 except that reading time was self-paced sentence by sentence. Accordingly, while we expected to replicate the basic patterns of performance observed in the pretesting and fact-study conditions of Experiment 1, we thought the performance in the reading-only condition of Experiment 2 might be lower than that of the reading-only (extended-reading) condition of Experiment 1 because we did not expect participants to spend four additional minutes reading this passage—although they were free to do so.

Experiment 2

Method

Participants Seventy individuals in the Washington University in St. Louis community participated for either pay or partial credit in a psychology course.

Design, materials, and procedure The materials were the same as those used in Experiment 1, and the design and procedure were the same except that participants read the passages, sentence by sentence, at their own pace. Reading times for each sentence—the main data of interest in this experiment—were recorded.

Results and discussion

Pretest performance Overall, participants chose the correct answer to 31% ($SD = 14\%$) of the multiple-choice questions on the multiple-choice pretests, which was significantly different from chance (25%), $t(35) = 2.77$, $p = .01$, and they produced the correct answer to 13% ($SD = 13\%$) of the cued-recall questions on the pretest.

Final-test performance for related items As we expected, given the self-paced presentation procedure used in Experiment 2, participants spent less time on the passage in the reading-only condition in Experiment 2 ($M = 4.8$ min, $SD = 1.6$) than what was allocated toward extended reading in

Experiment 1. Thus, as we expected, performance in the reading-only condition in the present experiment was numerically lower than that in the extended-reading condition in Experiment 1. With this exception, however, the basic pattern of results—that is, for pretested and pre-fact-study passages—was consistent across Experiments 1 and 2.

Overall performance on the final cued-recall test for information related to questions on the pretests or to the prestudied facts is shown in Fig. 4. A 2 (pre-reading activity: multiple-choice/facts with alternatives, cued-recall/facts without alternatives) $\times$ 3 (item type: pretest related, fact-study related, reading only) ANOVA revealed an interaction between pre-reading activity and item type on performance for related information, $F(2, 67) = 3.61$, $\eta^2 = .10$, $p = .03$. As indicated in Fig. 4 and replicating the results in Experiment 1, correct performance for items related to multiple-choice pretest questions ($M = 51\%$, $SE = 4\%$) was significantly better on the final test than was performance for items related to cued-recall pretest questions ($M = 41\%$, $SE = 3\%$), $t(68) = 2.05$, $p = .04$, $d = 0.50$, and significantly better than performance for questions related to facts with competitors ($M = 41\%$, $SE = 3\%$), $t(34) = 2.77$, $p < .01$, $d = 0.47$. Performance for questions related to previously tested multiple-choice questions was numerically but not reliably higher than that for questions from the reading-only condition ($M = 45\%$, $SE = 3\%$), $t(34) = 1.58$, $p = .12$.

Performance on the final cued-recall test following cued-recall pretesting ($M = 41\%$, $SE = 3\%$) was not better than performance following the comparable fact study ($M = 44\%$, $SE = 3\%$), $t(34) = 0.79$, $p = .44$, or following reading only ($M = 41\%$, $SE = 3\%$), $t(34) = 0.08$, $p = .94$.

Fig. 4 Correct performance percentages on the final cued-recall test in Experiment 2 for information related to pretested questions as a function of the type of pretest (multiple choice or cued recall, both with feedback) and corresponding fact-study condition. Facts corresponding to the multiple-choice pretest condition contained competitors; facts corresponding to the cued-recall pretest condition did not contain competitors. White bars show performance for initially nontested questions from the nontested passage that were tested in the first half of the final test. Error bars represent ± 1 SEM

Final-test performance for pretested items and previously studied facts Overall performance on the final cued-recall test for information that was either pretested or studied as facts prior to reading the passage is presented in Fig. 5. As indicated there, correct performance for items pretested as multiple-choice questions or studied as facts with competitors appears to be comparable to that for items pretested as cued-recall questions or prestudied as facts without competitors, respectively. A 2 (pre-reading activity: multiple-choice/facts with alternatives, cued-recall/facts without alternatives) $\times$ 3 (item type: pretested, fact-study control, reading only) ANOVA revealed a marginally significant interaction between pre-reading activity and item type on performance for previously tested or studied information, $F(2, 67) = 2.86$, $p = .06$, $\eta^2 = .08$. A significant main effect of item type, $F(2, 67) = 77.75$, $\eta^2 = .70$, $p < .001$, was obtained.

Correct performance for items pretested as cued-recall questions ($M = 77\%$, $SE = 3\%$) was marginally greater than that for items pretested as multiple-choice questions ($M = 69\%$, $SE = 3\%$), $t(68) = 1.91$, $p = .06$. The cued-recall pretest provided a performance advantage compared to studying facts ($M = 62\%$, $SE = 3\%$), $t(34) = 4.17$, $p < .001$, $d = 0.70$, but a multiple-choice pretest did not ($M = 66\%$, $SE = 4\%$), $t(34) = 0.65$, $p = 0.44$. Performance for both types of questions, however, was reliably better than for the corresponding questions in the reading-only condition ($M = 37\%$, $SE = 3\%$), $t(34) = 7.48$, $p < .001$, $d = 1.26$, in the multiple-choice condition; ($M = 41\%$, $SE = 3\%$), $t(34) = 9.33$, $p < .001$, $d = 1.58$, in the cued-recall condition.

