Psychonomic Bulletin & Review (2024) 31:411–441
https://doi.org/10.3758/s13423-023-02353-8

THEORETICAL/REVIEW

Guessing as a learning intervention: A meta-analytic review of the prequestion effect

Kyle J. St. Hilaire[1] · **Jason C. K. Chan**[1] · **Dahwi Ahn**[1]

Accepted: 24 July 2023 / Published online: 28 August 2023

Abstract

Giving students test questions before they have learned the correct answers (i.e., prequestions) enhances learning. However, existing research has provided conflicting evidence on whether the benefits of prequestions are *specific* to the initially tested material or if they *generalize* to new, nontested material. In this review, we summarize the literature on the prequestion effect, describe the attention-based account underlying this effect, report a meta-analysis of the magnitude of the specific and general effects, and explore theoretically and empirically relevant moderator variables that influence the size and direction of the prequestion effect. This preregistered meta-analysis demonstrated a moderate specific effect ($g = 0.54$, $k = 97$) but a virtually nonexistent general effect ($g = 0.04$, $k = 91$). Overall, the attention-based account received support from some theoretically relevant moderator analyses. Future researchers are encouraged to conduct theoretically motivated studies to help clarify the mechanisms that underlie the attention-enhancing effects of prequestions and to explore the benefits of prequestions in educational domains to establish the extent to which these effects translate into the classroom.

Keywords Memory · Prequestions · Learning and Instructions · Education

Employing research-supported teaching and learning interventions should promote student learning in the classroom. Because new learning depends on the successful encoding, storage, and retrieval of newly acquired course material, many classroom learning interventions have been devised to target one or more of these stages. One intervention that has received empirical support in recent years is the prequestion effect (PE), whereby asking students to guess the answers to test questions *before* they learn the correct answers (i.e., prequestions) enhances learning outcomes. In this paper, we (1) provide a qualitative review of the literature, (2) describe the dominant theoretical account, (3) report a meta-analysis, and (4) explore moderator variables that evaluate the strength of evidence in support of the theoretical account.

✉ Jason C. K. Chan
ckchan@iastate.edu

Kyle J. St. Hilaire
ksthilaire@chegg.com

Dahwi Ahn
dahn@iastate.edu

[1] Iowa State University, Ames, IA, USA

Prequestions are often used in the classroom as an assessment tool rather than a way to provide direct learning benefits to students. Indeed, prequestions can provide instructors with valuable information about students' academic readiness. First, prequestions are used *diagnostically* to evaluate how much prior knowledge students possess. For example, in content courses that build upon earlier ones, instructors often administer prequestion assessments to determine how much of the prerequisite course material should be reviewed before teaching novel content (Cook, 1984; Herman, 2010; Simkins & Allen, 2000). Second, prequestions may be used *predictively,* as performance on prequestion assessments positively correlates with course grades. Grover et al. (2009) observed that higher scores on a prequestion assessment in a college finance course (i.e., more prior knowledge) were associated with better end-of-semester course grades (see also Diederich et al., 2018). The diagnostic and predictive applications of prequestions can help instructors identify the extent to which students have the requisite knowledge; however, as suggested above, prequestions can also provide direct learning benefits. In the following, we review the relevant literature on the PE, whereby guessing the answers to prequestions enhances learning outcomes.

