



Prequestioning and Pretesting Effects: a Review of Empirical Research, Theoretical Perspectives, and Implications for Educational Practice

Steven C. Pan¹ · Shana K. Carpenter²

Accepted: 4 September 2023
© The Author(s) 2023

Abstract

Testing students on information that they do not know might seem like a fruitless endeavor. After all, why give anyone a test that they are guaranteed to fail because they have not yet learned the material? Remarkably, a growing body of research indicates that such testing—formally known as *prequestioning* or *pretesting*—can benefit learning if there is an opportunity to study the correct answers afterwards. This *prequestioning effect* or *pretesting effect* has been successfully demonstrated with a variety of learning materials, despite many erroneous responses being generated on initial tests, and in conjunction with text materials, videos, lectures, and/or correct answer feedback. In this review, we summarize the emerging evidence for prequestioning and pretesting effects on memory and transfer of learning. Uses of pre-instruction testing in the classroom, theoretical explanations, and other considerations are addressed. The evidence to date indicates that prequestioning and pretesting can often enhance learning, but the extent of that enhancement may vary due to differences in procedure or how learning is assessed. The underlying cognitive mechanisms, which can be represented by a three-stage framework, appear to involve test-induced changes in subsequent learning behaviors and possibly other processes. Further research is needed to clarify moderating factors, theoretical issues, and best practices for educational applications.

Keywords Prequestions · Pretesting effect · Errorful generation · Guessing · Practice testing · Pre-instruction testing · Test-enhanced learning

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

✉ Steven C. Pan
scp@nus.edu.sg

¹ Department of Psychology, Faculty of Arts and Social Sciences, College of Humanities and Sciences, National University of Singapore, 9 Arts Link, Singapore 117572, Singapore

² School of Psychological Science, Oregon State University, Corvallis, USA

In the science of learning, the concept of *test-enhanced learning* has become largely synonymous with *retrieval practice*, which is the strategy of taking practice tests on information that has previously been learned. That association is justified: Retrieval practice is well-established as a potent enhancer of learning for a plethora of materials across a wide array of educationally relevant circumstances (for reviews, see Carpenter, 2012; Carpenter et al., 2022; Dunlosky et al., 2013; Pan & Rickard, 2018; Roediger & Karpicke, 2006; Rowland, 2014, and others). Taking practice tests on information that has already been learned helps consolidate knowledge and increases its future accessibility (Bjork, 1975; Roediger & Butler, 2011). Research dating back to the 1960s and 1970s, however, suggests that an alternative approach that is, in a way, the opposite of retrieval practice—taking practice tests *before* information has been learned—may also be beneficial for learning. That approach, which is commonly known as *prequestioning* or *pretesting* (and sometimes *adjunct prequestioning*, *errorful generation*, *failed testing*, *unsuccessful testing*, among other appellations), is the subject of this review.

The effects of prequestioning were first addressed in the adjunct questions literature, a body of research that peaked in the 1970s and examined the pedagogical consequences of test questions encountered before, during, or after the reading of text materials (i.e., *adjunct prequestions*, *adjunct postquestions*, or *advance questions*; for reviews, see Anderson & Biddle, 1975; Frase, 1968; Hamilton, 1985; Rickards, 1979; for a meta-analysis, see Hamaker, 1986). In an early example, Rothkopf (1966) experimentally manipulated whether participants attempted practice questions or not prior to reading excerpts of a book on marine life. When asked to recall information targeted by practice questions on a subsequent posttest, participants that had engaged in prior testing well outperformed participants that had not. The implication is that testing led participants to better encode—and, as a consequence, better remember—at least some portions of the information that they were exposed to. That performance advantage, which is now known as the *prequestioning effect* or the *pretesting effect*, underscores the potential of prequestioning and pretesting as useful learning strategies in their own right. Evidence of prequestioning and pretesting effects was repeatedly observed in that literature (e.g., Berlyne, 1954; Hartley, 1973; Memory, 1983; Pressley et al., 1990; Rickards, 1976a, 1976b; Samuels, 1969), but not in all cases (e.g., Frase, 1968; Gustafson & Toole, 1970; Rothkopf & Bisbicos, 1967; Sagaria & Di Vesta, 1977). Studies in the adjunct questions literature, however, commonly lacked appropriate reference conditions, had no controls (or measures) of time-on-task, and intermixed prequestions with retrieval practice (for discussion see Carver, 1971), all of which complicate interpretation.

The modern prequestioning and pretesting literature, in contrast, incorporates more robust experimental controls and exclusively focuses on the case of practice testing prior to learning activities. It can be traced to Kornell et al. (2009) and Richland et al. (2009), which investigated prequestioning and pretesting for the cases of trivia facts/paired associate words and science texts, respectively, and reported robust learning benefits. Those results revitalized interest in prequestioning and pretesting, and in the decade-and-half since, over 60 peer-reviewed articles have been published on the topic. Aside from a recent introductory review (Carpenter et al., 2023) and other works that address selected aspects of this literature (e.g., retrieval

success as in Kornell & Vaughn, 2016; learning from errors as in Mera et al., 2022; Metcalfe, 2017), however, there has not yet been a comprehensive, up-to-date review that focuses specifically on the prequestioning and pretesting literature.

The present review examines that literature to address the overarching question: Are prequestioning and pretesting viable learning strategies? Anecdotally, a common perspective among some cognitive and educational psychologists is that any benefits of such testing are small in magnitude and ephemeral (i.e., only manifesting when learning is assessed very shortly afterwards), especially relative to retrieval practice. This review examines the literature for evidence of that perspective. It begins by defining common approaches that have been used to investigate prequestioning and pretesting. Studies that focus on benefits for memory as well as transfer of learning are then examined. These studies involve laboratory, online, or classroom settings, various forms of practice testing, and different timing and retention intervals. Next, metacognitive factors and theoretical explanations for prequestioning and pretesting effects are considered. Finally, other relevant learning strategies and phenomena, educational applications, directions for future research, and conclusions are discussed.

It should be noted that the terms “prequestion” and “pretest” have been used interchangeably in the literature. Historically, “prequestion” dates back to the adjunct questions literature, whereas “pretest” grew in prominence following Richland et al.’s (2009) titular mention of the “pretesting effect” and has been used in over half of the relevant studies published since. “Prequestion” has been used most frequently for studies involving text or video materials, although “pretesting” has also been used in cases involving such materials. This review includes studies that have used either of these terms, and as such, for ease of exposition, we henceforth use an all-encompassing alternative, “pre-instruction testing.”

How Effects of Pre-instruction Testing Are Investigated

Studies in the pre-instruction testing literature commonly feature an experimental design with two conditions: an experimental group and a control group. Participants may be randomly assigned to either group. Alternatively, in a within-participants design, all participants undergo an experimental condition and a control condition and, across those conditions, learn two or more sets of materials (with at least one set randomly assigned to each condition). Participants in the experimental group or condition will (1) engage in practice testing of information that is to be learned (i.e., target materials), (2) have an opportunity to learn the correct answers to the practice tests that they had experienced, and then (3) take a posttest. During practice testing, participants may attempt anywhere from one question to several dozen questions, and the pace of such testing may be carefully timed or self-paced. The posttest, which is also called a final test or criterial test, may occur at any time after (1) and (2) have completed, including immediately or after a delay ranging from several minutes to several weeks. The posttest is used to measure learning.

Although all studies in the literature commonly feature an experimental group or condition that undergoes steps (1) to (3), the precise implementation of (1) and (2)

varies between studies and often depends on the materials being learned. In some cases, (1) and (2) occur in separate phases, whereas in other cases, both are intermingled. Separation of (1) and (2) is more common when the materials are relatively content-rich, as in the case of text passages. The most common implementations of (1) and (2) are further considered in the next section.

The control group or condition typically does not engage in any practice testing. Rather, the initial activity usually is an opportunity to learn correct answers (or, more accurately, correct information that includes the answers to questions that will appear on a subsequent posttest). This learning opportunity is followed by a posttest that is essentially identical to the posttest that is administered to the experimental group. In some cases, however, a non-testing control activity may occur first. That control activity may simply control for time on task (e.g., solving math problems as in the case of Pan et al., 2020a, 2020b). Alternatively, the activity may also have the potential to facilitate learning of target materials (e.g., reading but not answering questions as in the case of Richland et al., 2009, experiment 5). In such cases, the control activity is followed by the opportunity to learn correct information.

The difference in performance between the experimental and control groups or conditions on the posttest, if any, serves as the measure of the effects of pre-instruction testing. In the literature, better performance in the experimental group or condition—that is, higher proportion correct on the posttest—has commonly been observed. The magnitude of that advantage is the focus of at least two recent quantitative meta-analyses (King-Shepard et al., n.d.; St. Hilaire et al., 2023).

Arrangement of Practice Testing and Subsequent Learning Opportunities

Schematics depicting the two most common implementations of pre-instruction testing, which differ primarily in the arrangement of (1) and (2), are presented in Fig. 1. In the upper panel, (1) and (2) occur in separate phases (i.e., all practice testing occurs prior to any opportunity to learn the correct answers). This approach has commonly been employed when pre-instruction tests are used to learn content from text passages (e.g., Little & Bjork, 2016), educational videos (e.g., Carpenter & Toftness, 2017), or live lectures (e.g., Carpenter et al., 2018). It is common for (2) to occur immediately after (1), although a delay of anywhere from several minutes to over several hours may occur. As the correct answers are embedded within a larger body of content, participants usually must discover those answers by engaging in searching behaviors—that is, by sifting through text, video, or lecture materials to find content that is directly relevant to the questions they had attempted. With this approach, participants are often exposed to more than just the correct answers, and as discussed in the next section, that exposure has repercussions for how learning can be assessed on the posttest.

In the lower panel, (1) and (2) are intermingled using practice testing with immediate *correct answer feedback*. Specifically, after each question, the correct answer is displayed for participants to study. This approach, which often requires computer presentation, has been used in the cases of learning word pairs (e.g., Huelser &

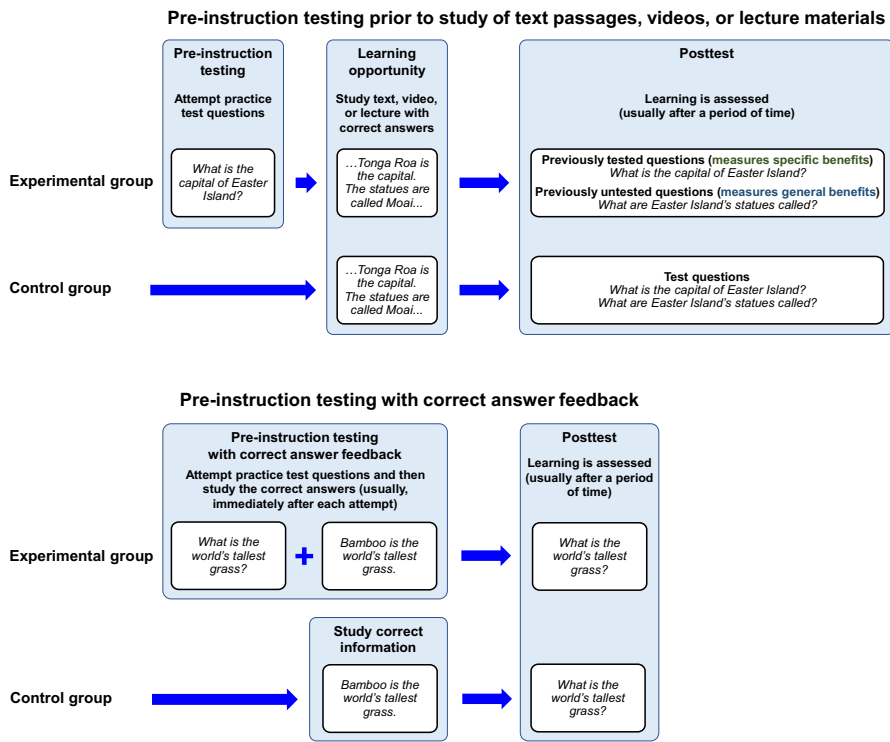


Fig. 1 Two common approaches for investigating the effects of pre-instruction testing on learning. In the upper panel, participants engage in practice testing prior to a learning opportunity or engage in that learning opportunity without any practice testing at all. This approach is commonly used with relatively information-rich materials such as text passages. In the lower panel, participants engage in practice testing with immediate correct answer feedback or study correct information without any practice testing at all. This approach is commonly used with simpler materials such as word pairs or facts. With both approaches, learning is assessed on a subsequent posttest

Metcalfe, 2012), foreign language translations (e.g., Seabrooke et al., 2019a, 2019b), and trivia facts or statements (e.g., Kornell, 2014). In such cases, learning the correct answer is a matter of attending to the provided feedback and no searching is needed. Moreover, exposure to additional materials beyond the correct answers is usually quite limited or nonexistent.

Although the two approaches depicted in Fig. 1 encompass most studies in this literature, other approaches exist. For instance, in some studies, the provision of correct answer feedback is not immediate, but rather delayed by several minutes or longer (e.g., Kornell, 2014), and there may be intervening activities beforehand (e.g., studying other items). In other studies, both immediate correct answer feedback and subsequent learning opportunities are provided (e.g., Pan & Sana, 2020, experiment 4). That is, after each practice question, the correct answer is displayed, and after completing practice testing with correct answer feedback, participants have a second opportunity to learn the correct answers from a text passage, video, lecture,

or other materials. In yet other cases, pre-instruction testing is combined with, or compared against, retrieval practice (e.g., Geller et al., 2017; Lima & Jaeger, 2020). Such approaches are considered later in this review.

Yet another approach involves *interpolated prequestioning* or *interpolated pre-testing*. It entails repeated cycles of practice testing followed by learning opportunities (e.g., Carpenter & Toftness, 2017; Pan et al., 2020a, 2020b). A schematic of this approach is displayed in Fig. 2. Interpolated testing can be used to divide a set of target materials, such as a video lecture, into a series of smaller segments. Participants still engage in practice testing before learning the correct answers, but do so one segment at a time. This approach originated with the adjunct questions literature, wherein text passages were often interspersed with questions preceding relevant content at intervals of one or more paragraphs (e.g., Rothkopf, 1966).