Fig. 5 Correct performance percentages on the final cued-recall test in Experiment 2 for previously tested information as a function of the type of pretest (multiple choice or cued recall, both with feedback) and corresponding fact-study condition. Facts corresponding to the multiple-choice pretest condition contained competitors; facts corresponding to the cued-recall pretest condition did not contain competitors. White bars show performance for initially nontested questions about the nontested passage that were tested in the second half of the final test. Error bars represent ± 1 SEM

Reading times Given the overall replication of the performance patterns observed in Experiment 1, we next analyzed the results directly pertinent to our main question of interest in Experiment 2: reading-time allocation to the critical sentences (i.e., those containing information pertaining to the pretested or prestudied information or containing information related to the pretested or prestudied information). Of key interest was whether individuals allocated more or less time to these critical sentences when a passage had been pretested or preceded by fact study versus when it had not. As previously discussed, the straightforward prediction would be that the better performance shown for related information in the multiple-choice pretest condition as compared to the cued-recall pretest condition or the comparable fact-study condition would be associated with increased time allocation; a more interesting possibility, however, was that better performance would not be driven by increased time allocation but by more effective or efficient processing of related information in the multiple-choice pretest condition.

Related items A 2 (pre-reading activity: multiple-choice/facts with alternatives, cued-recall/facts without alternatives) $\times$ 3 (item type: pretest related, fact-study related, reading only) ANOVA indicated a marginal interaction between pre-reading activity and item type on reading time allocated to related items, $F(2, 67) = 2.74$, $p = .07$, $\eta^2 = .08$. Planned paired-samples t tests showed that following a multiple-choice pretest, individuals spent less time (in ms) reading sentences containing information pertaining to the non-pretested related question ($M = 5,858$, $SE = 241$) than they spent reading comparable sentences from the reading-only passage ($M = 6,539$, $SE = 385$), $t(34) = 2.20$, $p = .04$, $d = 0.41$, and comparable sentences from the fact-study passage ($M = 6,613$, $SE = 405$), $t(34) = 2.18$, $p = .04$, $d = 0.40$. Following a cued-recall pretest, however, time to read sentences containing information pertaining to the non-pretested related question ($M = 6,309$, $SE = 380$) did not differ from time spent reading comparable sentences from the reading-only passage ($M = 6,426$, $SE = 411$), $t(34) = 0.40$, $p = .69$, or from the fact-study passage ($M = 6,031$, $SE = 315$), $t(34) = 1.01$, $p = .32$. The overall difference between the time spent in the multiple-choice condition and the cued-recall condition was not reliable, however, $t(68) = 1.00$, $p = .32$. Although the interaction of pre-reading activity and item type on reading times did not reach statistical significance, and a reliable difference in overall reading times did not emerge between the two pretest conditions, the intriguing result is that the highest performance on the final test (i.e., that related to multiple-choice pretesting) was associated with the least numerical amount of time spent in study.

Pretested and prestudied items A 2 (pre-reading activity: multiple-choice/facts with alternatives, cued-recall/facts

without alternatives) $\times$ 3 (item type: pretested, fact-study control, reading only) ANOVA did not indicate an interaction between pre-reading activity and item type on reading time allocated to pretested or prestudied items, $F(2, 67) = 0.04$, $p = .96$. A main effect of item type, however, did emerge, $F(2, 67) = 10.57$, $p < .001$, $\eta^2 = .24$. Paired-samples t tests confirmed that individuals spent less time reading sentences containing information that had been previously tested than reading comparable sentences from a passage that had not been preceded by a pre-reading activity of any type in both the multiple-choice condition ($M = 5,400$, $SE = 208$ vs. $M = 6,398$, $SE = 387$), $t(34) = 2.80$, $p < .01$, $d = 0.47$, and the cued-recall condition ($M = 5,367$, $SE = 280$ vs. $M = 6,335$, $SE = 370$), $t(34) = 3.27$, $p < .01$, $d = 0.55$. They also spent significantly less time, compared to the reading-only condition, on sentences with information pertaining to facts with competitors ($M = 5,478$, $SE = 306$ vs. $M = 6,398$, $SE = 387$), $t(34) = 3.37$, $p < .01$, $d = 0.57$, and on facts without competitors ($M = 5,423$, $SE = 272$ vs. $M = 6,335$, $SE = 370$), $t(34) = 2.50$, $p = .02$, $d = 0.42$. Preexposure in any form reduced the time participants spent during subsequent study on that directly preexposed information.

The goal of Experiment 2 was to gain further insight into how pretesting might direct subsequent study, especially for related information, and whether time spent studying would relate to later test performance. One clear possibility was that testing would direct learners' attention toward tested material: Learners would spend more time processing this material, thus leading to its enhanced retrieval on a later exam. A more interesting possibility was that testing would enhance later recall performance without learners having to spend more time on this information during subsequent study. Experiment 2 provides evidence consistent with this second possibility, and its results are especially intriguing in terms of retention of related information. Although participants could have learned the specific pretested information before study—that is, while taking the pretest—their learning of any specific information associated with incorrect alternatives would seem unlikely. They could learn, however, that such information may be important to attend to later, and the results suggest that they then learned such information more efficiently. That is, following a multiple-choice pretest, learners spent less time on related sentences (compared to that spent on corresponding sentences in the reading-only or pre-fact-study conditions), but (based on their later test performance) they learned such information better, at least compared to the fact-study control condition. Although following a multiple-choice pretest, learners did not spend significantly less time on sentences containing related information than they did following a cued-recall pretest, the efficacy of learning that information was clearly better in the case of the multiple-choice pretest than in the case of the cued-recall pretest, as indicated by the significantly better performance for such items on the final cued-recall test.

How is related information being processed more efficiently? On the one hand, it seems reasonable to assume that choosing incorrect answers would lead to interference and potentially the persistence of incorrect associations (see, e.g., research on the effect of choosing an incorrect photograph on later eyewitness identification; Gorenstein & Ellsworth, 1980). On the other hand, it also seems reasonable to assume that because the pretest is followed by learning (and not preceded by it), participants are likely to seek out corrections to their errors (knowing that they have likely made many), and thus incorrect alternatives are likely to garner more attention and processing during subsequent passage reading when they have been previously presented as part of a multiple-choice pretest question than as part of a prestudied fact. Furthermore, it also seems reasonable to assume that an incorrect alternative and its associated information would be processed more deeply during subsequent reading when it had been previously selected as the answer to the pretest question than when it had not been. The reasoning for this conjecture is that when learners do not already have sufficient knowledge to select the correct answer, they may turn to some other basis for selecting an answer that may involve deep processing. Perhaps, for example, they select the one that seems most pleasant to them or the one that is most relevant to their own personal lives—two types of processing that have been shown to improve subsequent recall of information even when subjects are not expecting a later test (see Craik & Tulving, 1975; Packman & Battig, 1978). Processing that leads to enhanced encoding of a given alternative may then lead to more efficient and effective learning of information pertaining to that alternative when encountered later in the subsequently studied passage.