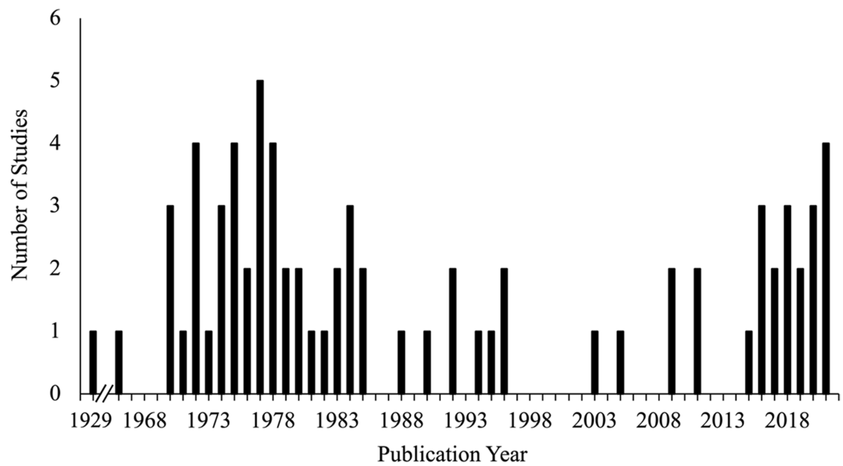


Fig. 1 Number of studies included in the meta-analysis as a function of publication year

Trends and related phenomena

We defined prequestions as asking learners questions before they encounter the study materials, which include both information that would and would not answer the prequestions. For example, an instructor might pose several prequestions to a class before lecturing, but the prequestions would only cover a portion (and likely a small portion) of the information to be covered in the lecture. The key is that for participants to benefit from prequestions, they must *search for and discover* the answers to the prequestions *amongst other relevant materials* that do not answer the prequestions.

Research interest in prequestions has come in noticeable waves. As evidenced by the papers included in this meta-analytic review (see Fig. 1), research interest in prequestions gained momentum during the second half of the 20th century. However, by the end of the century, interest in prequestions waned considerably until their resurgence in the mid-to-late 2000s. Attention towards the PE coincided with a broader rise in interest in other educationally relevant memory phenomena, including retrieval practice, distributed learning, and metacognition (see Bjork et al., 2013; Chan et al., 2018; Dunlosky et al., 2013; McDermott, 2021; Metcalfe, 2009; Pashler et al., 2007). To preview, prequestions are believed to promote learning by enhancing students' *encoding* of the study material. By contrast, other memory phenomena such as the testing effect (McDermott, 2021) bolster learning because they benefit the storage and retention of information over time (i.e., to slow down forgetting).

Despite dozens of studies on prequestions, the PE has not been consistently termed. Prequestions have been referred to as *adjunct prequestions* (Dowaliby, 1992), *inserted questions* (Smith, 1976), *pre-passage questions* (Wilhite, 1983,1984), *advance questions* (Kreiner, 1996), and *pretesting/pretest questions* (Richland et al., 2009; Welhaf et al., 2022). Most of these terms lacked staying power in the literature, except for "prequestions," which was used by the vast majority of researchers from the early days of this literature (Andrew, 1975; Berlyne, 1954a, b; Boker, 1974; Bull & Dizney, 1973; Carpenter & Toftness, 2017; Carpenter et al., 2018; Frase, 1967; Chan et al., 2022; Gustafson & Toole, 1970; Halpain et al., 1985; Hillman, 1979; Kastelic, 1977; Kirschner and Brink, 1979; Memory, 1981; Patrick, 1976; Peeck, 1970; Rickards, 1976a, b; Rickards et al., 1976; Sagaria & Di Vesta, 1978; Shanahan,1986; St. Hilaire et al., 2019; Swenson & Kulhavy, 1974; Wiesendanger & Wollenberg, 1978). Occasionally, researchers had used the term "pretest" in the context of prequestions (e.g., Hartley & Davies, 1976; Papay, 1971; Salmon et al., 1977), but these were exceptions rather than the rule. Moreover, some papers that used the term "pretest" were concerned with the topic of experimental design (i.e., pre–post testing in repeated measures design) rather than prequestion as an memory/educational intervention (e.g., Duncan & Harris, 1974).