Measuring Retention and Transfer of Learning

The posttest assesses the level of learning that was attained in the experimental and control groups or conditions. In so doing, it addresses the relative effectiveness of learning through pre-instruction testing followed by studying versus studying without such testing. As described next, memory for *directly tested materials* is almost always assessed on the posttest. In some instances, the posttest is also used to assess the learning of materials that were not directly tested—that is, *untested materials* that were either *presented for study* or *previously unrepresented*. Doing so addresses the extent of any pre-instruction testing benefit for those materials.

Memory for Directly Tested Materials

The posttest typically features questions that are identical or nearly identical to those that were used during pre-instruction testing. These questions measure *retention* or *memory* for the content targeted by the practice questions. For example, in Carpenter and Toftness (2017), participants attempted the practice question, “How many families

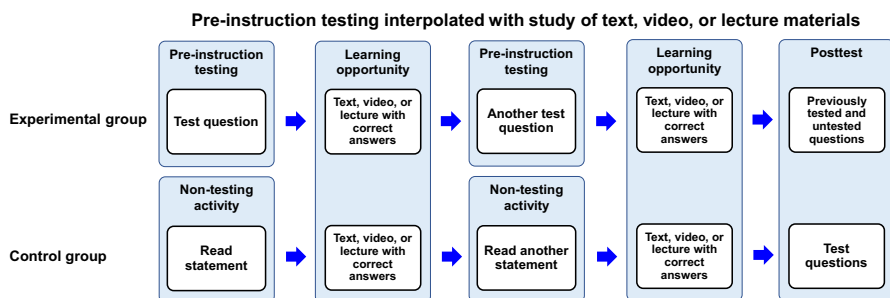


Fig. 2 Example implementation of *interpolated pre-instruction testing* versus a non-testing control condition. Participants engage in practice testing or a non-testing control activity prior to each of several segments of a text passage, video, or lecture. Learning is then assessed on a subsequent posttest

originally settled on the island of Rapa Nui?”, viewed an educational video about the history of Easter Island (which is also called Rapa Nui), and received the exact same question on a subsequent posttest. If pre-instruction testing was effective at enhancing memory for tested content, then performance on that posttest question should have been higher than among participants that had not practiced with that question (i.e., those in the control group), as was observed.

As a related matter, if prior knowledge of the target materials prior to the experiment is low, as is commonly the case, then performance during practice testing should also be low. Any improvement from practice test to posttest on the same questions, therefore, would indicate that learning occurred. Indeed, such improvements are commonly the case in the experimental group, indicating that the correct answers were successfully learned after practice testing had occurred.

Transfer to Untested Materials Presented for Study

In some cases, the posttest includes questions that participants had not previously seen or attempted. These questions commonly assess knowledge of content that was not directly targeted by practice questions, but was still available for participants to study during the subsequent learning opportunity. For example, in Carpenter and Toftness (2017), the posttest also included the question, “What was the approximate population of Rapa Nui from 1722 to the 1860s?”, which the aforementioned participants had not seen before. The correct answer to that question, however, was presented in one of the video segments that those participants did see. For the experimental group, such “new” questions measured whether the effects of practice testing generalize, or *transfer*, to untested but previously presented materials, as opposed to “old” questions that measured memory for previously tested content.

It is important to note that the distinction between “new” (retention) and “old” questions (transfer) on the posttest applies primarily to the experimental group. For the control group, all posttest questions are usually entirely “new.” For ease of comparison, however, posttest questions administered to the control group may be yoked (that is, given “old” and “new” classifications) to questions administered to the experimental group. In other cases, however, the control group’s performance on the posttest is reported in its entirety without separation into “old” and “new” categories.

The inclusion of posttest questions that assess knowledge of untested materials presented for study is common in studies where the correct answers to pre-instruction test questions are discovered by searching through text, video, or lecture materials. In studies where the correct answers are learned solely through correct answer feedback, which usually involve simpler materials such as word pairs or trivia facts, assessing knowledge of untested materials is rare (for exceptions, see Hays et al., 2013, Experiment 3; Pan et al., 2019).

Transfer to Untested and Previously Unpresented Materials

In several studies, the posttest features questions that address other forms of transfer, including drawing inferences (e.g., St. Hilaire et al., 2019) and classifying new

exemplars (e.g., Sana, Yan, et al., 2020a, 2020b). These questions might still be categorized as “new” given that they were not previously presented to participants. The correct answers to these questions, however, cannot be lifted directly from materials presented earlier in the study. Rather, answering these questions correctly requires drawing a conclusion or deriving knowledge that was not directly stated in the materials (in the case of inferences) or applying prior learning to newly presented information (in the case of classifying new exemplars). Research in the literature involving transfer of learning to new scenarios, situations, and other contexts (for related taxonomies of transfer, see Barnett & Ceci, 2002; Pan & Rickard, 2018) is currently in its infancy.

Specific Versus General Benefits of Pre-instruction Testing

When the posttest measures learning of directly tested and untested content—that is, by including “old” and “new” questions—an advantage of pre-instruction testing (i.e., a prequestioning effect or a pretesting effect) may occur across two common scenarios. In the first scenario, the experimental group outperforms the control group on both types of questions. If so, that would constitute a *general benefit* of prior testing wherein the learning of directly tested and untested materials is enhanced. In the second scenario, the experimental group outperforms the control group but only on questions that previously appeared on the prior practice test (i.e., “old” questions). If so, that would constitute a *specific benefit* of prior testing wherein only the learning of directly tested materials is enhanced (Carpenter et al., 2023; see also Anderson & Biddle, 1975; Hartley & Davies, 1976). Whether pre-instruction testing commonly yields specific or general benefits is a major question in this literature.

Effects of Pre-instruction Testing on Learning

The following sections summarize the evidence for the effects of pre-instruction testing across a host of circumstances. First, effects on directly tested and untested information are addressed. Second, studies conducted in classrooms and other authentic educational settings, which constitute exceptions to the standard practice of relying on laboratory settings or online experiment platforms, are summarized. Next, effects involving different test formats are considered. Finally, studies that investigated different temporal intervals between practice testing and subsequent learning opportunities, as well as studies that featured a posttest that was administered after an extended retention interval, are discussed (Table 1).

Except where noted, all of the studies reviewed here feature a separate non-testing control group or condition against which the efficacy of pre-instruction testing was determined (cf. de Lima & Jaeger, 2020; Geller et al., 2017; McDaniel et al., 2011; Overoye et al., 2021). In such groups or conditions, learning without any practice testing occurred in a separate group of participants or a separate set of materials. Moreover, retrieval practice was not implemented prior to the posttest (cf. Welhaf et al., 2022). Finally, in most studies, the participants were undergraduate students

Table 1 Pre-instruction testing effects on memory across short versus extended retention intervals

Learning materials	Retention interval	Prequestioning or pretesting effect	Reference
Lecture content	2–8 weeks	$d=0.40$	Soderstrom and Bjork (2023)
Paired associate words	5 min	$d=0.38$	Kornell et al. (2009), experiment 4
Paired associate words	5 min	$d=1.15$	Yan et al. (2014), experiment 1
Paired associate words	24 h	$d=0.94$	Kornell et al. (2009), experiment 5
Paired associate words	24 h	$d=0.44$	Kornell et al. (2009), experiment 6
Paired associate words	48 h	$d=0.64$	Yan et al. (2014), experiment 3
Triple associate words	5 min	$d=0.89$	Pan et al. (2019), experiment 2
Triple associate words	48 h	$d=0.40$	Pan et al. (2019), experiment 1
Trivia facts	3 min	$d=0.56$	Kornell (2014), experiment 2
Trivia facts	24 h	$d=0.86$	Kornell (2014), experiment 3a
Text passage	No delay	$d=1.10$	Richland et al. (2009), experiment 1
Text passage	7 days	$d=0.45$	Richland et al. (2009), experiment 4
Text passage	5 min	$d=1.20$	Pan and Sana (2021), experiment 3, cued recall condition
Text passage	5 min	$d=1.44$	Pan and Sana (2021), experiment 3, cued recall with feedback condition
Text passage	48 h	$d=0.78$	Pan and Sana (2021), experiment 4, cued recall condition
Text passage	48 h	$d=2.04$	Pan and Sana (2021), experiment 4, cued recall with feedback condition
Text passage	1 min	$d=1.27$	Kliegl et al. (2022), experiment 2
Text passage	30 min	$d=1.84$	Kliegl et al. (2022), experiment 2
Text passage	7 days	$d=2.68$	Kliegl et al. (2022), experiment 2

or individuals of similar age. Studies involving other age groups are specifically noted as such.

Effects on Directly Tested Information

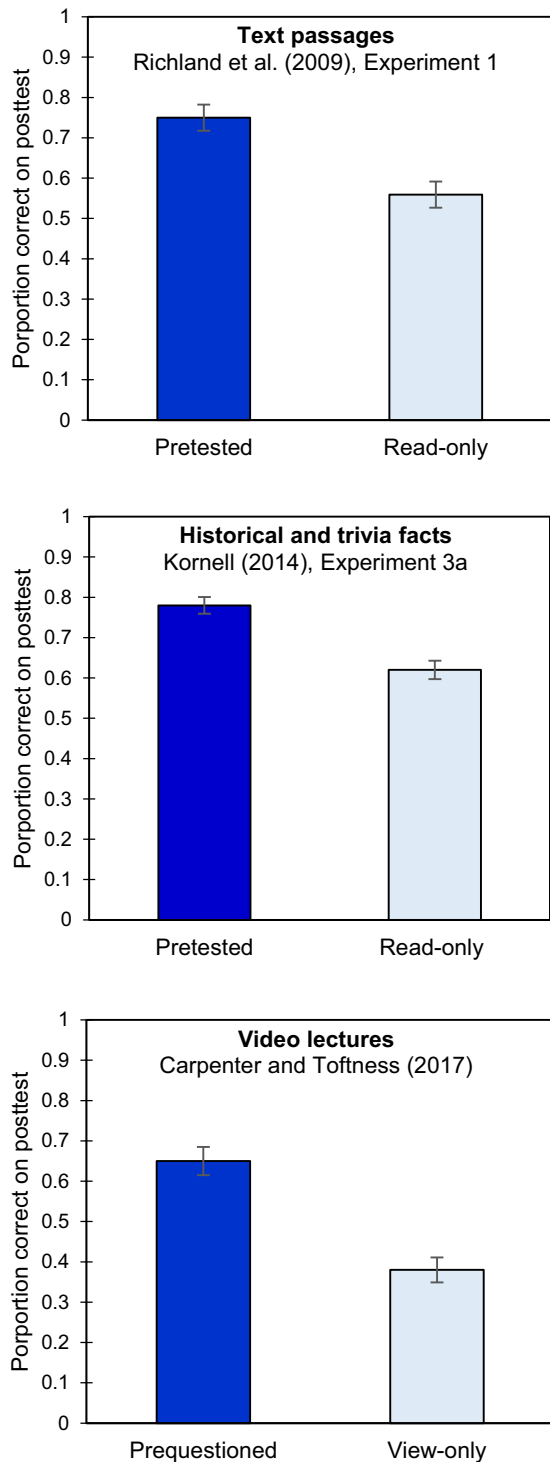
Nearly all studies of pre-instruction testing have addressed effects on directly tested information. The following studies reported data for posttest questions targeting memory of such information. As described next, these studies involved text or video materials.

Text Materials

Pre-instruction testing has been shown to enhance memory for directly tested information drawn from text passages. A prominent example is that of Richland et al. (2009), wherein participants read a two-page passage about a visual disorder, cerebral achromatopsia. In the experimental group, participants spent two minutes attempting five practice questions drawn from the passage (e.g., “What is total color blindness caused by brain damage called?”), then read the passage for eight minutes. In the control group, participants spent the entire ten-minute period reading the passage. Prior knowledge of the text passage was low; in the experimental group, participants typically answered three-quarters or more of the practice questions incorrectly. When the test questions were re-administered on an immediate posttest (experiments 1–3) or on a one-week delayed posttest (experiment 4), however, participants in the experimental group consistently outperformed the control group (see Fig. 3, upper panel). These results compellingly demonstrate the capacity of pre-instruction testing to enhance memory for the correct answers when those answers are subsequently presented in a text passage. Multiple studies have since reported similar results with text passages involving such topics as biographies, geography, history, oceanography, physics, science fiction, statistics, and weather (e.g., Hausman & Rhodes, 2018, Experiment 1; James & Storm, 2019, Experiments 1–4; Kliegl et al., 2022, Experiment 2; Little & Bjork, 2016; Sana et al., 2020a, 2020b; Sana & Carpenter, 2023, Experiment 1; St. Hilaire et al., 2019; St. Hilaire & Carpenter, 2020).

Enhanced memory for directly tested information drawn from text statements (e.g., trivia or other types of facts) has also been demonstrated following pre-instruction testing. For instance, in Kornell et al., (2009; experiment 1), participants learned a series of 20 fictional trivia facts, all presented one at a time in question-and-answer form (e.g., “Q: What treaty ended the Calumet War? A: Harris”). For half of the facts, the question and answer were presented simultaneously for 5 s, during which participants read both; for the other half, participants first attempted to answer the question for 8 s before being presented with the correct answer for 5 s (i.e., correct answer feedback). During practice testing, participants were unable to answer any questions correctly. On an immediate posttest, however, participants better remembered the answers to previously tested, as opposed to untested, trivia facts. Improved memory for factual content learned via pre-instruction testing with

Fig. 3 Example pre-instruction testing effects on memory for text passages, trivia facts, and video lectures. Results reproduced from Richland et al., (2009, experiment 1), Kornell (2014, experiment 3a), and Carpenter and Toftness (2017)



correct answer feedback has since been replicated across multiple other studies (e.g., Kornell, 2014; see Fig. 3, middle panel), including at a delay of up to 48 h (Vaughn et al., 2017), with use of search engines to obtain correct answers (e.g., Storm et al., 2022), and with various types of facts (e.g., obscure trivia, fictional facts, historical facts).