As a way of assessing this possibility, we calculated the rates of correct performance on non-pretested related questions on the final test conditional upon the answers to such questions having been incorrectly chosen in the prior multiple-choice pretest versus having not been chosen on that test. The results of this analysis for Experiments 1 and 2, as well as for the to-be-presented Experiment 3, are shown in Table 1. As revealed there by the percentages shown in the last two columns, participants were more likely to answer a related question correctly when its answer had been incorrectly chosen on the pretest than when it had not been chosen (i.e., the correct answer for the pretested question or one of the other incorrect alternatives had been chosen instead). We do not see this pattern of results arising because participants are simply choosing the same answer twice for two different questions (participants, in fact, were unlikely to provide their previously chosen answer for both questions). Instead, we believe this pattern of results to be consistent with our conjecture that the processing of related information is enhanced during subsequent reading of the passage as a consequence of thinking more deeply about all the alternatives during the initial test and, in these particular cases, having specifically chosen that alternative as

Table 1 Likelihood of Choosing the Answer to the Non-Pretested Related Question as an Incorrect Response on the Pretest and Resulting Conditional Correct Performance on Non-Pretested Related Questions on the Final Test in Experiments 1–3

Experiment	Percentage of times chosen on pretest ^a	Final test correct performance chosen on pretest	Final test correct performance not chosen on pretest ^b
1	23%	61%	47%
2	19%	65%	45%
3 (5 min)	19%	66%	50%
3 (48 hr)	20%	44%	27%

Note. ^a Percentage of times an incorrect alternative that will later be a correct answer to a nontested related question was incorrectly chosen as an answer on the pretest. ^b Includes correct choices on pretest.

the correct answer. (Note also that these results are inconsistent with the possibility that participants are correctly recalling information pertaining to the incorrect alternatives during the pretest, with such recall providing the benefit in retention evident on the final test, as this benefit is strongest when learners incorrectly chose the incorrect alternative on the pretest that will later be the correct answer to a related non-pretested question on the final cued-recall test.)

One might also wonder how the selection of an incorrect alternative influenced the time spent studying information pertaining to that alternative during subsequent reading. Thus, for Experiment 2, we also analyzed the reading times associated with correct performance, as a function of whether participants chose the incorrect alternative on the initial test that would later become the answer to the related question on the final test versus when they chose another alternative on the initial test. When participants incorrectly chose an alternative that would later become the answer to the related question and then answered that related question correctly, the average reading time ($M = 4,376$, $SE = 620$) was shorter than when participants chose any other alternative and then answered the question correctly ($M = 5,865$, $SE = 310$), $t(34) = 2.34$, $p = .03$. This finding is consistent with the idea that selection of an answer made it easier for participants to learn the information related to that choice when it was later encountered in their reading of the passage; when another answer had been selected, participants could still answer the related question correctly, but their learning of the information necessary to do so was associated with more time to process it when encountered in their subsequent reading of the passage. The results thus far are consistent with the notion that answering multiple-choice questions leads to processing of competitive incorrect alternatives that is deeper than that occurring when such alternatives are provided within facts. Furthermore, this benefit for the learning of related information is especially clear for the item that is chosen, both in terms of the level of performance and

the efficiency of learning. Neither cued-recall pretesting nor exposure to the alternatives in facts directs reading in such an effective manner.

Experiment 3

The results of Experiments 1 and 2 demonstrate a potentially unique benefit of using multiple-choice pretests to promote learning in educational contexts. Namely, not only are multiple-choice pretests about as beneficial as cued-recall pretests or studying comparable facts for the retention of the pretested information, they are actually more effective than these other activities for the learning of competitive related information. This latter advantage of multiple-choice pretests would, in particular, seem to be a valuable asset for their fruitful application to educational settings. After all, it seems likely that most instructors would typically want to ask similar, but not identical, questions on a later exam to assess whether students learned concepts more generally rather than simply being able to remember specific answers to previously tested questions.

In educational contexts, however, delays between learning activities and a final test of the to-be-learned information are often longer than those we used in the present Experiments 1 and 2. Thus, in Experiment 3, we used a modified procedure to assess whether benefits of pretesting would persist across longer delays and, in particular, whether the ability of multiple-choice pretesting to foster retention of nontested related information would persist, as compared to cued-recall pretesting. If the performance benefits observed for previously tested and related information in Experiments 1 and 2 resulted as the consequence of deep processing, then we should expect such benefits to persist over time.

In the modified procedure of Experiment 3, participants were first given pretests for two different passages, with one being a multiple-choice pretest and the other being a cued-recall pretest. Then, participants read three passages; two of the passages had received a pretest and the third had not received a pretest, thus serving as a reading-only baseline. A final cued-recall test was then administered for information presented in each of the three passages at either a short (5 min) or long (48 hrs) delay to assess the persistence of pretesting effects on retention for both pretested and non-pretested related information.

Method

Participants and design Seventy-two individuals in the Washington University in St. Louis community participated for either pay or partial credit in a psychology course and were all recruited for a two-session study. All participants took two pretests (one multiple choice and one cued recall), followed by

the reading of three passages. On the final cued-recall test, participants answered pretested questions and related non-pretested questions from the two pretested passages, and nontested questions from the passage not receiving an initial pretest. We randomly manipulated whether participants were given the final test after a 5-min or 48-hour delay.