Of particular relevance here is that more recently, a number of studies have investigated the phenomenon of the *pretesting effect*, which differs from the *prequestion effect* considered in this review. Somewhat problematically, researchers have sometimes used the terms pretesting and prequestion interchangeably (e.g., James & Storm, 2019; Richland et al., 2009), but we feel that a distinction must be drawn between studies that used the pretesting design and those that used the prequestion design. In *pretesting* effect studies, participants attempt to answer a question and are then provided the answer to *that question* (i.e., feedback). Participants then answer another question, receive

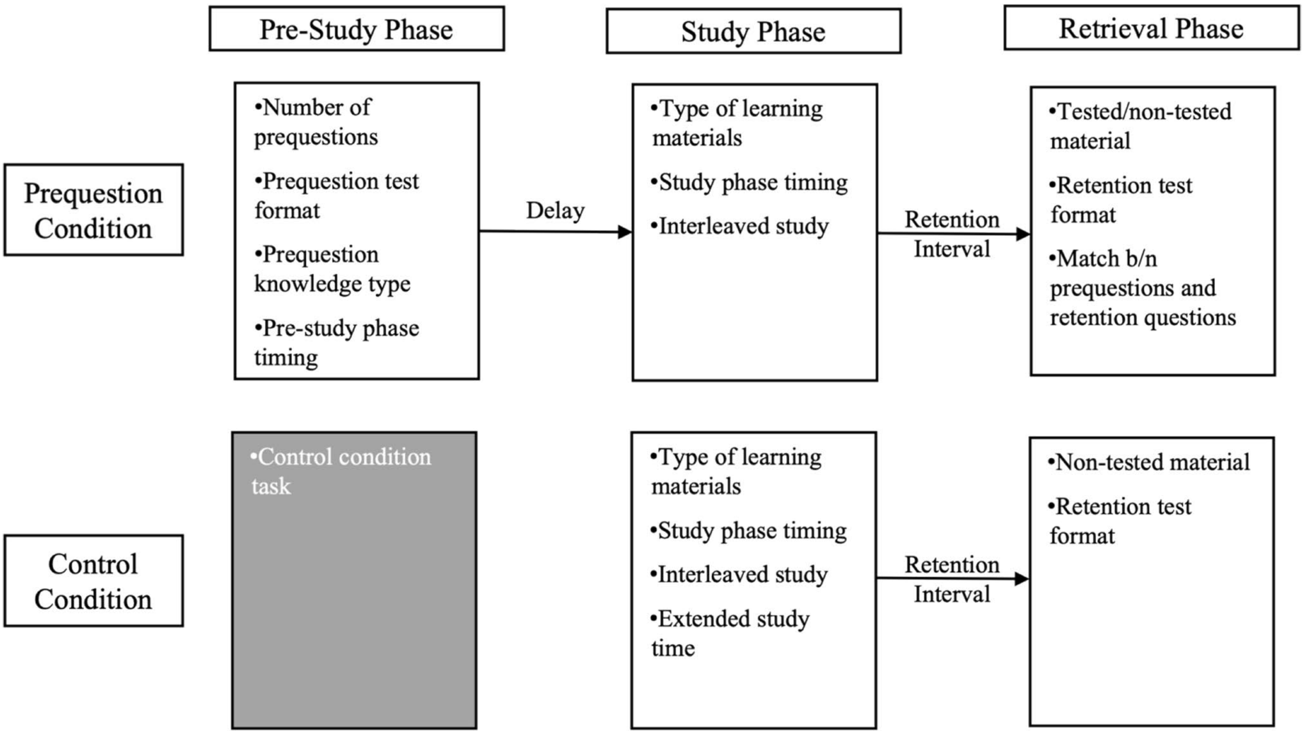


Fig. 2 Prequestion effect paradigm with common moderator variables

the answer to that question, and so on. There are several major differences between the prequestion and pretesting designs (e.g., most pretesting studies implemented no delay between presentation of the question and the answer, whereas there was almost always a delay between answering prequestions and encountering the information relevant to them). But most critically, unlike the prequestion paradigm, the pretesting paradigm does not require participants to search for an answer to the prequestions amongst other relevant materials. Specifically, the prequestion paradigm examines learning through discovery, whereas the pretesting paradigm examines learning through feedback, which makes these fundamentally different learning environments. To draw an analogy from the visual attention domain, prequestion is similar to performing visual search in a complex environment, whereas pretesting is similar to attending to a single stimulus at a time.[1]