A large portion of the studies in this literature has used paired associate words (e.g., *doctor-nurse*) as learning materials. A characteristic finding has been that taking practice tests on the target word of a pair that had not previously been learned (e.g., *doctor-???*), followed by immediate correct answer feedback, can enhance memory for the target word. That result, however, is predicated on the two words in a given pair having at least a weak semantic association (e.g., Hays et al., 2013; Kliegl et al., 2022; Kornell, 2014; Kornell et al., 2009; Zawadzka et al., 2023; and others). If that semantic association is absent (e.g., *door-shoe*), then no advantage of prior testing may occur if the posttest involves cued recall (e.g., Grimaldi & Karpicke, 2012; Huelser & Metcalfe, 2012; Knight et al., 2012). Benefits of pre-instruction testing for weakly associated word pairs have since been replicated with children of kindergarten and early elementary school age, but not preschool age (Carneiro et al., 2018). Further studies have shown benefits for the case of word triplets (e.g., *gift, rose, wine*), indicating that the benefit of pre-instruction testing for cue-and-target materials is not limited to paired associates (Pan et al., 2019; see also Metcalfe & Huelser, 2020).

Other studies have focused on learning the definitions of foreign language vocabulary or unfamiliar words in one's native language. For instance, in Potts and Shanks (2014), participants learned Euskara-English translations (e.g., *urmael—pond*) or the definitions of obscure English words (e.g., *frampold—quarrelsome*). Each translation or definition was learned in multiple ways, of which two were as follows: The word and its definition (or translation) were read in their entirety, or participants guessed the definition prior to viewing the correct answer (i.e., practice testing with feedback). On a multiple-choice posttest, the definitions of words that had previously been tested were better recognized than those that had been read. Improved recognition of definitions following pre-instruction testing has been replicated in several other studies (e.g., Potts et al., 2019; Seabrooke et al., 2019a, 2019b, 2021a, 2021b). When the posttest involves cued recall, however, a benefit of pre-instruction testing has typically not been observed (Butowska et al., 2021; Seabrooke et al., 2019a, 2019b). That result bears some similarity to the aforementioned findings involving paired associates that lack a semantic association (although, see Carpenter et al., 2012 for an exception).

Video Materials

Studies of pre-instruction testing with video-based materials have also reported memory benefits. In Carpenter and Toftness (2017), for instance, participants viewed a seven-minute educational video about Easter Island. In the experimental condition, six practice questions were interpolated amongst two-minute segments of video (two questions before each segment), whereas in the control condition, no practice questions were administered. On an immediate posttest, participants in the

experimental condition exhibited better memory for directly tested information relative to the control condition (see Fig. 3, bottom panel). Benefits of pre-instruction testing for directly tested video content have since been repeatedly observed (e.g., James & Storm, 2019; Pan et al., 2020a, 2020b; Sana & Carpenter, 2023; St. Hilaire & Carpenter, 2020; Toftness et al., 2018). The length of the videos for which such benefits have been demonstrated has ranged from 5 to 30 min (e.g., James & Storm, 2019; St. Hilaire & Carpenter, 2020), and the addressed topics have included psychology, statistics, and history. Administering all practice questions prior to the entire video (e.g., Toftness et al., 2018) and interpolating questions throughout segments of video (e.g., Pan et al., 2020a, 2020b) have both been effective at enhancing memory.

Effects on Untested Information

A subset of the literature has investigated the effects of pre-instruction testing on the learning of information that is not directly tested. Studies addressing this issue can be classified into two categories based on the nature of the untested materials. In the more common category, untested materials were presented for participants to study during the experiment. In the other category, untested materials were not directly available for study and participants had to draw inferences or apply what they had learned to classify new information on the posttest.

Untested Materials Presented for Study

Unlike with directly tested materials, the evidence involving untested materials presented for study is highly inconsistent. Those results are summarized next. Possible explanations are considered later in this review.

Among studies involving text passages, results have been mixed. For instance, in Richland et al. (2009), enhanced learning of untested information presented for study (i.e., available in the provided text passage) was observed in one experiment but not in any of the other four similarly designed experiments. That result contrasts with the benefits of pre-instruction testing for directly tested information that was observed in all experiments. No benefit or decrement for untested information, relative to a non-testing control condition, has been observed in several other studies involving text passages (e.g., James & Storm, 2019, experiments 1–4; St. Hilaire et al., 2019, experiment 1; St. Hilaire & Carpenter, 2020). Exceptions include Little and Bjork (2016), Sana and Carpenter (2023), and St. Hilaire et al., (2019, experiment 2), wherein benefits of pre-instruction testing were observed for both tested and untested materials.

Among studies involving simpler materials, namely word pairs or triplets, two studies have addressed whether guessing a target word and studying the correct answer will yield improved memory for the cue word or words. Across two experiments, Pan et al. (2019) found that such testing did enhance learning of untested cues from word triplets (e.g., guessing the answer to *gift, rose, ???* which is *wine*, will enhance the ability to correctly answer *???, rose, wine*). Hays et al., (2013,

experiment 3) observed a similar result for the case of word pairs. These findings suggest that pre-instruction testing can foster learning that is, to a degree, transferable from the answers (to simple questions) to aspects of the questions themselves.

Among studies involving video materials, results have also been mixed. Whereas James and Storm (2019, experiment 5), Pan et al., (2020a, 2020b, experiment 1), St. Hilaire and Carpenter (2020), and Toftness et al. (2018) observed no benefits of pre-instruction testing for untested materials presented in videos, Carpenter and Toftness (2017), Pan et al., (2020a, 2020b, experiment 2), and Sana and Carpenter (2023, experiment 1) did observe such benefits. As with studies involving text passages, however, a benefit for directly tested materials was observed in all of those studies.

Untested and Previously Unpresented Materials

Evidence regarding the effects of pre-instruction testing on the ability to draw inferences from a text passage (e.g., after reading a text passage about different types of brakes, inferring the primary difference between mechanical brakes and hydraulic brakes) is relatively limited. Of the two studies that have addressed the issue to date, Hausman and Rhodes (2018) found no evidence of improved inferences following pre-instruction testing on expository text passages, whereas St. Hilaire et al. (2019) observed improved performance on inference questions but only when participants were required to make inferences during practice testing *and* search for the answers to those inference questions while reading. Based on those limited results, pre-instruction testing does not guarantee improved inference ability, although such improvements are possible.

At least one study to date has investigated the efficacy of pre-instruction testing for learning to classify new exemplars of previously studied categories. In Sana, Yan et al., (2020a, 2020b, experiment 1), participants learned to recognize three types of statistical procedures (chi-squared test, Kruskal–Wallis test, and Wilcoxon signed-rank test) and, for each procedure, studied example scenarios or attempted four practice questions prior to studying the scenarios. On a 5-min delayed classification test (wherein participants were presented with new scenarios and had to correctly identify the appropriate statistical procedure that should be used), participants that engaged in practice testing outperformed those that had not. A second experiment observed the same results even when the control condition studied the practice questions and their answers beforehand. This result suggests that pre-instruction testing has the potential to enhance classification skills.

Effects in Authentic Educational Settings

At least four studies conducted in authentic educational settings have compared pre-instruction testing against a separate non-testing control group or condition. In Beckman (2008), undergraduate students in two sections of an aerospace course completed three units of the course; in one section, each unit was prefaced by a

pretest targeting the learning objectives for that course, whereas in the other section, no pretests were administered. Each pretest was scored before being returned to the student. On high-stakes tests conducted at the end of each unit, students that had engaged in pretesting scored an average of 9–12% higher.

Benefits of pre-instruction classroom testing as measured on subsequent post-tests but using more robust experimental designs (i.e., random assignment of students to groups or conditions) have also been reported by Carpenter et al. (2018) and Soderstrom and Bjork (2023). Both studies entailed the administration of practice quizzes at the start of lectures. Carpenter et al. (2018) had students in an undergraduate introductory psychology laboratory course attempt or not attempt a single prequestion individually on laptop computers at the start of a lecture period. No correct answer feedback was provided. At the end of the lecture period, all students completed a two-question posttest on which benefit of prior testing was observed (approximately 7% better performance in the prequestion condition for previously tested content). A week later, however, when the posttest was administered a second time, test performance was numerically, but not significantly, higher for students who had engaged in prequestioning. These results suggest that pre-instruction testing can enhance learning at least over the short term, but as implemented by Carpenter et al., such testing did not reliably improve long-term learning over that due to a single posttest (i.e., retrieval practice).

Soderstrom and Bjork (2023) administered multiple-choice pretests via pen and paper at the start of several lectures during a 10-week undergraduate psychology research methods course. No correct answer feedback was provided. On a high-stakes final exam administered at the end of the course, students performed 8–9% better on questions addressing content that had been previously pretested. That benefit was observed both for final exam questions that were identical to those that were used during pretesting, as well as final exam questions that targeted related but different information from that which was pretested. In this case, the results suggest that pre-instruction testing can yield learning improvements that persist for at least several weeks.

Janelli and Lipnevich (2021) investigated the effects of pre-instruction testing in a Massive Open Online Course (MOOC) about climate change. Students enrolled in the MOOC took a pretest at the start of each of five course modules or did not take any pretests at all. Depending on condition assignment, students received various forms of feedback or no feedback at all. The results showed no overall benefit of pretesting or feedback on subsequent exam performance, but with an important caveat: Students that had received pretests were more likely to drop out of the MOOC as it progressed, but among the students that persisted until the end of the MOOC, exam performance was better if they had received pretests. The authors concluded that pretesting in MOOCs has both negative and positive effects—that is, it can not only reduce persistence, perhaps by affecting motivation, but also enhance learning among the students that persevere.

Lima and Jaeger (2020) had fourth and fifth grade students read an age-appropriate encyclopedic text passage with key words partially removed and replaced with blanks. Students guessed the missing words, then re-read the passage with all words intact. When presented with the text passage again after one week, this time with

previously tested keywords and other keywords removed, students were better able to recall the words that had been tested. That result highlights the potential of such testing to enhance learning at the elementary school level.

In a recent study of pre-instruction testing in clinical settings, Willis et al. (2020) had medical students learn a laparoscopic procedure in a simulation laboratory by attempting the procedure before watching an instructional video (the “struggle first” group) or watching the video prior to attempting the procedure (the “instruction-first” group). On an immediate posttest, participants in the struggle-first group were able to complete the procedure more quickly and with fewer errors. The authors attributed that result to a pre-instruction testing effect for procedural skills (alternatively, invention activities or productive failure may have played a role; both strategies are discussed later in this review).

Studies conducted in authentic educational settings but without a separate non-testing control group or condition—in which pre-instruction testing did not occur—provide further evidence. In McDaniel et al., (2011; in a middle school science course; see also Geller et al., 2017, in an undergraduate chemical engineering course), students completed practice questions prior to or at the beginning of lecture sessions. When the questions were re-presented on subsequent posttests, performance was improved. That result indicates that learning occurred following the administration of the practice questions. The conclusions that can be drawn from that study, however, is limited due to the lack of control groups.

Effects Involving Different Test Formats

Effects of pre-instruction testing have been investigated using practice tests and posttests in a variety of test formats. With few exceptions, benefits of pre-instruction testing have been observed regardless of test format.

Practice Test Format

Practice tests in cued recall or short answer (e.g., Carpenter et al., 2018; Richland et al., 2009), fill-in-the-blank (e.g., Richland et al., 2009), and multiple-choice format (e.g., Toftness et al., 2018) have all been shown to elicit pre-instruction testing effects. Table 2 features examples of each test format. Pan and Sana (2021; experiments 2–4), for example, manipulated whether participants received multiple-choice or cued recall pretests; the format of the practice test did not substantially influence the magnitude of the resulting pre-instruction testing effects.

In some circumstances, however, the test format used for practice testing can be influential. For instance, in Little and Bjork (2016), participants took multiple-choice or cued recall pretests on information drawn from text passages about the planet Saturn, Yellowstone National Park, and stimulant drugs. The multiple-choice questions featured competitive lures, which were incorrect answer alternatives referencing information in the passage (e.g., “Q: What is the tallest geyser in Yellowstone National Park – Old Faithful, Steamboat Geyser, Castle Geyser, or Daisy Geyser? A: Steamboat Geyser.”), whereas the cued recall questions, by their very nature,

Table 2 Test formats used in pre-instruction testing studies

Learning materials	Practice test		Posttest		Reference
	Format	Example	Format	Examples Question on tested materials	Question on untested materials
Paired associate words	Cued recall	<i>Frog-???</i>	Cued recall	Identical to practice test question	???-pond Hays et al. (2013)
Paired associate words	Cued recall	<i>Gadoid (What do you think this word means?)</i>	Recognition	<i>Gadoid (Did you see this word before? Yes or No)*</i>	Not applicable Seabrooke et al., (2019a, 2019b)
Text passage	Cued recall (short answer)	<i>What is total color blindness caused by brain damage called?</i>	Cued recall (short answer)	Identical to practice test question	<i>How does Mr. I distinguish red and green traffic lights?</i> Richland et al. (2009)
Text passage	Fill-in-the-blank	<i>V4 specializes in responding to ____</i>	Fill-in-the-blank	Identical to practice test question	<i>Cells in V1 respond specifically to ____</i> Richland et al. (2009)
Text passage	Multiple-choice	<i>The first method to scientifically study autobiographical memory was called:</i> a. Diary technique b. Word cueing technique c. Interview technique d. Reminiscence technique	Multiple-choice	Identical to practice test question	<i>The reminiscence bump is the phenomenon of particularly good memory for:</i> a. Early life events b. Events during school years c. Young adulthood d. Events that happened most recently Toftness et al. (2018)

Example questions are paraphrased, where necessary, for brevity

*Example trial with foil word not shown

did not include any lures. The competitive lures (e.g., “Castle Geyser”) would ultimately be the correct answers to transfer questions on a cued recall posttest. On the posttest, whereas a pre-instruction testing effect for directly tested information was observed regardless of practice test format, a transfer effect for untested information was only observed following multiple-choice pretesting. A crucial factor for successful transfer in that case appears to have been prior exposure to untested information, in a practice testing context, via the lures that were present only in the multiple-choice pretests. Simply viewing the untested information in a non-testing context (i.e., studying) did not yield the same degree of learning (Little & Bjork, 2016). More broadly, these results suggest that not only is test format potentially influential for pre-instruction testing effects, but so is the manner with which a given format is implemented (for related discussions in the adjunct questions literature, including possible effects on the learning of directly tested and untested information, see Anderson & Biddle, 1975; Hamaker, 1986).