Materials and procedure The materials were the same as those used in Experiments 1 and 2, and the procedure was the same except for the following differences. First, participants took two pretests before reading any passages: one consisting of multiple-choice questions for one of the three to-be-presented passages and the other consisting of cued-recall questions for another one of the three to-be-presented passages. Importantly, and unlike the procedures of Experiments 1 and 2 (but like the procedure used by Little & Bjork, 2011, in their Experiment 1), the pretests did not provide feedback. After administration of the two pretests, all three passages were read in immediate succession and for 6 min each. Which passages were pretested or not pretested, the type of pretest received, and the order of the two pretests were counterbalanced across participants. After a 5-min distractor task, participants randomly assigned to the 5-min delay condition were given a final cued-recall test for both pretested and non-pretested related information. Those participants randomly assigned to the 24-hour delay condition were dismissed after the 5-min distractor task and asked to return in 2 days.

As in the present Experiments 1 and 2, non-pretested related questions and the corresponding questions from the reading-only passage appeared in the first half of the final cued-recall test, whereas pretested questions and their corresponding questions from the reading-only passage appeared in the second half of the final cued-recall test.

Results and discussion

Pretest performance For the multiple-choice pretest, participants chose the correct answer to 34% ($SD = 15\%$) and 33% ($SD = 15\%$) of the questions in the 5-min delay condition and the 48-hour delay condition, respectively, with this level of performance being greater than chance performance (25%), $t(71) = 4.81$, $p < .01$. For the cued-recall pretest, participants provided the correct answer to 10% ($SD = 9\%$) and 9% ($SD = 11\%$) of the questions in the 5-min delay condition and the 48-hour delay condition, respectively.

Final-test performance for related items Overall correct performance on non-pretested related questions is shown in Fig. 6. As indicated there, the advantage of using multiple-choice pretests over using cued-recall pretests appears to occur, regardless of the delay used. That is, a 2 (delay: short, long) $\times$ 3 (pre-reading activity: multiple-choice pretest, cued-

Fig. 6 Correct performance percentages on the final cued-recall test in Experiment 3 for information related to pretested questions as a function of the type of pre-reading activity (multiple-choice pretest, cued-recall pretest, no pretest/reading only) and retention interval. White bars show performance for initially nontested questions from the nontested topic, which were tested in the first half of the final test. Error bars represent ± 1 SEM

recall pretest, no pretest/reading only) ANOVA did not indicate an interaction between delay and pre-reading activity on retention of related information, $F(2, 69) = 0.05$, $p = .955$. There were, however, main effects of both pre-reading activity, $F(2, 69) = 5.11$, $p < .01$, $\eta^2 = .13$, and delay, $F(1, 70) = 46.81$, $p < .001$, $\eta^2 = .40$. A paired-samples t test confirmed that, regardless of delay, taking a multiple-choice pretest ($M = 43\%$, $SE = 3\%$) was better for the retention of related information than was taking a cued-recall pretest ($M = 37\%$, $SE = 2\%$), $t(71) = 2.46$, $p = .02$, $d = 0.29$. Additionally, whereas taking a multiple-choice pretest improved performance for related questions as compared to that for corresponding questions in the reading-only condition ($M = 36\%$, $SE = 2\%$), $t(71) = 2.94$, $p < .01$, $d = .35$, taking a cued-recall pretest did not produce such a benefit for related questions as compared to that for corresponding questions in the reading-only condition, $t(71) = 0.57$, $p = .57$.

Final-test performance for pretested items Performance obtained on the final cued-recall test for information initially pretested on a multiple-choice test, a cued-recall test, or not pretested (reading only) at both the 5-min and 48-hr retention intervals is presented in Fig. 7. As indicated there, correct performance on questions pretested with either type of test appears to be higher than that for comparable non-pretested questions, regardless of retention interval.

A 2 (delay: short, long) $\times$ 3 (pre-reading activity: multiple-choice pretest, cued-recall pretest, no pretest/reading only) ANOVA did not indicate an interaction between delay and pre-reading activity on retention of pretested information, $F(2, 69) = 1.36$, $p = 0.26$, although there were main effects of both pre-reading activity, $F(2, 69) = 44.20$, $p < .001$, $\eta^2 = .56$, and delay, $F(1, 70) = 31.11$, $p < .001$, $\eta^2 = .31$. Indeed, as confirmed by planned-comparison tests, overall final-test performance for information pretested with either a multiple-choice

Fig. 7 Correct performance percentages on the final cued-recall test in Experiment 3 for previously tested information as a function of the type of pre-reading activity (multiple-choice pretest, cued-recall pretest, no pretest/reading only) and the retention interval. White bars show performance for initially nontested questions about the nontested topic, which were tested in the second half of the final test. Error bars represent ± 1 SEM

test ($M = 57\%$, $SE = 3\%$) or a cued-recall test ($M = 52\%$, $SE = 2\%$) was better than that for corresponding information from the reading-only passage ($M = 34\%$, $SE = 2\%$), $t(71) = 8.49$, $p < .01$, $d = 1.00$, and $t(71) = 7.94$, $p < .01$, $d = 0.94$, respectively.