[1] An additional paradigm that bears resemblance to the PE is the interpolated testing paradigm, during which participants learn, say, two sets of materials, and between the encoding of these two sets of materials they either answer some (post)questions or they complete some filler tasks. In general, answering interpolated questions promotes learning of the later set of materials (e.g., Ahn & Chan, in press, 2022; Chan et al., 2020, 2022; Manley & Chan, 2019; Wilford et al., 2014).

This review only investigated the learning benefits of prequestions. Readers interested in the *pretesting effect* are directed to the meta-analytic review by Chan et al. (2018) and the qualitative review by Kornell and Vaughn (2016).

The prequestion effect paradigm

In a typical PE study, participants are either assigned to a *prequestion condition*, in which they receive prequestions before the study phase, or to a *no-prequestions control condition*, in which no prequestions are asked before the study phase. Following the study phase, participants in both conditions complete a retention test containing the initially tested items. Variants of this paradigm have demonstrated that participants in the prequestion condition score better on the retention test than participants in the control condition (see Fig. 2).

Below, we provide an overview of the notable manipulations that have been implemented in the prequestion paradigm. We organized this review based on the three phases of its design, including the pre-study phase, the study phase, and the retrieval phase. This overview also serves as a preview to our moderator analyses.

The pre-study phase

As depicted in Fig. 2, only participants in the prequestion condition are provided with prequestions. First, researchers have manipulated the number of prequestions that participants receive, asking as few as one (Johnson, 2015) or two prequestions (James & Storm, 2019; Memory, 1981, 1983), but as many as 41 prequestions (Salmon et al., 1977). Second, although a majority of studies have assessed the PE using short answer prequestions, some have also employed multiple choice prequestions (Dowaliby, 1992; Felker & Dapra, 1975; Little & Bjork, 2011, 2016; Peeck, 1970). Other variants include using factual-style prequestions (i.e., prequestions with verbatim answers provided *directly* in the learning materials; e.g., Bull & Dizney, 1973; Denner & McGinley, 1992; Frase et al., 1970; Janelli & Lipnevich, 2021; Peeck, 1970) and conceptual prequestions (i.e., inference-based prequestions that are not directly answered in the learning materials; e.g., Felker & Dapra, 1975; Hausman & Rhodes, 2018; Hillman, 1979; Kastelic, 1977; Stolz, 1985; Wilkerson, 1982). Third, some researchers have provided participants with a timed pre-study phase (Kirschner and Brink, 1979; Manchester, 1984; Pan et al., 2020; Papay, 1971; Peeck, 1970; Richland et al., 2009; Truog, 1977; Yang et al., 2021), but others have employed an untimed, self-paced pre-study phase (Benya, 1980; Boker, 1974; Khoynejad, 1980; Lee, 2011; Swenson & Kulhavy, 1974; Wiesendanger & Wollenburg, 1978). In addition to manipulating timing during the pre-study phase, researchers have also occasionally manipulated the time interval between the pre-study and study phases (e.g., Little, 2021; Salmon et al., 1977), but these delays are uncommon ($k = 2$ amongst the studies included in this meta-analysis).

Participants in the no-prequestions control condition do not receive prequestions and generally begin the experiment with the study phase. However, some researchers have provided control participants with mock pretests comprising content-*irrelevant* prequestions to equate for time spent during the pre-study phase. For example, in Anderson (1978), participants enrolled in a sex education class either answered sex education prequestions or content-irrelevant prequestions about curricular theory (see also Karjala, 1984). Similarly, Pan et al. (2020) had control participants complete algebra questions to control for time on task before the study phase.