Posttest Format

Pre-instruction testing effects have also been observed on posttests involving recall or short answer (e.g., Carpenter et al., 2018; Richland et al., 2009), fill-in-the-blank (e.g., Richland et al., 2009), multiple-choice (e.g., Toftness et al., 2018), and recognition format (e.g., Potts et al., 2019). Table 2 features examples of each test format. These results suggest that the posttest format may not be crucially important to detect benefits of pre-instruction testing. An exception, however, involves the cases of semantically-unrelated paired associate words and unfamiliar foreign language translations. In studies using such materials wherein both cued recall and recognition posttests have been administered, a benefit of prior testing has only been observed in the case of recognition tests (e.g., Seabrooke et al., 2021a, 2021b). That result has yet to be fully explained, although one account posits that pretesting enhances attention to correct answer feedback regardless of the degree with which a cue and target are semantically associated.

Effects Across Different Timing and Retention Intervals

Most studies in the literature have featured little-to-no interval of time between pre-instruction testing and subsequent learning opportunities. The retention interval prior to the posttest is, in many cases, also as short as a few minutes or none. Exceptions to these patterns, which provide insights into implementation issues and the durability of observed effects, are discussed next.

Interval Between Practice Testing and Subsequent Learning Opportunities

In the case of paired associates, immediate correct answer feedback is often necessary for improved memory of pretested target words to occur. In Grimaldi and Karpicke (2012), Vaughn and Rawson (2012), and Hays et al. (2013), the correct answer was displayed immediately after participants had entered their guess for

a given word pair or withheld until guesses had been made for multiple or all to-be-learned word pairs. In all three studies, only the provision of immediate correct answer feedback yielded a pretesting effect relative to a non-testing control condition on a subsequent cued recall posttest and delaying feedback nullified that advantage. These results contrast with findings involving text passages and other materials, wherein participants may not discover the correct answers until a learning opportunity that is at least several minutes later and after other information has been encountered in the intervening time period (e.g., Richland et al., 2009). (An alternative explanation suggested by a reviewer, however, is that in the aforementioned studies with paired associates, the lack of a pretesting effect in the delayed feedback case may be due to a spacing effect in the non-testing control condition).

To address the discrepancy between studies of paired associates versus other materials, Kornell (2014) investigated the efficacy of delayed feedback following pre-instruction testing with materials that are more semantically detailed than paired associates, namely trivia facts. Across three experiments, pre-instruction testing enhanced learning regardless of whether feedback was withheld until other intervening items had been encountered or even 24 h after participants had finished practice testing on all to-be-learned items (see Fig. 3, middle panel, for a depiction of the results). In fact, when the timing of feedback was manipulated within a single experiment, there was no difference in pre-instruction testing effect magnitude between immediate versus delayed feedback conditions (Kornell, 2014, experiment 2). These results suggest that benefits of pre-instruction testing for non-paired associate materials (i.e., more semantically detailed or complex materials as in the case of facts and text passages) can manifest even in cases where learning of the correct answers is delayed.

Retention Interval Prior to the Posttest

Among studies that have assessed learning after a retention interval of at least 24 h, there continues to be evidence of pre-instruction testing effects. As illustrated in Table 1, studies involving paired or triple associate words (e.g., Kornell et al., 2009; Pan et al., 2019; Yan et al., 2014; see also Potts et al., 2019), trivia facts (e.g., Kornell, 2014), and text passages (e.g., Kliegl et al., 2022; Little & Bjork, 2016) have all shown memory improvements that persist for at least 24 h, and in the case of text passages, at least 7 days. The reported memory improvements after at least 24 h, in effect size terms, range from Cohen's $d=0.44$ to well over 2.0 (for discussion of interpreting effect size magnitude, see Kraft, 2020).

Few of those studies, however, have addressed transfer to untested materials seen during study. In the case of triple associate words, Pan et al. (2019) reported successful transfer to cue words at a 48-h retention interval. Using text passages, Richland et al. (2009) reported no evidence of transfer at extended retention intervals, whereas Little and Bjork (2016) reported successful transfer at a 48-h retention interval. In all three studies, the transfer results at a delay mirrored patterns observed with shorter retention intervals.

Kliegl et al. (2022) investigated the efficacy of pre-instruction testing for text passages across three retention intervals—1 min, 30 min, and 1 week—within a single

experiment. The magnitude of the pre-instruction testing effect grew with retention interval and was largest at a 1-week interval. That result led the authors to conclude that the benefits of pre-instruction testing manifest more fully after longer retention intervals. If so, that would suggest that the magnitude of the effects that have been reported in the current literature, most involving very brief or immediate retention intervals, may be understating the potency of pre-instruction testing for memory.

Additionally, pre-instruction testing effects have been observed in three classroom studies (Beckman, 2008; Janelli & Lipnevich, 2021; Soderstrom & Bjork, 2023), wherein the duration between practice testing and a subsequent, high-stakes test ranged from a few days to many weeks. Those results further reinforce the conclusion that the benefits of pre-instruction testing persist across educationally meaningful retention intervals.

Metacognitive Considerations

Many learners do not recognize the benefits of pre-instruction testing even after experiencing and benefiting from it. For example, Huelser and Metcalfe (2012; experiment 2) had participants learn semantically related and unrelated word pairs via pretesting and reading, take a posttest on those pairs, and then rank order the effectiveness of the techniques that they had used. Although pretested word pairs were the best remembered (in the case of semantically related words), participants nevertheless ranked pretesting as the least effective. Similarly, when asked to make predictions about the likelihood of future recall (i.e., judgments of learning or JOLs) of word pairs that were pretested or read, participants tend to give lower predictions to pretested pairs (e.g., Pan & Rivers, 2023; Potts & Shanks, 2014; Yang et al., 2017; Zawadzka & Hanczakowski, 2019). Those predictions belied subsequent post-test results.

The inability to recognize the benefits of pre-instruction testing, which constitutes a potent metacognitive illusion (Bjork et al., 2013), may stem from a tendency to eschew (or at least not favor) pre-instruction testing and other error-prone strategies. In a survey by Yang et al., (2017; experiment 3), for example, when presented with a hypothetical scenario involving the use of pretesting or reading to learn word pairs, 78% of respondents rated reading as more effective. In another survey by Pan et al., (2020a, 2020b), when presented with a hypothetical scenario that entailed learning an academic subject via a guess-and-study approach or a study-only approach, 56% of respondents rated the former and 44% rated the latter as more likely to be effective. Although more favorable towards pre-instruction testing, that result is far from a strong endorsement. In the same survey, 81% of respondents also endorsed the importance of avoiding errors during the learning process, which is a hallmark of pre-instruction testing. Overall, survey data suggest that learners often hold beliefs about learning that may predispose them to favor non-testing methods instead (for related discussions, see Huelser & Metcalfe, 2012; Metcalfe, 2017).

Two studies have investigated ways to promote metacognitive awareness of the benefits of pre-instruction testing. In Yang et al., (2017; experiment 4), participants read a brief explanation about the benefits of “errorful generation” prior to learning

word pairs via pretesting and reading; doing so resulted in item-level JOLs that were less biased towards reading (but still favoring it) and JOLs over all items that favored pretesting. In Pan and Rivers (2023; experiments 3–5), participants completed two rounds of pretesting and reading on word pairs, making JOLs, and taking a posttest; after receiving performance feedback on the first-round posttest (i.e., how they scored on pretested versus read items), participants' second-round JOLs favored pretesting outright. These results suggest that providing learners with explanations or other information about the benefits of pre-instruction testing can help correct meta-cognitive misperceptions.

Theoretical Perspectives

Researchers have proposed a variety of theoretical accounts to explain pre-instruction testing effects. Some accounts are relatively general, reflecting the incipient state of theoretical understanding in this literature, whereas others are more detailed and suggest specific cognitive mechanisms. Prominent and recent perspectives in the pre-instruction testing literature are summarized next. A list of those accounts, which vary in their emphasis on processes occurring during the event of taking a practice test and/or during subsequent learning opportunities, is presented in Table 3. Additional theoretical perspectives involving potentially related learning strategies and phenomena, including test-potentiated new learning and retrieval practice, are addressed later in this review.

Pre-instruction Testing with Immediate Correct Answer Feedback

At least four theoretical accounts have been proposed to explain pre-instruction testing effects on paired associate words with immediate correct answer feedback (for reviews, see Mera et al., 2022; Metcalfe, 2017). One such account, the *mediator* or mediation account, involves the generation of mediators (i.e., words that link cues and targets) during practice testing. This account was previously developed as an explanation for the retrieval practice effect (Carpenter, 2011; Carpenter & Yeung, 2017; Pyc & Rawson, 2010; see also Soraci et al., 1999). In the case of pre-instruction testing, the mediator account proposes that incorrect responses made during practice testing take the form of mediator words (e.g., for the word pair *mother–child*, the mediator word *father*). These mediators, in turn, support improved recall of the correct answers on a posttest. On such posttests, the mediator words are presumably recalled and used as mental “stepping stones” to the correct answer (Huelser & Metcalfe, 2012; Kornell et al., 2009; cf. Metcalfe & Huelser, 2020). The mediator account is broadly consistent with the idea of multiple potential retrieval routes to a correct answer (Kornell et al., 2009).

The *search set* account posits that the attempt to answer a practice test question activates a set of potential candidate answers (i.e., the search set), one of which is the correct answer. When the correct answer is presented in the form of immediate feedback, the encoding of the correct answer (which is presumably already activated

Table 3 Theoretical accounts of pre-instruction testing effects

Account	Description	Reference(s)
Mediator hypothesis	Practice testing causes the generation of words that link cues and targets	Huelser and Metcalfe (2012)
Search set account	Practice testing activates a set of potential candidate answers, one of which is the correct answer; encoding of that answer is enhanced when feedback is presented	Grimaldi and Karpicke (2012)
Recursive reminding account	Experience of generating an error and learning the correct answer are encoded in the same episodic event	Jacoby and Wahlheim (2013)
Prediction error account	Discrepancy between erroneous response and correct answer triggers an error signal	Kang et al. (2011)
Mathemagenic hypothesis	Practice testing leads to a general increase in interest, curiosity, or attention to/for subsequently presented materials	Rothkopf (1966) and various (e.g., Pan et al., 2020a, 2020b)
Answer search	Learners engage in information seeking behaviors to find the answers to specific practice questions	Lewis and Mensink (2012); St. Hilaire and Carpenter (2020)
Attentional window hypothesis	Practice testing leads to increased attention that persists until the correct answers to practice questions are discovered	Sana and Carpenter (2023)

in memory) is enhanced (Grimaldi & Karpicke, 2012). By this account, the correct answer must be presented immediately after a practice test trial, while the search set is still active, in order for improved learning to result. Both the search set and mediator accounts are consistent with spreading activation accounts of semantic memory (Collins & Loftus, 1975).

The *recursive reminding* account, which is adapted from theoretical explanations of change detection in multi-list, paired associate learning studies (Jacoby & Wahlheim, 2013; Wahlheim & Jacoby, 2013), proposes that the experience of generating an error and learning the correct answer through feedback are encoded in the same episodic event (Mera et al., 2022; Metcalfe, 2017). On a posttest, when individuals recall the context in which they generated an error, they should also be able to remember the correct answer given that it was encoded in the same episodic event. By this account, it is recursive reminding that gives rise to improved posttest performance.

The *prediction error* account suggests that the discrepancy between an erroneous response and subsequent correct answer feedback triggers an error signal, which in turn enhances attention and subsequent learning (Kang et al., 2011; Metcalfe, 2017; see also Brod, 2021). This account draws from the cognitive neuroscience research on learning processes involved in detecting and correcting errors (e.g., Ergo et al., 2020; Wang & Yang, 2023). It is reminiscent of the *hypercorrection effect*, which is the finding that errors made with high confidence are more likely to be corrected on a posttest (Butterfield & Metcalfe, 2001) and possibly due to the greater level of surprise that is felt upon the discovery of such errors (Butterfield & Mangels, 2003). (Owing to differences in experimental procedure, however, the confidence that a learner may have in their answers to a pretest question, and the corresponding level of surprise that they may experience upon learning the correct answer, may not always correspond to the conditions that give rise to the hypercorrection effect.)

The foregoing accounts potentially could be applied to scenarios involving materials other than paired associate words and without immediate correct answer feedback. Doing so, however, may encounter some difficulties. For example, it is unclear whether mediators would be generated when the learning materials are not individual cues and targets. Moreover, it seems unlikely that a search set would remain active in the case of delayed feedback (cf. Kornell, 2014). The recursive reminding and prediction error accounts, in contrast, appear to be applicable to a wider range of stimuli and learning scenarios without need for substantial modification.

Pre-instruction Testing with Text Passage, Video, and Lecture Materials

Several theoretical accounts have been applied to pre-instruction testing effects for text passage, video, and lecture materials, as well as cases where immediate correct answer feedback is not provided (for discussion see Carpenter et al., 2023). These accounts, which are often descriptive in nature, are largely consistent with the *mathemagenic hypothesis* from the adjunct questions literature (Rothkopf, 1966; Rothkopf & Bisbicos, 1967; see also Anderson & Biddle, 1975; Bull, 1973). The mathemagenic hypothesis posits that the experience of

answering adjunct questions causes learners to process subsequently presented text materials more thoroughly. Among such accounts, it has been suggested that pre-instruction testing affects interest (Little & Bjork, 2011), curiosity (Geller et al., 2017; see also Berlyne, 1954), and/or attention (Pan et al., 2020a, 2020b), any of which may cause learners to profit more from subsequent learning opportunities.