The pattern of results observed in Experiment 3 demonstrates that a pretesting benefit for the retention of pretested information (as compared to non-pretested information) can persist over a delay that would be considered more educationally realistic (i.e., 48 hrs) and, additionally, in the case of a multiple-choice pretest, that a pretesting benefit for the retention of related information can also persist over such a delay. In evaluating this outcome, it is important to note that although we presented participants with the same textual materials as those used in Experiments 1 and 2, we introduced changes to the design and procedure of Experiment 3 that could have been expected to impact the previously observed benefits of pretesting. First, in Experiment 3, all participants received both a multiple-choice pretest and a cued-recall test, which they took before presentation of any of the to-be-studied passages. Thus, whereas in Experiments 1 and 2 a given passage was always presented for study immediately after participants had taken the relevant pretest, a given passage was unlikely to be presented for study immediately after participants had taken the relevant pretest in Experiment 3. In fact, for some participants, as much as 20 minutes of filled activity (e.g., taking another pretest, reading other passages) may have occurred between their taking of a given pretest and their opportunity to read the relevant passage—a procedural difference that would seem likely to diminish the participants' ability to direct their attention to previously tested or related information during the reading of the relevant passage. This procedure also served to directly compare effects of cued-recall pretests to multiple-choice pretests in the same

participants and without specific fact-study conditions being tied to those pretest conditions. Second, in Experiment 3, we did not provide feedback during the pretests, which also might be expected to produce a smaller pretesting benefit for tested information. Despite these changes, however, a pretesting benefit emerged in Experiment 3, which we believe underscores the strength and persistence of such effects.

Another procedural change made in Experiment 3 that we believe to be important, as it speaks to the generality of the pretesting effect, was our change from using an extended-reading condition (i.e., in Experiment 1, the time allotted for studying the reading-only passage was the sum of the time allowed for the study of a pretested passage plus the time taken for administering its pretest) to giving all passages the same amount of time for reading. Our use of an extended-reading condition in Experiment 1 models the situation in which students might elect to spend a given amount of time for their study activities (or instructors have allowed only a fixed amount of classroom time for such activities), and, thus, taking a pretest before reading a passage would take away from the time they have available for reading/studying the passage. Students, however, may not always follow such total-time study routines. They might, for example, sometimes elect to engage in extra study activities, such as taking a pretest, in addition to the total time they spend studying or reading a passage or chapter.

General discussion

Increasingly in classrooms, instructors are using pretests given with the help of classroom response systems, such as clickers, with the goal of increasing attention and learning in their courses (Buhay et al., 2010; Glass et al., 2008; Lantz & Stawiski, 2014; Mayer et al., 2009; Smith et al., 2009). Our findings suggest that using multiple-choice questions for such purposes would make the subsequent learning of related competitive information more effective than would the use of cued-recall questions (Experiments 1–3) or even studying facts beforehand (Experiments 1 and 2), and that this benefit of multiple-choice pretesting over cued-recall pretesting persists over a delay of 48 hours and even when feedback is not provided on the pretest (Experiment 3).¹ Furthermore, our findings from Experiment 2 provide evidence that the enhanced recall of related information observed on a later test is not associated with increased time devoted to the study of that information, suggesting that, following a multiple-choice pretest, information pertaining to the incorrect alternatives on

a previously given multiple-choice test is learned more efficiently. Finally, it should be noted that this benefit of a multiple-choice pretest occurs even though the format of the later test is cued recall.

How does multiple-choice pretesting (compared to cued-recall pretesting or studying facts) benefit the retention of related information?

As discussed in the Results and Discussion section for Experiment 2, the present results are consistent with the notion that multiple-choice pretesting *can* direct attention more broadly during subsequent study than can cued-recall pretesting—both toward information pertaining to the question at hand and toward information pertaining to the incorrect alternatives. It, of course, may not be the case that all multiple-choice questions presented in pretests would serve as effective learning tools in this way. Rather, they may need to possess certain qualities in order to do so. The alternatives, for example, may need to be competitive. This quality has been shown to be necessary for the learning of related, nontested information to be enhanced in prior research by Little and Bjork (2015; see also, Little et al., 2012) when multiple-choice tests are given following study rather than prior to study. The mechanism proposed for the enhanced retention of related, nontested information observed by these investigators, however, does not readily apply to the present situation of giving multiple-choice pretests, as it involves the assumption that test takers think back to the previously studied passage in search of information related to the alternatives to help them in their selection of the correct answer. By so doing, information about the competitive incorrect alternatives can be retrieved and thereby strengthened in addition to information related to the specifically correct answer. Then, if such nontested information becomes the bases for related questions on the final cued-recall test, the test taker's ability to answer such questions will be enhanced. In the present situation, however, when the multiple-choice tests are presented before study, accessing of information contained in the passage is not possible during the taking of the test. Thus, in light of this inherent difference in the pretest versus the posttest situation, it remains unclear whether the incorrect alternatives in multiple-choice pretests questions need to be competitive in order to produce the enhanced processing of information related to them during the test takers' subsequent reading of the passage, and the resolution of this issue will require future research.

For the time being, however, our results are consistent with the notion that presenting competitive information as multiple-choice alternatives leads to deeper processing of that information (particularly with respect to the item that is selected) than does presenting competitive information within a fact. We argue that learners, although not able to successfully answer most questions, may bring to mind any information

¹ Since the conclusion of this laboratory experiment, evidence consistent with our finding that the potentiation of learning resulting from multiple-choice pretests can last for educationally realistic durations has also been obtained in an actual classroom (Soderstrom & Bjork, 2015).

that they can—whether entirely relevant to the to-be-studied information or not—in order to help them decide upon an answer. They might, for example, in trying to answer a question asking for the name of the oldest geyser in Yellowstone Park, turn to semantic knowledge (e.g., *Old Faithful* has “old” in its name—thus, maybe it is the oldest) or they might choose an answer for reasons that are arbitrary yet helpful for memory (e.g., selecting the most pleasant one or the one that is most relevant to their own personal lives; Craik & Tulving, 1975; Packman & Battig, 1978). Then, when participants see these items in the text, they will be more likely to (a) notice them and (b) encode them, perhaps using the information that they had previously used to select them as mediator(s) for their encoding (see, e.g., Pyc & Rawson, 2010, and also Kornell, Hays, & Bjork, 2009; Richland et al., 2009, for related arguments about why pretesting might improve learning of pretested information). It is possible, however, that other entirely different mechanisms also play a role in enhancing the learning of related information when encountered during subsequent study, and this possibility should be examined in future research.