The study phase

As shown in Fig. 2, participants in *both* the prequestion and control conditions complete a study phase. The type of learning materials employed by researchers have included prose passages (Bull & Dizney, 1973; Dickerson, 1988; Frase et al., 1970; Kastelic, 1977; Pressley et al., 1990; Rickards, 1976a, b) and video lectures (Carpenter & Toftness, 2017; James & Storm, 2019; Janelli & Lipnevich, 2021; Kirschner and Brink, 1979; Pan, Schmitt, et al., 2020; Smith, 1996; St. Hilaire & Carpenter, 2020a; Toftness et al., 2018; Yang et al., 2021) on a wide variety of topics, including historical figures such as Joan of Arc (James & Storm, 2019) and William James (Patrick, 1976); geographical topics including Greece (Peeck, 1970) and Yellowstone National Park (Little & Bjork, 2011, 2016; Pan & Sana, 2021); biological concepts such as human development (Pressley et al., 1990), colorblindness (Richland et al., 2009), and vitamins and drugs (Sagaria & Di Vesta, 1978); psychology principles such as autobiographical memory (Toftness et al., 2018) and operant conditioning (Felker & Dapra, 1975); and other topics including signal detection theory (Toftness et al., 2018), vehicular brakes (St. Hilaire et al., 2019), educational history (Bull & Dizney, 1973), and fictional information (Manchester, 1984; McCrudden et al., 2005; Rickards, 1976b; Rickards et al., 1976; Rickards & McCormick, 1988; Snowman & Cunningham, 1975).

Researchers have used study intervals of varying lengths, providing learners with both experimenter-paced (Haimowitz, 1972; Hillman, 1979; Little & Bjork, 2011, 2016; Pan, Schmitt, et al., 2020; Richland et al., 2009; St. Hilaire & Carpenter, 2020a) and self-paced study phases (Denner & McGinley, 1992; Dowaliby, 1992; Feil, 1977; Janelli & Lipnevich, 2021; Johnson, 2015; Kreiner, 1996; Stolz, 1985). Although the study phase typically comprises a single *massed* study interval (e.g., they might answer 10 prequestions and then read an essay), some researchers have interleaved brief pre-study and study phases repeatedly (e.g., two prequestions before each of five short study segments; Carpenter & Toftness, 2017; Feil, 1977; Miyagi, 1995; Mottley, 1972; Patrick, 1976; Shanahan, 1986; Shavelson et al., 1974). Rather than using mock pretests to equate experimental time-on-task across the two conditions (see the prior section of the pre-study phase), some researchers have provided control participants with *extended study time* during the study phase (Little & Bjork, 2011, 2016; Peeck, 1970; Richland et al., 2009; St. Hilaire, 2017; St. Hilaire et al., 2019).

Finally, researchers have implemented retention intervals following the study phase ranging from 5 min (Kealy et al., 2003; Little & Bjork, 2011, 2016; Pan et al., 2020; St. Hilaire, 2017; St. Hilaire et al., 2019; Toftness et al., 2018) to several days (Anderson, 1978; Bull & Dizney, 1973; Karjala, 1984; Salmon et al., 1977; Truog, 1977).

The retrieval phase

During the retrieval phase, participants complete a criterial retention test that assesses their memory for the materials (see Fig. 2). For participants in the prequestion condition, the retention test usually comprises the same questions as the pre-study phase (i.e., *tested material*); however, it is common to also ask *new* questions (i.e., *nontested material*). Researchers have also varied the format of the retention questions. Although the retention questions generally match the format of the prequestions (e.g., multiple choice for both), some researchers have switched question formats across phases (e.g., multiple choice prequestions and short answer retention questions; Little & Bjork, 2011, 2016). Because control participants do *not* receive prequestions, all questions during their retrieval phase comprise nontested material; moreover, the "match" in question formats from pretest to retrieval is irrelevant for control participants.