Recent empirical research provides support for some of these accounts. For instance, there is evidence that pre-instruction testing may improve attention. In Pan et al., (2020a, 2020b), participants viewed a 26-min, four-segment video lecture about signal detection theory accompanied by pretests or no pretests (a control activity, solving math problems, occurred instead). They rated their level of attention after each video segment. For each segment, self-reported attention was higher in the pretested condition, which performed better on a subsequent posttest; a mediation analysis further revealed links between attention levels and posttest performance. These results suggest that improved attention—or reductions in the tendency to mind wander away from the task at hand—may underpin pre-instruction testing effects involving video and other types of materials (although see Welhaf et al., 2022 for a critique of that conclusion).

At least two eye tracking studies suggest that pre-instruction testing gives learners an idea of the kinds of information that they should seek out or pay attention to, shaping subsequent learning behaviors as a result. Lewis and Mensink (2012) had participants attempt or not attempt pre-instruction questions prior to reading text passages about space travel, physiology, or epidemiology; in most cases, engaging in prior testing caused learners to spend more time rereading and revisiting sentences that were relevant to the questions. Similarly, Yang et al. (2021) found that pre-instruction testing caused participants to fixate longer on portions of a 5-min video lecture about nutrition that pertained to the questions (although that effect was limited to participants with high achievement motivation). The results of both studies are broadly consistent with the suggestion that pre-instruction testing provides a “metacognitive ‘reality check’” (Carpenter & Toftness, 2017, p. 105), highlighting what one does and does not know and, in turn, encouraging information seeking or exploratory behaviors that lead to improved learning (see also Little & Bjork, 2016, Experiment 2). As with the attention account, a metacognitive ‘reality check’ could manifest following pre-instruction testing on a wide range of materials ranging from text passages to video lectures.

There is also evidence that memories for the practice test questions themselves play a role in pre-instruction testing effects. In St. Hilaire and Carpenter (2020; Experiments 2–4; see also Pan et al., 2020a, 2020b, experiment 2), participants attempted a dozen questions prior to watching a 31-min video lecture about information theory. While watching the video, participants were asked to take notes on information relevant to the questions (experiment 2) or write down/identify the answers to the questions (experiments 3–4). A pre-instruction testing effect relative to a non-testing control condition was observed only in cases where participants were able to recall the questions they had attempted (as evident by their notes or answer identification behaviors during the video). That result suggests that episodic memories for specific questions are an important contributor to pre-instruction

testing effects (although it should be noted that a specific as opposed to general effect of pre-instruction testing on memory was observed in that study).

The *attentional window* hypothesis recently proposed by Sana and Carpenter (2023) provides a way to integrate the aforementioned accounts—that is, whether processes posited by the mathemagenic hypothesis and/or other theoretical perspectives of pre-instruction testing effects occur as a result of testing—and address conflicting data regarding specific versus general effects (i.e., inconsistency across studies). This hypothesis posits that pre-instruction testing opens an attentional window wherein learners actively search for the answers to practice test questions, and once those answers are discovered, the window closes. An implication of that account is that practice questions targeting information that will be encountered at a later point in a subsequent learning opportunity will increase the likelihood of a general benefit when untested information appears earlier in the learning opportunity. In other words, untested information benefits by being “in the attentional window” when it occurs prior to the tested information. In two experiments with practice questions that targeted information appearing early or late in a subsequently presented text passage or video, Sana and Carpenter reported results consistent with the attentional window hypothesis (i.e., specific versus general benefits, respectively, when the questions targeted content that appeared early prior to untested content, versus late after untested content, in the learning materials). The purported attentional window could entail greater levels of interest, curiosity, and potentially other cognitive processes that lead to improved learning.

A Three-Stage Theoretical Framework for Pre-instruction Testing Effects

Additional research is needed to further explicate the theoretical mechanisms that are responsible for pre-instruction testing effects. The foregoing theoretical perspectives, however, are largely consistent with the following three-stage framework. We propose this framework, which is depicted in Fig. 4, to help organize the perspectives proposed thus far, some of which fit into specific stages, and to clarify the psychological processes that may be involved when pre-instruction testing is used.

In the first stage, the event of taking a practice test engages one or more psychological processes that are not triggered, or triggered to a lesser extent, by non-testing methods. These processes may be relatively general, such as an enhanced state of curiosity; alternatively, they may be more specific, such as forming a memory for a particular practice question, generating a set of possible answers, developing a mental framework for the information that is to be learned, or generating more retrieval routes to a particular answer. In the second stage, which occurs during a subsequent learning opportunity or during the presentation of correct answer feedback, the encoding of information into long-term memory is optimized to a greater extent than in a non-testing condition. These optimizations may involve improved attention, searching for specific answers, connecting answers to memories for practice questions, and other processes that do not occur or only occur to a lesser degree in a non-testing condition. In the third and final stage, which occurs during a subsequent posttest, improved performance in the tested (i.e., experimental) condition is

Three-stage theoretical framework for pre-instruction testing effects

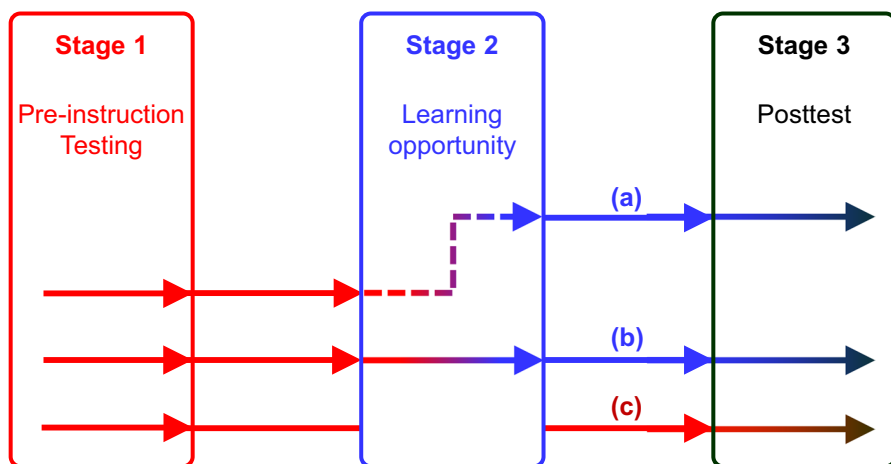


Fig. 4 Three-stage theoretical framework showing three potential routes to improved posttest performance. **(a)** Taking a practice test triggers a general psychological process or state, such as enhanced curiosity, that affect subsequent learning behaviors indirectly. **(b)** Taking a practice test causes memories to be formed (e.g., for a particular question) that drive specific learning behaviors (e.g., answer search). **(c)** Taking a practice test causes memories to be formed that act as retrieval cues on a subsequent posttest (largely but not entirely bypassing the second stage)

driven by better encoded information from the second stage, and/or with potential assistance from specific memories (e.g., mediators) or expanded retrieval routes that were formed in the first stage.

The processes described above can be categorized into three potential routes for improved posttest performance. As detailed in Fig. 4, pre-instruction testing may directly or indirectly affect subsequent learning behaviors (routes a and b), leading to improved recall. As an alternative (and perhaps less likely) possibility, memories for the practice test event itself may benefit posttest performance even without having a major impact on subsequent learning behaviors. For instance, the generation of guesses on a pretest—or consideration of the possibilities, principles, or theories addressed by a prequestion—may better equip a learner to answer a subsequent posttest question than if they had not engaged in such guessing or consideration.

The three-stage framework outlines ways that pre-instruction testing may confer a learning benefit on information that was specifically tested. As described previously, however, pre-instruction testing can also confer a general benefit on information from the learning episode that was not tested. Although the general benefit is far less common and theoretical progress toward understanding this effect is in its very early stages, the evidence so far points to a mechanism whereby pre-instruction questions orient attention to both tested and untested information during a learning episode. Consistent with this idea, pre-instruction testing is more likely to lead to both specific and general benefits when the untested information shares a noticeable relationship with the tested information (e.g., a pretest over a particular type of geyser leads to better learning of other, untested information about geysers, Little

& Bjork, 2016), or when untested information is placed before tested information in the learning episode (Sana & Carpenter, 2023). Both manipulations increase the chances that pre-instruction testing will, by virtue of the relatedness and proximity between tested and untested content, stimulate attention to both types of information. The fact that these general benefits do not occur automatically, but instead appear to be sensitive to the learning materials and how they support a learner's engagement, would suggest that the general benefits arise during the second stage. Although more research is certainly needed, these findings highlight attentional processes as candidate mechanisms contributing to the general benefits of pre-instruction testing.

As an additional point, the theoretical accounts for specific or general benefits of pre-instruction testing are not necessarily mutually exclusive. It is possible that multiple cognitive mechanisms and processes contribute to these effects. The relative contributions of these processes at each stage may also vary according to the materials being learned, the implementation of pre-instruction testing that was used, and other factors. Finally, for related theorizing that addresses the case of incorrect or correct responses to cued recall tests, see Kornell et al., (2015; also Kornell & Vaughn, 2016; Rickard & Pan, 2018).

Comparable Learning Strategies and Phenomena

At least several other learning strategies resemble or might be considered as a variant of pre-instruction testing. Studies involving these strategies provide further context for theoretical and practical issues. Additionally, some studies have investigated the effects of pre-instruction testing in combination with, or in comparison against, organizational signals (e.g., text headings or previews) or retrieval practice. Such studies exemplify an emerging line of research that addresses two or more evidence-based learning strategies simultaneously (for related discussions, see McDaniel, 2023; Roelle et al., 2023).

Test-Potentiated Learning

Taking a test prior to a study opportunity can lead to improved learning from that study opportunity. The literature on that phenomenon, which is known as *test-potentiated learning*, has typically been conducted using materials that are studied prior to being tested (i.e., retrieval practice) and then restudied (e.g., Arnold & McDermott, 2013; Izawa, 1970). By definition, pre-instruction testing effects could be regarded as forms of test-potentiated learning. More recently, a growing body of literature has also demonstrated that taking tests can improve subsequent learning of new materials, a phenomenon known as *test-potentiated new learning* or the *forward testing effect* (for reviews see Chan et al., 2018; Pastötter & Bauml, 2014; Yang et al., 2018; see also Boustani & Shanks, 2022). It has been suggested that the pre-instruction testing effect constitutes an example of test-potentiated new learning (e.g., Chan et al., 2018), although unlike most studies in the literature on that

phenomenon, pre-instruction testing does not entail a study opportunity prior to an initial test and does not typically involving practice testing on information that is markedly different from subsequently studied information.

A host of theoretical explanations has been proffered for test-potentiated learning and test-potentiated new learning (for discussions see Chan et al., 2018; Yang et al., 2018). The explanations that are arguably most applicable to pre-instruction testing involve learners adopting more optimal encoding or retrieval strategies after practice testing. For instance, it has been suggested that the experience of taking a test causes learners to exert greater effort at encoding subsequently presented information. Alternatively, learners may focus on encoding the types of information that they were previously tested on. Those changes in learning behaviors may stem from knowledge of the type of test to follow, improved awareness of one's own state of knowledge, or from the experience of retrieval failures.

Retrieval Practice

At least two studies to date have investigated whether pre-instruction testing prior to a lecture, followed by retrieval practice on content from that lecture, is more effective than retrieval practice alone. In Geller et al. (2017) and Carpenter et al. (2018), undergraduate students attempted a single pre-instruction question prior to a given lecture, during which the answer to the question was learned, and then attempted that question again at the end of the lecture (as a form of retrieval practice). Another, previously unpracticed question was also included at the end of lecture for retrieval practice. On a posttest occurring up to one week later, students attempted both questions again. Performance was not significantly better for the question that had been practiced twice (first as a pre-instruction question and then as retrieval practice) versus the question that had been practiced one (as retrieval practice only). These results suggest that the effects of pre-instruction testing—at least in a low-dosage form and without any further enhancements (e.g., specialized instructions or detailed feedback)—do not augment, nor detract from, the effects of retrieval practice. The theoretical basis for these results remains to be determined. The authors of both studies, however, speculated that how pre-instruction testing was implemented and characteristics of the course materials may have been factors (Carpenter et al., 2018; Geller et al., 2017).

Three studies to date have directly compared the efficacy of pre-instruction testing with retrieval practice. In Latimier et al. (2019), online participants took a multiple-choice practice test with immediate correct answer feedback either before (i.e., pretesting) or after (i.e., retrieval practice) reading biology text passages. On a one-week delayed posttest, both pretesting and retrieval practice effects were observed, but the effects were larger in the case of retrieval practice. De Lima and Jaeger (2020) included a retrieval practice condition wherein elementary school students attempted to fill in missing keywords in an encyclopedia text after having read that text with all words intact; that condition outperformed a pretesting condition on a one-week delayed test. Across five experiments, Pan and Sana (2021) had participants complete practice questions in cued recall or multiple-choice format before or after reading an encyclopedic text passage. On posttests conducted after a 5-min

or 48-h delay, performance was consistently higher in the pretesting versus retrieval practice conditions (with an advantage of $d_s = 0.30$ and 0.14 , across all experiments, for pretesting versus retrieval practice).

The results of these three studies—two showing an advantage for retrieval practice and one showing an advantage for pre-instruction testing—present an interesting theoretical puzzle. The mixed results raise the question of whether retrieval practice and pre-instruction testing evoke the same or different cognitive processes. On one hand, Kornell and Vaughn (2016) have argued that retrieval practice and pre-instruction testing represent two ways to tap into a common, two-stage mechanism for learning from retrieval attempts, whereas Pan and Sana (2021) have speculated that separate mechanisms may be involved (i.e., retrieval practice helps consolidate prior learning in long-term memory and pre-instruction testing spurs improvements in the subsequent encoding of information).