Multiple-choice pretesting and the retention of pretested information

In the present research, performance on the final cued-recall test did not differ for questions about information that was either directly tested in a multiple-choice pretest question or studied as a fact before subsequent passage reading. This finding is intriguing because, in the present context, one might have expected that later performance on questions for information previously studied as a fact would be better than that for information previously tested in a multiple-choice question. To clarify: Our fact-study activity essentially provided participants with 24 seconds of access to the question and correct answer prior to their being asked that same question on the final test. In contrast, in our multiple-choice pretesting activity, participants were presented with the correct answer for only 4 seconds after spending 20 seconds considering the question and, typically, choosing an incorrect answer for it. Nonetheless, the ability of participants to recall such tested information was not impaired in the later final cued-recall test compared to their ability to recall the same information when it had, instead, appeared in the fact-study condition. That lack of a difference between later recall of previously tested and previously studied information, combined with the finding that participants’ ability to answer questions about related information was significantly better in the multiple-choice pretest condition than in the fact-study condition, would seem to suggest that trying to answer a multiple-choice question about to-be-studied information, even if one cannot do so correctly, would function as a more powerful way to potentiate overall future learning than simply being preexposed to that

information in the form of facts. Additionally, this pattern of results would seem to call for a reconsideration of claims made by proponents of errorless learning, who would argue that the making of errors should be avoided during learning activities, with the assumption that such errors will tend to persist and thereby impair future correct learning (e.g., Guthrie, 1952; Skinner, 1958). We also see the pattern of results obtained in the present Experiments 1 and 2 to be consistent with findings observed by Kornell et al. (2009; see also, Carpenter, 2011; Grimaldi & Karpicke, 2012; Hays, Kornell, & Bjork 2013; Huelser & Metcalfe, 2012; Knight, Ball, Brewer, DeWitt, & Marsh, 2012; Kornell, 2014; Vaughn & Rawson, 2012; Yue et al. 2015) in their research on test-potentiated learning.

Concluding comments

Trying to answer multiple-choice pretest questions—at least ones constructed with competitive alternatives, as was used in the present research—can provide unique benefits to the learning of subsequently studied related and competitive information that outweighs that of other activities, such as cued-recall pretests and/or prestudy of to-be-encountered facts. This unique advantage of multiple-choice pretests may well come about owing to the operation of two simultaneous but different processes stemming from the taking of such tests. First, the answering of multiple-choice questions with competitive alternatives induces deeper processing of the information contained in the question than what results from simple exposure to that competitive information (e.g., in the form of facts with competitors), and second, it directs attention more broadly (but without requiring more time) during subsequent reading than does answering cued-recall pretest questions. From an educational perspective, then, the present results suggest an effective, efficient, and practical way for instructors to enhance their students learning: giving them multiple-choice pretests containing competitive alternatives before their study of potentially confusable topics.

Author Note This research was supported by a Collaborative Activity Award from the James S. McDonnell Foundation (Grant 29192G). We thank Robert Bjork, Nate Kornell, Barbara Knowlton, and members of CogFog for helpful insights. Experiment 1 appears as part of the dissertation of Jeri L. Little.

Appendix

Following are the passage and 20 multiple-choice questions constructed about Yellowstone National Park. Correct answers are shown in boldface.

Yellowstone National Park

Yellowstone, America's first national park, is located at the headwaters of the Yellowstone River, for which it takes its name. In the 18th century, French trappers named the river *Roche Jaune*, which is a translation of the Minnetaree name for "Rock Yellow River."

Approximately 96% of Yellowstone National Park is located in the state of Wyoming, but the park extends into neighboring states of Idaho and Montana. The Yellowstone region is traditionally home to several Native American tribes, including the Nez Perce, Crow, and Shoshone. European explorers first entered the region in the early 19th century. One such explorer, John Colter, explored the region with the Lewis and Clark Expedition. In 1806, he left the expedition to explore the region with a group of fur trappers. Colter explored the region for 4 years, until two of his partners were killed by Blackfoot Indians. After an 1856 exploration, mountain man Jim Bridger reported observing boiling springs, spouting water, and a mountain of glass and yellow rock. These reports were largely ignored because Bridger was known for being a "spinner of yarns." His stories did arouse the interest of explorer and geologist Ferdinand Vandeveer Hayden, who, in 1895, started a 2-year survey of the upper Missouri River region. Bridger and United States Army surveyor W. F. Reynolds acted as guides. After exploring the Black Hills region in what is now the state of South Dakota, the party (Bridger, Hayden, and Reynolds) neared the Yellowstone River, but heavy snows forced them to turn back.

In 1872, Yellowstone officially became a national park. President Ulysses S. Grant signed a bill into law that deemed Yellowstone America's first national park. Nathaniel Langford was appointed as the park's first superintendent, serving for 5 years although denied a salary, funding, and staff. Langford's lack of resources left Yellowstone vulnerable to poachers, vandals, and others seeking to raid its resources. In 1916, pioneering industrialist and conservationist Stephen Tyng Mather, along with journalist and writer Robert Sterling Yard, spearheaded a publicity campaign to promote the creation of a federal agency to oversee the national parks. Mather helped to establish other parks, and in 1917, administration of Yellowstone was transferred to the National Park Service.

Yellowstone is the centerpiece of the 20-million acre Greater Yellowstone Ecosystem and is home to a variety of animals, including elk, moose, and bison. Threatened species include the endangered gray wolf, the threatened lynx, and grizzly bears. Only an estimated 600 grizzly bears live in the Greater Yellowstone Ecosystem. Population figures for elk are in excess of 30,000, making them the largest population of any large mammal species in Yellowstone. Bison are also well represented in the park.