Invention Activities and Productive Failure

Having students attempt to develop formulas or procedures, followed by instruction on the correct formulas or procedures, can yield better learning than receiving such instruction outright. Evidence for the benefits of *invention* or *inventing activities* comes from studies in the domains of statistics and physics (e.g., Chin et al., 2016; Schwartz et al., 2011; Schwartz & Martin, 2004; see also Brydges et al., 2022). In alignment with some of the aforementioned theories of pre-instruction testing effects, the benefits of engaging in invention activities have been attributed to positive influences of those activities on subsequent learning behaviors (Schwartz et al., 2011). Alternatively, engaging in invention activities may prevent misconceptions that arise from receiving instruction over procedures and concepts before an opportunity to apply those procedures and concepts (Schwartz et al., 2011).

The *productive failure* literature (e.g., Kapur, 2008, 2015; for a meta-analysis, see Sinha & Kapur, 2021) indicates that having learners attempt to solve complex problems prior to instruction can be more effective than receiving instruction from the outset. Studies of productive failure have been conducted primarily in the domains of mathematics, physical sciences, and medicine and, unlike most studies of pre-instruction testing, often involve group activities (e.g., Kapur & Bielaczyc, 2012). Explanations of productive failure effects include scaffolding via the activation of relevant knowledge, attention being focused on critical problem features, and highlighting knowledge gaps (Loibl & Rummel, 2014; Sinha & Kapur, 2021). These explanations also resemble some of the theoretical accounts in the pre-instruction testing literature.

Organizational Signals and Learning Objectives

The use of *signaling devices* in text materials, including headings, summaries, previews, overviews, and typographical cues (e.g., underlying, bolding, italics, and capitalization), can improve memory for the information targeted by such devices (for a review, see Lorch, 1989). Such devices appear to do so by orienting readers'

attention to specific portions of a given text, which also resembles some of the theorizing in the pre-instruction testing and adjunct questions literatures (e.g., Hartley & Davies, 1976, Lewis & Mensink, 2012; see also Hamilton, 1985). Directly addressing that suggestion, Richland et al., (2009; experiments 3–5; cf. James & Storm, 2019, experiment 4) compared the effects of pretesting versus reading on a text passage that featured key sentences in italics or keywords presented in bold. Across all experiments, pretesting enhanced memory over a read-only condition. Those results suggest that the effects of pre-instruction testing for text materials surpass that conferred by signaling devices in the form of typographical cues.

Sana et al., (2020a, 2020b; experiment 2) compared the efficacy of reading *learning objectives* (i.e., statements about specific information that students should learn) versus taking pretests on learning objectives prior to reading text passages on neuroscience topics. On a 5-min delayed posttest, participants were better able to recall content targeted by the learning objectives when they had previously taken pretests. Similar to Richland et al.'s (2009) findings for typographical cues, these results suggest that the benefits of pretests are beyond that conferred by learning objectives and possibly other statements read in advance of a learning opportunity.

Educational Applications

The present review summarizes a rich body of literature showing that asking students questions is a powerful way to enhance their learning. Given the longstanding history of using questions for examination purposes in education, the idea of using questions for learning purposes is fairly new and sometimes counterintuitive. In particular, asking students questions over a topic they have not yet learned is almost certain to result in failure to answer those questions successfully. Thus, students and teachers alike may raise the very reasonable question of what purpose these pre-instruction tests serve for learning. While the idea of using prequestions or pretests to check students' knowledge is quite intuitive and can help with lesson planning, it is not obvious or intuitive that such tests can also be used to directly enhance learning.

Findings from empirical research are thus critically important in considering educational practices designed to improve learning. The current review focuses on pre-instruction testing as an effective, even if non-intuitive, pedagogical practice. The emerging research in this area highlights a number of ways that pre-instruction testing can be applied in educational settings and important factors for instructors to consider in the implementation of pre-instruction tests in real-world learning situations.

In authentic educational settings, pre-instruction testing has been shown to be beneficial in a variety of situations, including elementary school students learning about space exploration (de Lima & Jaeger, 2020), middle school students learning science (McDaniel et al., 2011), undergraduate students learning psychology (Carpenter et al., 2018; Soderstrom & Bjork, 2023) and chemical engineering (Geller et al., 2017), and medical students learning surgical procedures (Willis et al., 2020). These studies also demonstrate that pre-instruction testing can be delivered

in a number of different ways, from low-tech paper-and-pencil activities, to digital methods such as in-class “clicker” questions and pre-instruction quizzes embedded in online course management systems that can suit larger classes. Benefits of pre-instruction testing have been observed across multiple delivery methods (although as previously noted, the choice of test format and other implementation factors may influence such benefits; further research is needed to clarify those possibilities). Overall, the research from real educational environments demonstrates that pre-instruction testing is an effective and flexible approach for enhancing student learning across a variety of topics and levels of education. Research shows as well that these benefits apply across a variety of question types (see Table 2). Instructors can thus straightforwardly implement and adapt pre-instruction testing to any course material using resources available to them.

Important points to consider when implementing pre-instruction testing in educational settings are as follows. First, pre-instruction testing is likely to involve situations where students get the answers wrong, or simply do not know the answers, to the questions they are asked. Although these errors are certainly reasonable and should be expected for tests that occur prior to a learning opportunity, it is important to anticipate that the experience of making errors could be undesirable for students. Indeed, students prefer to avoid making errors, even if they know that errors are part of the learning process (Pan et al., 2020a, 2020b). Other research shows that students strongly prefer to read information rather than engage in testing, especially if they feel they do not yet know the information well (Kirk-Johnson et al., 2019). Consistent with these findings, research confirms that students are largely unaware of the learning benefits of pre-instruction testing, but they can become more aware through direct experience and through instruction aimed at increasing their understanding of the positive role of errors in learning (Pan & Rivers, 2023; Yang et al., 2017). Thus, students may be reluctant at first to engage with pre-instruction testing, and instructors may anticipate that students may need some “warm up” time to familiarize with the approach and have an opportunity to experience its benefits in their own learning.

Second, implementing pre-instruction testing requires time on the part of instructors. Creating the questions themselves is a non-trivial time investment, especially if instructors wish to use these practice questions on a regular basis across an academic term, which could mean the creation of several dozen practice questions per course. Implementing pre-instruction testing could affect time spent in class as well. In particular, if questions are posed to students at the beginning of class, students will need time to answer those questions, and instructors will need to adjust accordingly by shortening the time for the subsequent lesson. These are important considerations, as time is a critical resource in education. Although creating practice questions requires time, fortunately, this time is front-loaded such that once created, those practice questions can be used again each time the course is taught.

Furthermore, the in-class time required for students to answer pre-instruction questions should be considered in light of the learning gains afforded by those questions. Decades of research from the learning sciences shows that it is less about how much time is spent, and more about *how* that time is spent. Learning can be substantially and significantly boosted by using an effective learning strategy, for example,

trying to retrieve information, than by spending the same amount of time reading, reviewing, or simply being presented the information (e.g., Carpenter et al., 2022). Although answering pre-instruction questions shifts time from instructor-led activities over the lesson, these questions in themselves are an effective learning activity that can significantly enhance students' retention and comprehension of the lesson. From a cost–benefit perspective, therefore, starting a lesson with pre-instruction questions is time well spent. It is also notable that pre-instruction testing has been beneficial in studies wherein time-on-task has been strictly controlled—hence, taking time away from studying to engage in pretesting or prequestioning can be uniquely beneficial.

Finally, it is worth noting that pre-instruction testing is likely to benefit learning of information that is explicitly included in a lesson. Most of the research on pre-instruction testing involves learning material that provides fairly straightforward answers to the test questions. For example, if a pre-instruction question asks, “What is the oldest geyser in the world?”, a subsequent reading passage about Yellowstone National Park includes, among other information, a clear statement that Castle Geyser is believed to be the oldest geyser (Little & Bjork, 2016). Although some studies have shown benefits of pre-instruction testing on more complex questions that require integrating information from different parts of a reading passage, the benefits are still more common for fairly discrete factual information that is directly stated in one place (St. Hilaire et al., 2019). This pattern is likely due to the fact that learners need to remember the pre-instruction question itself, plus identify the relevant information in the reading or lesson that provides the answer to that question, in order to successfully learn it (St. Hilaire & Carpenter, 2020).

In situations where these connections are less clear—such as when the pre-instruction questions are too complex, or the learning material itself is difficult to comprehend—pre-instruction questions may not benefit learning. Indeed, this pattern could be why pre-instruction questions that require complex inferences—e.g., deducing a conclusion or premise that is not directly stated in the learning material—tend not to benefit learning (Hausman & Rhodes, 2018). Future research can shed more light on this phenomenon, as very few studies have been conducted on the effects of complex inference questions. Presently, however, it is clear that pre-instruction testing is a useful and effective learning strategy for material that is fairly concrete and straightforward. As it can be the case that particular learning strategies are sometimes best for particular types of material (see Dunlosky et al., 2013), instructors will likely find it beneficial to use pre-instruction testing particularly in cases where the learning of basic facts and concepts is needed. Moreover, uses of pre-instruction testing do not necessarily preclude the incorporation of other learning strategies that may also lead to positive educational outcomes.

Currently, pre-instruction testing is commonly implemented for the purposes of formative assessment (for a review, see Black & Wiliam, 1998), wherein instructors and students use such testing to help guide future teaching activities and study behaviors, respectively. For instance, a pretest may be administered at the start of a class to help teachers measure students' preexisting knowledge, to identify students that need extra assistance, and to identify areas to focus on during subsequent study activities (Bennett, 2011; Heritage, 2007). In such cases, there is usually no

consideration of any direct impacts that practice testing might have on memory, comprehension, or transfer of learning (e.g., Boston, 2002; Heritage, 2007). The research discussed in this review, however, suggests that such testing has considerable potential to directly improve learning.

Conclusions and Future Directions

This comprehensive review of pre-instruction testing shows that answering questions about to-be-learned information benefits learning of that information. These benefits apply across fairly simple materials, such as paired associates, to more complex materials, such as reading passages, videos, and live lectures. Benefits of pre-instruction testing have been shown across a variety of test formats and time intervals and in real educational environments with different age groups and different subject matters.

Theoretical investigations of pre-instruction testing and related effects have identified candidate mechanisms that serve to increase the effectiveness of encoding during a subsequent learning opportunity. Although support for these mechanisms can be found in various studies, additional research is encouraged that can contribute further theoretical insights and establish boundary conditions for the benefits of pre-instruction testing. In particular, many studies on pre-instruction testing typically measure memory for learned material following testing versus no testing, and use memory performance to infer how that material may have been processed. Although some insights can and have been gleaned from such studies, a fruitful direction for future research is to collect measures of real time processing during a learning activity.

Following pre-instruction testing versus no testing, for example, direct measures of processing—such as reading times for specific content, eye tracking, or dual task conditions to measure attentional resources devoted to the learning material—can reveal new insights into what learners do while engaging with material following pre-instruction testing. Although a couple of studies have sometimes included reading times and eye tracking (Lewis & Mensink, 2012; Little & Bjork, 2016), such real time measures can be explored more extensively and systematically. Such measures may be especially important in understanding the processing that contributes to the general benefits of pre-instruction testing on untested material.

Considering the practical applications of pre-instruction testing, additional research in authentic educational environments is encouraged that can help inform how best to use pre-instruction testing for educational purposes. Indeed, the literature on pre-instruction testing effects in the classroom remains relatively small compared to that of other, better-established learning strategies such as retrieval practice. Moreover, although much of the research compares pre-instruction testing to a control condition involving no testing and no alternative learning activity, additional research comparing pre-instruction testing to active controls—such as learning objectives (Sana et al., 2020a, 2020b) or other preparatory activities such as reflecting on the importance or interest in the material to be learned—would shed new light on the efficacy of pre-instruction testing relative to other educationally relevant

activities. As real learning situations include a number of choices and options, additional research into potential combined benefits of pre-instruction testing and other learning strategies (e.g., retrieval practice, generating one's own questions over the material to be learned, summarization or reflection) are also worth exploring. In addition, more direct comparisons of pre-instruction testing versus other learning strategies are needed to clarify the relative potency of such testing. More studies exploring the durability of pre-instruction effects over time will also be helpful in understanding and anticipating the time course of these benefits.

Finally, a relevant question for future research is how best to incorporate pre-instruction testing into modern learning environments. Online or hybrid courses are more flexible and rely on student-initiated actions more so than traditional classroom-based courses. In these unsupervised environments, are there optimal ways to deliver pre-instruction testing that ensures students are engaged with the process? Are there drawbacks to pre-instruction testing—such as the experience of errors or the belief that such testing is not beneficial for learning (e.g., Pan & Rivers, 2023; Yang et al., 2017)—that might lead students to disengage from it when given the choice? Given the rapidly increasing availability of non-traditional learning opportunities, exploring the efficacy of pre-instruction testing in these contexts will shed important light on its utility as an educational tool in modern contexts.

Funding

This research was supported by a Faculty of Arts & Social Sciences (FASS) grant from the National University of Singapore. This research is also based upon work supported by the James S. McDonnell Foundation 21st Century Science Initiative in Understanding Human Cognition, Collaborative Grant No. 220020483.

Acknowledgements Thanks to Ahmad Naufal, Benedict Yew, Faith Siau, Justina Tan, Marcus Wong, and Shruthi Kumar for assistance with verifying references.

Data Availability Given that this manuscript is a review of existing research, no generation of new datasets or analysis of such a dataset was performed.

Declarations Disclaimer.

Any opinions, findings, conclusions, or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the James S. McDonnell Foundation.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Anderson, R. C., & Biddle, W. B. (1975). On asking people questions about what they are reading. In *Psychology of Learning and Motivation* (Vol. 9, pp. 89–132). Elsevier. [https://doi.org/10.1016/S0079-7421\(08\)60269-8](https://doi.org/10.1016/S0079-7421(08)60269-8)
- Arnold, K. M., & McDermott, K. B. (2013). Test-potentiated learning: distinguishing between direct and indirect effects of tests. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 39(3), 940–945. <https://doi.org/10.1037/a0029199>
- Barnett, S. M., & Ceci, S. J. (2002). When and where do we apply what we learn?: A taxonomy for far transfer. *Psychological Bulletin*, 128(4), 612–637. <https://doi.org/10.1037/0033-2909.128.4.612>
- Beckman, W. (2008). Pre-testing as a method of conveying learning objectives. *Journal of Aviation/Aerospace Education & Research*. <https://doi.org/10.15394/jaaer.2008.1447>
- Bennett, R. E. (2011). Formative assessment: A critical review. *Assessment in Education: Principles, Policy & Practice*, 18(1), 5–25. <https://doi.org/10.1080/0969594X.2010.513678>
- Berlyne, D. E. (1954). An experimental study of human curiosity. *British Journal of Psychology: General Section*, 45(4), 256–265. <https://doi.org/10.1111/j.2044-8295.1954.tb01253.x>
- Bjork, R. A. (1975). Retrieval as a memory modifier. In R. Solso (Ed.), *Information processing and cognition: The Loyola Symposium* (pp. 123–144). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Bjork, R. A., Dunlosky, J., & Kornell, N. (2013). Self-regulated learning: Beliefs, techniques, and illusions. *Annual Review of Psychology*, 64(1), 417–444. <https://doi.org/10.1146/annurev-psych-113011-143823>
- Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice*, 5(1), 7–74. <https://doi.org/10.1080/0969595980050102>
- Boston, C. (2002). The concept of formative assessment. *Practical Assessment, Research, and Evaluation*. <https://doi.org/10.7275/KMCCQ-DJ31>
- Boustani, S., & Shanks, D. R. (2022). Heterogeneity and publication bias in research on test-potentiated new learning. *Collabra: Psychology*, 8(1), 31996. <https://doi.org/10.1525/collabra.31996>
- Brod, G. (2021). Predicting as a learning strategy. *Psychonomic Bulletin & Review*. <https://doi.org/10.3758/s13423-021-01904-1>
- Brydges, R., Fiume, A., & Grierson, L. (2022). Mastery versus invention learning: Impacts on future learning of simulated procedural skills. *Advances in Health Sciences Education*, 27(2), 441–456. <https://doi.org/10.1007/s10459-022-10094-x>
- Bull, S. G. (1973). The role of questions in maintaining attention to textual material. *Review of Educational Research*, 43(1), 5.
- Butowska, E., Hanczakowski, M., & Zawadzka, K. (2021). You won't guess that: On the limited benefits of guessing when learning a foreign language. *Memory & Cognition*. <https://doi.org/10.3758/s13421-021-01254-2>
- Butterfield, B., & Mangels, J. (2003). Neural correlates of error detection and correction in a semantic retrieval task. *Brain Research. Cognitive Brain Research*, 17, 793–817. [https://doi.org/10.1016/S0926-6410\(03\)00203-9](https://doi.org/10.1016/S0926-6410(03)00203-9)
- Butterfield, B., & Metcalfe, J. (2001). Errors committed with high confidence are hypercorrected. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 27(6), 1491–1494. <https://doi.org/10.1037/0278-7393.27.6.1491>
- Carneiro, P., Lapa, A., & Finn, B. (2018). The effect of unsuccessful retrieval on children's subsequent learning. *Journal of Experimental Child Psychology*, 166, 400–420. <https://doi.org/10.1016/j.jecp.2017.09.010>
- 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(6), 1547–1552. <https://doi.org/10.1037/a0024140>
- Carpenter, S. K. (2012). Testing enhances the transfer of learning. *Current Directions in Psychological Science*, 21(5), 279–283. <https://doi.org/10.1177/0963721412452728>
- Carpenter, S. K., & Toftness, A. R. (2017). The effect of prequestions on learning from video presentations. *Journal of Applied Research in Memory and Cognition*, 6(1), 104–109. <https://doi.org/10.1016/j.jarmac.2016.07.014>
- Carpenter, S. K., & Yeung, K. L. (2017). The role of mediator strength in learning from retrieval. *Journal of Memory and Language*, 92, 128–141. <https://doi.org/10.1016/j.jml.2016.06.008>

- Carpenter, S. K., Sachs, R. E., Martin, B., Schmidt, K., & Looft, R. (2012). Learning new vocabulary in German: The effects of inferring word meanings, type of feedback, and time of test. *Psychonomic Bulletin & Review*, 19(1), 81–86. <https://doi.org/10.3758/s13423-011-0185-7>
- Carpenter, S. K., Rahman, S., & Perkins, K. (2018). The effects of prequestions on classroom learning. *Journal of Experimental Psychology: Applied*, 24(1), 34–42. <https://doi.org/10.1037/xap0000145>
- Carpenter, S. K., Pan, S. C., & Butler, A. C. (2022). The science of effective learning with spacing and retrieval practice. *Nature Reviews Psychology*. <https://doi.org/10.1038/s44159-022-00089-1>
- Carpenter, S. K., King-Shepard, Q., & Nokes-Malach, T. (2023). The prequestion effect: Why it is useful to ask students questions before they learn. In *In their own words: What scholars want you to know about why and how to apply the science of learning in your academic setting* (pp. 74–82). Division 2 American Psychological Association.
- Carver, R. P. (1971). A critical review of mathemagenic behaviors and the effect of questions upon the retention of prose materials. *Journal of Reading Behavior*, 4(2), 93–119. <https://doi.org/10.1080/10862967109546983>
- Chan, J. C. K., Meissner, C. A., & Davis, S. D. (2018). Retrieval potentiates new learning: A theoretical and meta-analytic review. *Psychological Bulletin*, 144(11), 1111–1146. <https://doi.org/10.1037/bul0000166>
- Chin, D. B., Chi, M., & Schwartz, D. L. (2016). A comparison of two methods of active learning in physics: Inventing a general solution versus compare and contrast. *Instructional Science*, 44(2), 177–195. <https://doi.org/10.1007/s11251-016-9374-0>
- Collins, A., & Loftus, E. (1975). A spreading activation theory of semantic processing. *Psychological Review*, 82, 407–428. <https://doi.org/10.1037//0033-295X.82.6.407>
- de Lima, N. K., & Jaeger, A. (2020). The effects of prequestions versus postquestions on memory retention in children. *Journal of Applied Research in Memory and Cognition*, 9(4), 555–563. <https://doi.org/10.1016/j.jarmac.2020.08.005>
- 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. <https://doi.org/10.1177/1529100612453266>
- Ergo, K., De Loof, E., & Verguts, T. (2020). Reward prediction error and declarative memory. *Trends in Cognitive Sciences*, 24(5), 388–397. <https://doi.org/10.1016/j.tics.2020.02.009>
- Frase, L. T. (1968). Effect of question location, pacing, and mode upon retention of prose material. *Journal of Educational Psychology*, 59(4), 244–249. <https://doi.org/10.1037/h0025947>
- Geller, J., Carpenter, S. K., Lamm, M. H., Rahman, S., Armstrong, P. I., & Coffman, C. R. (2017). Prequestions do not enhance the benefits of retrieval in a STEM classroom. *Cognitive Research: Principles and Implications*, 2(1), 42. <https://doi.org/10.1186/s41235-017-0078-z>
- Grimaldi, P. J., & Karpicke, J. D. (2012). When and why do retrieval attempts enhance subsequent encoding? *Memory & Cognition*, 40(4), 505–513. <https://doi.org/10.3758/s13421-011-0174-0>
- Gustafson, H. W., & Toole, D. L. (1970). Effects of adjunct questions, pretesting, and degree of student supervision on learning from an instructional text. *The Journal of Experimental Education*, 39(1), 53–58. <https://doi.org/10.1080/00220973.1970.11011231>
- Hamaker, C. (1986). The effects of adjunct questions on prose learning. *Review of Educational Research*, 56(2), 212–242. <https://doi.org/10.3102/00346543056002212>
- Hamilton, R. J. (1985). A framework for the evaluation of the effectiveness of adjunct questions and objectives. *Review of Educational Research*, 55(1), 47–85.
- Hartley, J. (1973). The effect of pre-testing on post-test performance. *Instructional Science*, 2(2), 193–214. <https://doi.org/10.1007/BF00139871>
- Hartley, J., & Davies, I. (1976). Preinstructional strategies: The role of pretests, behavioral objectives, overviews and advance organizers. *Review of Educational Research*, 46(2), 239–265.
- Hausman, H., & Rhodes, M. G. (2018). When pretesting fails to enhance learning concepts from reading texts. *Journal of Experimental Psychology: Applied*, 24(3), 331–346. <https://doi.org/10.1037/xap0000160>
- 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(1), 290–296. <https://doi.org/10.1037/a0028468>
- Heritage, M. (2007). Formative assessment: What do teachers need to know and do? *Phi Delta Kappan*, 89(2), 140–145. <https://doi.org/10.1177/003172170708900210>

- Huelser, B. J., & Metcalfe, J. (2012). Making related errors facilitates learning, but learners do not know it. *Memory & Cognition*, 40(4), 514–527. <https://doi.org/10.3758/s13421-011-0167-z>
- Izawa, C. (1970). Optimal potentiating effects and forgetting-prevention effects of tests in paired-associate learning. *Journal of Experimental Psychology*, 83(2, Pt.1), 340–344. <https://doi.org/10.1037/h0028541>
- Jacoby, L. L., & Wahlheim, C. N. (2013). On the importance of looking back: The role of recursive reminders in recency judgments and cued recall. *Memory & Cognition*, 41(5), 625–637. <https://doi.org/10.3758/s13421-013-0298-5>
- James, K. K., & Storm, B. C. (2019). Beyond the pretesting effect: What happens to the information that is not pretested? *Journal of Experimental Psychology: Applied*, 25(4), 576–587. <https://doi.org/10.1037/xap0000231>
- Janelli, M., & Lipnevich, A. A. (2021). Effects of pre-tests and feedback on performance outcomes and persistence in Massive Open Online Courses. *Computers & Education*, 161, 104076. <https://doi.org/10.1016/j.compedu.2020.104076>
- Kang, S. H. K., Pashler, H., Cepeda, N. J., Rohrer, D., Carpenter, S. K., & Mozer, M. C. (2011). Does incorrect guessing impair fact learning? *Journal of Educational Psychology*, 103(1), 48–59. <https://doi.org/10.1037/a0021977>
- Kapur, M. (2008). Productive failure. *Cognition and Instruction*, 26(3), 379–424. <https://doi.org/10.1080/07370000802212669>
- Kapur, M. (2015). Learning from productive failure. *Learning: Research and Practice*, 1(1), 51–65. <https://doi.org/10.1080/23735082.2015.1002195>
- Kapur, M., & Bielaczyc, K. (2012). Designing for productive failure. *Journal of the Learning Sciences*, 21(1), 45–83. <https://doi.org/10.1080/10508406.2011.591717>
- King-Shepard, Q. W., Nokes-Malach, T. J., Carpenter, S. K., & Fraundorf, S. H. (In Prep). The effects of prequestions on learning: a meta-analysis.
- Kirk-Johnson, A., Galla, B. M., & Fraundorf, S. H. (2019). Perceiving effort as poor learning: The misinterpreted-effort hypothesis of how experienced effort and perceived learning relate to study strategy choice. *Cognitive Psychology*, 115, 101237. <https://doi.org/10.1016/j.cogpsych.2019.101237>
- Kliegl, O., Bartl, J., & Bäuml, K.-H.T. (2022). The pretesting effect comes to full fruition after prolonged retention interval. *Journal of Applied Research in Memory and Cognition*. <https://doi.org/10.1037/mac0000085>
- Knight, J. B., Hunter Ball, B., 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(4), 731–746. <https://doi.org/10.1016/j.jml.2011.12.008>
- 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(1), 106–114. <https://doi.org/10.1037/a0033699>
- Kornell, N., & Vaughn, K. E. (2016). How retrieval attempts affect learning: A review and synthesis. In B. H. Ross (Ed.), *Psychology of learning and motivation* (Vol. 65, pp. 183–215). Academic Press. <https://doi.org/10.1016/bs.plm.2016.03.003>
- Kornell, N., Hays, M. J., & Bjork, R. A. (2009). Unsuccessful retrieval attempts enhance subsequent learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 35(4), 989–998. <https://doi.org/10.1037/a0015729>
- Kornell, N., Klein, P. J., & Rawson, K. A. (2015). Retrieval attempts enhance learning, but retrieval success (versus failure) does not matter. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 41(1), 283–294. <https://doi.org/10.1037/a0037850>
- Kraft, M. A. (2020). Interpreting effect sizes of education interventions. *Educational Researcher*, 49(4), 241–253. <https://doi.org/10.3102/0013189X20912798>
- Latimier, A., Riegert, A., Peyre, H., Ly, S. T., Casati, R., & Ramus, F. (2019). Does pre-testing promote better retention than post-testing? *Npj Science of Learning*, 4(1), 15. <https://doi.org/10.1038/s41539-019-0053-1>
- Lewis, M. R., & Mensink, M. C. (2012). Prereading questions and online text comprehension. *Discourse Processes*, 49(5), 367–390. <https://doi.org/10.1080/0163853X.2012.662801>
- Little, J. L., & Bjork, E. L. (2011). Pretesting with multiple-choice questions facilitates learning. *Proceedings of the Annual Meeting of the Cognitive Science Society*.
- Little, J. L., & Bjork, E. L. (2016). Multiple-choice pretesting potentiates learning of related information. *Memory & Cognition*, 44(7), 1085–1101. <https://doi.org/10.3758/s13421-016-0621-z>