Yellowstone National Park rests at an average altitude of 8,000 feet above sea level. The highest point in the park is atop Eagle Peak, which is almost 11,400 feet above sea level. The most prominent summit in the Yellowstone Plateau (although

not the highest) is Mount Washburn, at 10,243 feet. Nearby mountain ranges include the Beartooth Mountains to the north and the Absaroka Range to the east. Forests comprise 80% of the park's land area—the remaining land area is primarily comprised of grasslands. Of the eight conifer tree species documented, lodgepole pine forests cover 80% of the total forested area. Other conifers, such as the Douglas fir and whitebark pine, are found in scattered groves throughout the park. As of 2007, the whitebark pine is threatened by a fungus known as the white pine blister rust. Aspen and willow are the most common species of deciduous trees in the park.

Yellowstone National Park sits atop the Yellowstone Caldera, an active volcano, and the park is well known for its geothermal features. Hot springs are a common hydrothermal features in the park. Geysers are a special type of hot spring that erupts periodically. There are only 300 geysers in Yellowstone, making them the least common of the park's 10,000 geothermal features. The most famous geyser in the park is Old Faithful. The tallest active geyser in the park, as well as in the world, is Steamboat Geyser. Daisy Geyser is known for its predictability; it usually erupts every 90–110 minutes. Castle Geyser is thought to be the oldest geyser in the world. Besides geysers, Yellowstone has other geothermal features, including mud pots and fumaroles. A fumarole is an opening in the earth's crust that emits steam and gasses.

Set A

1. What explorer left the Lewis and Clark Expedition to explore the region with a group of fur trappers?
 - a. Reynolds
 - b. Colter
 - c. Bridger
 - d. Hayden
2. About 600 of what threatened species live within the Greater Yellowstone Ecosystem?
 - a. elk
 - b. bison
 - c. grizzly bears
 - d. gray wolf
3. Attacks by what tribe caused Colter to leave the Yellowstone region?
 - a. Minnetaree
 - b. Shoshone
 - c. Blackfoot
 - d. Nez Perce
4. What is the tallest geyser in Yellowstone National Park?

- a. Old Faithful
 - b. Steamboat Geyser
 - c. Castle Geyser
 - d. Daisy Geyser
5. In what year did Colter first explore Yellowstone with a group of fur trappers?
- a. 1806
 - b. 1856
 - c. 1872
 - d. 1917
6. The majority of Yellowstone National Park resides in what state?
- a. Idaho
 - b. South Dakota
 - c. Wyoming
 - d. Montana
7. What is the highest peak in Yellowstone National Park?
- a. Eagle Point
 - b. Mt. Washburn
 - c. Beartooth
 - d. Absoroka
8. Who was Yellowstone National Park's first superintendent, who served for 5 years devoid of salary, funding, and staffing?
- a. Mather
 - b. Yard
 - c. Grant
 - d. Langford
9. Which species of tree covers 80% of the total forested areas in Yellowstone National Park?
- a. Douglas fir
 - b. Lodgepole pine
 - c. Whitebark pine
 - d. Aspen
10. What type of geothermal feature is an opening in the earth's crust that emits steam and gasses?
- a. Geyser
 - b. Hot Spring
 - c. Fumarole
 - d. Mudpot

Set B

1. What mountain man reported observing boiling springs, sprouting water, and a mountain of yellow rock, but was largely ignored due to a reputation of being a "spinner of yarns?"
- a. Raynolds
 - b. Colter
 - c. Bridger
 - d. Hayden
2. What species makes up the largest population of a large mammal species in Yellowstone National Park?
- a. elk
 - b. bison
 - c. grizzly bears
 - d. gray wolf
3. French trappers named Yellowstone River *Roche Jaune*, probably a translation of what Native American tribe's name for Yellow Rock River?
- a. Minnetaree
 - b. Shoeshone
 - c. Blackfeet
 - d. Nez Perce
4. What geyser is thought to be the oldest in the world?
- a. Old Faithful
 - b. Steamboat Geyser
 - c. Castle Geyser
 - d. Daisy Geyser
5. In what year did Yellowstone become a national park?
- a. 1806
 - b. 1856
 - c. 1872
 - d. 1917
6. The Black Hills region is found primarily in what state?
- a. Idaho
 - b. South Dakota
 - c. Wyoming
 - d. Montana
7. At 10,243, what is the most prominent peak (but not the highest) in Yellowstone National Park?

- a. Eagle Peak
 - b. Mt. Washburn
 - c. Beartooth
 - d. Absoroka
8. Who signed a bill into law making Yellowstone the first national park?
- a. Mather
 - b. Yard
 - c. Grant
 - d. Langford
9. As of 2007, which tree, found only in scattered groves, is threatened by a specific fungus?
- a. Douglas fir
 - b. Lodgepole pine
 - c. Whitebark pine
 - d. Aspen
10. With only about 300 examples in the park, what is the least common type of hydrothermal feature in Yellowstone?
- a. Geyser
 - b. Hot Spring
 - c. Fumarole
 - d. Mudpot