- Loibl, K., & Rummel, N. (2014). Knowing what you don't know makes failure productive. *Learning and Instruction*, 34, 74–85. <https://doi.org/10.1016/j.learninstruc.2014.08.004>
- Lorch, R. F. (1989). Text-signaling devices and their effects on reading and memory processes. *Educational Psychology Review*, 1(3), 209–234. <https://doi.org/10.1007/BF01320135>
- McDaniel, M. A. (2023). Combining Retrieval Practice with Elaborative Encoding: Complementary or Redundant?. *Educational Psychology Review* 35, 75. <https://doi.org/10.1007/s10648-023-09784-8>
- McDaniel, M. A., Agarwal, P. K., Huelser, B. J., McDermott, K. B., & Roediger, H. L. (2011). Test-enhanced learning in a middle school science classroom: The effects of quiz frequency and placement. *Journal of Educational Psychology*, 103(2), 399–414. <https://doi.org/10.1037/a0021782>
- Memory, D. M. (1983). Main idea prequestions as adjunct aids with good and low-average middle grade readers. *Journal of Reading Behavior*, 15(2), 37–48. <https://doi.org/10.1080/10862968309547482>
- Mera, Y., Rodríguez G., & Marin-Garcia, E. (2022). Unraveling the benefits of experiencing errors during learning: Definition, modulating factors, and explanatory theories. *Psychonomic Bulletin & Review*, 29, 753–765.
- Metcalfe, J. (2017). Learning from errors. *Annual Review of Psychology*, 68(1), 465–489. <https://doi.org/10.1146/annurev-psych-010416-044022>
- Metcalfe, J., & Huelser, B. J. (2020). Learning from errors is attributable to episodic recollection rather than semantic mediation. *Neuropsychologia*, 138, 107296. <https://doi.org/10.1016/j.neuropsychologia.2019.107296>
- Overoye, A. L., James, K. K., & Storm, B. C. (2021). A little can go a long way: Giving learners some context can enhance the benefits of pretesting. *Memory*, 29(9), 1206–1215. <https://doi.org/10.1080/09658211.2021.1974048>
- Pan, S. C., & Rickard, T. C. (2018). Transfer of test-enhanced learning: Meta-analytic review and synthesis. *Psychological Bulletin*, 144(7), 710–756. <https://doi.org/10.1037/bul0000151>
- Pan, S. C., & Rivers, M. L. (2023). Metacognitive awareness of the pretesting effect improves with self-regulation support. *Memory & Cognition*. <https://doi.org/10.3758/s13421-022-01392-1>
- Pan, S. C., & Sana, F. (2021). Pretesting versus posttesting: Comparing the pedagogical benefits of errorful generation and retrieval practice. *Journal of Experimental Psychology: Applied*, 27(2), 237–257. <https://psycnet.apa.org/doi/10.1037/xap0000345>
- Pan, S. C., Lovelett, J., Stoeckenius, D., & Rickard, T. C. (2019). Conditions of highly specific learning through cued recall. *Psychonomic Bulletin & Review*, 26(2), 634–640. <https://doi.org/10.3758/s13423-019-01593-x>
- Pan, S. C., Sana, F., Samani, J., Cooke, J., & Kim, J. A. (2020a). Learning from errors: Students' and instructors' practices, attitudes, and beliefs. *Memory*, 28(9), 1105–1122. <https://doi.org/10.1080/09658211.2020.1815790>
- Pan, S. C., Sana, F., Schmitt, A. G., & Bjork, E. L. (2020b). Pretesting reduces mind wandering and enhances learning during online lectures. *Journal of Applied Research in Memory and Cognition*, 9(4), 542–554. <https://doi.org/10.1016/j.jarmac.2020.07.004>
- Pastötter, B., & Bauml, K.-H. T. (2014). Retrieval practice enhances new learning: The forward effect of testing. *Frontiers in Psychology*, 5. <https://doi.org/10.3389/fpsyg.2014.00286>
- Potts, R., & Shanks, D. R. (2014). The benefit of generating errors during learning. *Journal of Experimental Psychology: General*, 143(2), 644–667. <https://doi.org/10.1037/a0033194>
- Potts, R., Davies, G., & Shanks, D. R. (2019). The benefit of generating errors during learning: What is the locus of the effect? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 45(6), 1023–1041. <https://doi.org/10.1037/xlm0000637>
- 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(1), 27–35. [https://doi.org/10.1016/0361-476X\(90\)90003-J](https://doi.org/10.1016/0361-476X(90)90003-J)
- Pyc, M. A., & Rawson, K. A. (2010). Why testing improves memory: Mediator effectiveness hypothesis. *Science*, 330(6002), 335–335. <https://doi.org/10.1126/science.1191465>
- Richland, L. E., Kornell, N., & Kao, L. S. (2009). The pretesting effect: Do unsuccessful retrieval attempts enhance learning? *Journal of Experimental Psychology: Applied*, 15(3), 243–257. <https://doi.org/10.1037/a0016496>
- Rickard, T. C., & Pan, S. C. (2018). A dual memory theory of the testing effect. *Psychonomic Bulletin & Review*, 25(3), 847–869. <https://doi.org/10.3758/s13423-017-1298-4>
- Rickards, J. P. (1976a). Interaction of position and conceptual level of adjunct questions on immediate and delayed retention of text. *Journal of Educational Psychology*, 68(2), 210–217.

- Rickards, J. P. (1976b). Type of verbatim question interspersed in text: A new look at the position effect. *Journal of Reading Behavior*, 8(1), 37–45. <https://doi.org/10.1080/10862967609547164>
- Rickards, J. P. (1979). Adjunct postquestions in text: A critical review of methods and processes. *Review of Educational Research*, 49(2), 181–196.
- Roediger, H. L., & Butler, A. C. (2011). The critical role of retrieval practice in long-term retention. *Trends in Cognitive Sciences*, 15(1), 20–27. <https://doi.org/10.1016/j.tics.2010.09.003>
- Roediger, H. L., & Karpicke, J. D. (2006). The power of testing memory: Basic research and implications for educational practice. *Perspectives on Psychological Science*, 1(3), 181–210. <https://doi.org/10.1111/j.1745-6916.2006.00012.x>
- Roelle, J., Endres, T., Obergassel, N., Nückles, M., & Renkl, A. (2023). Happy together? On the relationship between research on retrieval practice and generative learning using the case of follow-up learning tasks. *Educational Psychology Review*.
- 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(4), 241–249.
- Rothkopf, E. Z., & Bisbicos, E. E. (1967). Selective facilitative effects of interspersed questions on learning from written materials. *Journal of Educational Psychology*, 58(1), 56–61. <https://doi.org/10.1037/h0024117>
- Rowland, C. A. (2014). The effect of testing versus restudy on retention: A meta-analytic review of the testing effect. *Psychological Bulletin*, 140(6), 1432–1463. <https://doi.org/10.1037/a0037559>
- Sagaria, S. D., & Di Vesta, F. J. (1977). *Additive effects from interspersed adjunct questions in prose text*. Annual Meeting of the American Educational Research Association.
- Samuels, J. (1969). The effect of post-test relevant pre-tests and discussion-type feedback on learning and retention. *Psychonomic Science*, 16(2), 67–68.
- Sana, F., & Carpenter, S. K. (2023). Broader benefits of the pretesting effect: Placement matters. *Psychonomic Bulletin & Review*. <https://doi.org/10.3758/s13423-023-02274-6>
- Sana, F., Forrin, N. D., Sharma, M., Dubljevic, T., Ho, P., Jalil, E., & Kim, J. A. (2020). Optimizing the efficacy of learning objectives through pretests. *CBE—Life Sciences Education*, 19(3), ar43. <https://doi.org/10.1187/cbe.19-11-0257>
- Sana, F., Yan, V. X., Clark, C. M., Bjork, E. L., & Bjork, R. A. (2020b). Improving conceptual learning via pretests. *Journal of Experimental Psychology: Applied*. <https://doi.org/10.1037/xap0000322>
- Schwartz, D. L., & Martin, T. (2004). Inventing to prepare for future learning: The hidden efficiency of encouraging original student production in statistics instruction. *Cognition and Instruction*, 22(2), 129–184. https://doi.org/10.1207/s1532690xc2202_1
- Schwartz, D. L., Chase, C. C., Oppezzo, M. A., & Chin, D. B. (2011). Practicing versus inventing with contrasting cases: The effects of telling first on learning and transfer. *Journal of Educational Psychology*, 103(4), 759–775. <https://doi.org/10.1037/a0025140>
- Seabrooke, T., Hollins, T. J., Kent, C., Wills, A. J., & Mitchell, C. J. (2019a). Learning from failure: Errorful generation improves memory for items, not associations. *Journal of Memory and Language*, 104, 70–82. <https://doi.org/10.1016/j.jml.2018.10.001>
- Seabrooke, T., Mitchell, C. J., Wills, A. J., Waters, J. L., & Hollins, T. J. (2019b). Selective effects of errorful generation on recognition memory: The role of motivation and surprise. *Memory*, 27(9), 1250–1262. <https://doi.org/10.1080/09658211.2019.1647247>
- Seabrooke, T., Mitchell, C. J., Wills, A. J., & Hollins, T. J. (2021a). Pretesting boosts recognition, but not cued recall, of targets from unrelated word pairs. *Psychonomic Bulletin & Review*, 28(1), 268–273. <https://doi.org/10.3758/s13423-020-01810-y>
- Seabrooke, T., Mitchell, C. J., Wills, A. J., Inkster, A. B., & Hollins, T. J. (2021b). The benefits of impossible tests: Assessing the role of error-correction in the pretesting effect. *Memory & Cognition*. <https://doi.org/10.3758/s13421-021-01218-6>
- Sinha, T., & Kapur, M. (2021). When problem solving followed by instruction works: Evidence for productive failure. *Review of Educational Research*, 91(5), 761–798. <https://doi.org/10.3102/00346543211019105>
- Soderstrom, N. C., & Bjork, E. L. (2023). Pretesting Enhances Learning in the Classroom. *Educational Psychology Review*, 35(3), 88.
- Soraci, S. A., Carlin, M. T., Chechile, R. A., Franks, J. J., Wills, T., & Watanabe, T. (1999). Encoding variability and cuing in generative processing. *Journal of Memory and Language*, 41(4), 541–559. <https://doi.org/10.1006/jmla.1999.2661>

- St. Hilaire, K. J., & Carpenter, S. K. (2020). Prequestions enhance learning, but only when they are remembered. *Journal of Experimental Psychology: Applied*, 26(4), 705–716. <https://doi.org/10.1037/xap0000296>
- St. Hilaire, K. J., Carpenter, S. K., & Jennings, J. M. (2019). Using prequestions to enhance learning from reading passages: The roles of question type and structure building ability. *Memory*, 27(9), 1204–1213. <https://doi.org/10.1080/09658211.2019.1641209>
- St. Hilaire, K., Chan, J., & Ahn, D. (2023). Guessing as a learning intervention: A meta-analytic review of the prequestion effect. *Psychonomic Bulletin & Review*. <https://doi.org/10.3758/s13423-023-02353-8>
- Storm, B. C., James, K. K., & Stone, S. M. (2022). Pretesting can be beneficial even when using the internet to answer questions. *Memory*, 30(4), 388–395. <https://doi.org/10.1080/09658211.2020.1863990>
- Toftness, A. R., Carpenter, S. K., Lauber, S., & Mickes, L. (2018). The limited effects of prequestions on learning from authentic lecture videos. *Journal of Applied Research in Memory and Cognition*, 7(3), 370–378. <https://doi.org/10.1016/j.jarmac.2018.06.003>
- Vaughn, K. E., & Rawson, K. A. (2012). When is guessing incorrectly better than studying for enhancing memory? *Psychonomic Bulletin & Review*, 19(5), 899–905. <https://doi.org/10.3758/s13423-012-0276-0>
- Vaughn, K. E., Hausman, H., & Kornell, N. (2017). Retrieval attempts enhance learning regardless of time spent trying to retrieve. *Memory*, 25(3), 298–316. <https://doi.org/10.1080/09658211.2016.1170152>
- Wahlheim, C. N., & Jacoby, L. L. (2013). Remembering change: The critical role of recursive reminders in proactive effects of memory. *Memory & Cognition*, 41(1), 1–15. <https://doi.org/10.3758/s13421-012-0246-9>
- Wang, L., & Yang, J. (2023). Learning from errors: Distinct neural networks for monitoring errors and maintaining corrects through repeated practice and feedback. *NeuroImage*, 271, 120001. <https://doi.org/10.1016/j.neuroimage.2023.120001>
- Welhaf, M. S., Phillips, N. E., Smeekens, B. A., Miyake, A., & Kane, M. J. (2022). Interpolated testing and content pretesting as interventions to reduce task-unrelated thoughts during a video lecture. *Cognitive Research: Principles and Implications*, 7(1), 26. <https://doi.org/10.1186/s41235-022-00372-y>
- Willis, R., Erwin, D., & Adelaja, F. (2020). Struggling prior to a teaching event results in superior short-term skills acquisition in novice learners. *Journal of Surgical Education*, 77(1), 34–39. <https://doi.org/10.1016/j.jsurg.2019.07.022>
- Yan, V. X., Yu, Y., Garcia, M. A., & Bjork, R. A. (2014). Why does guessing incorrectly enhance, rather than impair, retention? *Memory & Cognition*, 42(8), 1373–1383. <https://doi.org/10.3758/s13421-014-0454-6>
- Yang, C., Potts, R., & Shanks, D. R. (2017). Metacognitive unawareness of the errorful generation benefit and its effects on self-regulated learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 43(7), 1073–1092. <https://doi.org/10.1037/xlm0000363>
- Yang, C., Potts, R., & Shanks, D. R. (2018). Enhancing learning and retrieval of new information: A review of the forward testing effect. *Npj Science of Learning*, 3(1), 1. <https://doi.org/10.1038/s41539-018-0024-y>
- Yang, J., Zhang, Y., Pi, Z., & Xie, Y. (2021). Students' achievement motivation moderates the effects of interpolated pre-questions on attention and learning from video lectures. *Learning and Individual Differences*, 91, 102055. <https://doi.org/10.1016/j.lindif.2021.102055>
- Zawadzka, K., & Hanczakowski, M. (2019). Two routes to memory benefits of guessing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 45(10), 1748–1760. <https://doi.org/10.1037/xlm0000676>
- Zawadzka, K., Zaborowska, O., Butowska, E., Piątkowski, K., & Hanczakowski, M. (2023). Guessing can benefit memory for related word pairs even when feedback is delayed. *Memory & Cognition*. <https://doi.org/10.3758/s13421-022-01385-0>