References

- Anderson, R. C., & Biddle, W. B. (1975). On asking people questions about what they are reading. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 9, pp. 89–132). New York, NY: Academic Press.
- 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.
- Buhay, D., Best, L. A., & McGuire, K. (2010). The effectiveness of library instruction: Do student response systems (clickers) enhance learning? *The Canadian Journal for the Scholarship of Teaching and Learning*, 1, 5.
- Carpenter, S. K. (2011). Semantic information activated during retrieval contributes to later retention: Support for the mediator effectiveness hypothesis of the testing effect. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 37, 1547–1552.
- Craik, F. I. M., & Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General*, 104, 268–294.
- Foos, P. W., & Fisher, R. P. (1988). Using tests as learning opportunities. *Journal of Educational Psychology*, 80, 179–183.
- Frase, L. T. (1967). Learning from prose material: Length of passage, knowledge of results and position of questions. *Journal of Educational Psychology*, 58, 266–272.
- Glass, A. L., Brill, G., & Ingate, M. (2008). Combined online and in-class pretesting improves exam performance in general psychology. *Educational Psychology*, 28, 483–503.
- Gorenstein, G. W., & Ellsworth, P. C. (1980). Effect of choosing an incorrect photograph on a later identification by an eyewitness. *Journal of Applied Psychology*, 65, 616–622.
- Grimaldi, P. J., & Karpicke, J. D. (2012). When and why do retrieval attempts enhance subsequent encoding? *Memory & Cognition*, 40, 505–513.
- Guthrie, E. R. (1952). *The psychology of learning*. Oxford, UK: Harper.
- Hamaker, C. (1986). The effects of adjunct question on prose learning. *Review of Educational Research*, 56, 212–242.
- Hays, M. J., Kornell, N., & Bjork, R. A. (2013). When and why a failed test potentiates the effectiveness of subsequent study. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 39, 290–296.
- Huelser, B. J., & Metcalfe, J. (2012). Making related errors facilitates learning, but learners do not know it. *Memory & Cognition*, 40, 514–527.
- Kane, J. H., & Anderson, R. C. (1978). Depth of processing and interference effects in learning and remembering sentences. *Journal of Educational Psychology*, 70, 626–635.
- Knight, J. B., Ball, B. H., Brewer, G. A., DeWitt, M. R., & Marsh, R. L. (2012). Testing unsuccessfully: A specification of the underlying mechanisms supporting its influence on retention. *Journal of Memory and Language*, 66, 731–746.
- Kornell, N. (2014). Attempting to answer a meaningful question enhances subsequent learning even when feedback is delayed. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 40, 106–114.
- Kornell, N., Hays, M. J., & Bjork, R. A. (2009). Unsuccessful retrieval attempts enhances subsequent learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 35, 989–998.
- Lantz, M. E., & Stawiski, A. (2014). Effectiveness of clickers: Effect of feedback and the timing of questions on learning. *Computers in Human Behavior*, 31, 280–286.
- Little, J. L., & Bjork, E. L. (2011). Pretesting with multiple-choice questions facilitates learning. In L. Carlson, C. Hölscher, & T. Shipley (Eds.), *Proceedings of the 33rd annual conference of the Cognitive Science Society* (pp. 294–299). Austin, TX: Cognitive Science Society.
- Little, J. L., & Bjork, E. L. (2015). Optimizing multiple-choice tests as tools for learning. *Memory & Cognition*, 43, 14–26.
- Little, J. L., Bjork, E. L., Bjork, R. A., & Angello, G. (2012). Multiple-choice tests exonerated, at least of some charges: Fostering test-induced learning and avoiding test-induced forgetting. *Psychological Science*, 23, 1337–1344.
- Little, J. L., & McDaniel, M. A. (2015). Metamemory monitoring and control following retrieval-practice for text. *Memory & Cognition*, 43, 85–98.
- Little, J. L., Storm, B. C., & Bjork, E. L. (2011). The costs and benefits of testing text materials. *Memory*, 19, 346–359.
- Mayer, R. E., Stull, A., DeLeeuw, K., Almeroth, K., Bimber, B., Chun, D., & Zhang, H. (2009). Clickers in college classrooms: Fostering learning with questioning methods in large lecture classes. *Contemporary Educational Psychology*, 34, 51–57.
- Packman, J. L., & Battig, W. F. (1978). Effects of different kinds of semantic processing on memory for words. *Memory & Cognition*, 6, 502–508.
- Pressley, M., Tanenbaum, R., McDaniel, M. A., & Wood, E. (1990). What happens when university students try to answer prequestions that accompany textbook material? *Contemporary Educational Psychology*, 15, 27–35.
- Pyc, M. A., & Rawson, K. A. (2010). Why testing improves memory: Mediator effectiveness hypothesis. *Science*, 330, 335–335.

- Reynolds, R. E., & Anderson, R. C. (1982). Influence of questions on the allocation of attention during reading. *Journal of Educational Psychology*, 74, 623–632.
- Reynolds, R. E., Standiford, S. N., & Anderson, R. C. (1979). Distribution of reading time when questions are asked about a restricted category of text information. *Journal of Educational Psychology*, 71, 183–190.
- Richland, L. E., Kornell, N., & Kao, L. S. (2009). The pretesting effect: Do unsuccessful retrieval attempts enhance learning? *Journal of Experimental Psychology: Applied*, 15, 243–257.
- Roediger, H. L., & Schmidt, S. R. (1980). Output interference in the recall of categorized and paired-associate lists. *Journal of Experimental Psychology: Human Learning and Memory*, 6, 91–105.
- Rothkopf, E. Z. (1966). Learning from written instructive materials: An exploration of the control of inspection behavior by test-like events. *American Educational Research Journal*, 3, 241–249.
- Skinner, B. F. (1958). Teaching machines. *Science*, 128, 969–977.
- Smith, A. D. (1971). Output interference and organized recall from long-term memory. *Journal of Verbal Learning and Verbal Behavior*, 10, 400–408.
- Smith, M. K., Wood, W. B., Adams, W. K., Wieman, C., Knight, J. K., Guild, N., & Su, T. T. (2009). Why peer discussion improves student performance on in-class concept questions. *Science*, 323, 122–124.
- Soderstrom, N., & Bjork, E. L. (2015). *Using pretesting to enhance learning in the classroom*. Manuscript in preparation.
- Vaughn, K. E., & Rawson, K. A. (2012). When is guessing incorrectly better than studying for enhancing memory? *Psychonomic Bulletin & Review*, 19, 899–905.
- Yue, C. E., Soderstrom, N. C., & Bjork, E. L. (2015). Partial testing can potentiate learning of tested and untested material from multimedia lessons. *Journal of Educational Psychology*. doi:[10.1037/edu0000031](https://doi.org/10.1037/edu0000031). Advance online publication.