The specific and general effects of prequestions

The magnitude of the PE is determined by comparing how well participants in the prequestion condition perform at the retention test against those in the control condition. Thus, there are two questions of interest: (1) do participants in the prequestion condition outperform participants in the control condition when they answer the *same tested questions* during the retrieval phase (i.e., is there a significant Specific Prequestion Effect?), and (2) do participants in the prequestion condition outperform participants in the control condition when they answer *new, nontested questions* during the retrieval phase (i.e., is there a significant General Prequestion Effect?).

There is considerable evidence that the PE is *specific*. When participants receive prequestions, they typically demonstrate a robust advantage relative to control participants when they answer the tested questions during retrieval (Haimowitz, 1972; James & Storm, 2019; Richland et al., 2009; Smith, 1976; St. Hilaire & Carpenter, 2020a). Hartley and Davies (1976) emphasized the importance of the specific effect and highlighted three key functions that prequestions serve: prequestions introduce students to the to-be-learned materials, provide students with an organizational framework to learn that material, and highlight what students do not know. Hartley and Davies (1976) thus argued that prequestions should be used *selectively* to only highlight the important details that students need to learn. For example, if students are expected to know and learn specific facts and figures, then instructors should write prequestions that highlight those details.

In contrast, whether the PE reliably *generalizes* to nontested material has yet to be ascertained. Although several studies have demonstrated neither a benefit nor detriment of prequestions on nontested material (Bull & Dizney, 1973; Felker & Dapra, 1975; Hausman & Rhodes, 2018; James & Storm, 2019; Kirschner and Brink, 1979; Pressley et al., 1990; St. Hilaire & Carpenter, 2020a; Toftness et al., 2018; Wilkerson, 1982), some studies have shown that answering prequestions enhanced learning of the nontested material—a *positive* general PE (Carpenter & Toftness, 2017; Denner & McGinley, 1992; Lee, 2011; Little & Bjork, 2011, 2016; McCrudden et al., 2005; Osman & Hannafin, 1994; Pan, Schmitt, et al., 2020; St. Hilaire, 2017; St. Hilaire et al., 2019), whereas others have found the opposite—a *negative* general PE (Boker, 1974; Frase et al., 1970; Haimowitz, 1972; Hillman, 1979; Patrick, 1976; Peeck, 1970; Sagaria & Di Vesta, 1978; Smith, 1976; Strollo, 1972; Truog, 1977). Thus, the present quantitative review would help to establish the overall magnitude of the general effect and determine what manipulations, if any, tend to produce a positive or a negative general PE.

An attention-based account of prequestions

Despite many demonstrations establishing the direct benefits of prequestions, the theoretical mechanisms underlying the PE are still poorly understood. Most studies in this literature have been driven by applied questions, and researchers have typically offered post hoc explanations that speculate about why prequestions enhanced learning in specific experiments. Despite the absence of an organizing theoretical framework, extant explanations have coalesced around a broad attention-based account in which prequestions influence *how* learners engage with the study materials. That is, answering prequestions is believed to enhance learning during the study phase because those questions drew participants' attention to the relevant material. However, because attentional processes are generally *not* measured in the prequestion literature, it can be difficult to generate testable hypotheses for the attention-based account. Studies that examined attention during the study phase have sometimes provided support for the idea that answering prequestions promote learners' attention toward relevant materials (Little & Bjork, 2016; McCrudden et al., 2005; but see Morasky & Willcox, 1970; Welhaf et al., 2022). However, a key issue with the attention account is that it does not address *why* prequestions promote such attentive encoding, and in the absence of such an understanding, arguing that "prequestions enhance learning because they promote attention" borders on restating the results (cf. Richland et al., 2009, who examined whether prequestions promote learning due to processes beyond merely highlighting important concepts). To this end, we have identified two potential *precursors* to the attention account: a *memory* precursor account and a *curiosity* precursor account.

Because descriptions of these two precursor explanations have been somewhat unsystematic, there is some subjectivity regarding how the accounts can be interpreted and what predictions might emerge from them. In our view, the memory and curiosity precursor accounts are *not* exclusive or exhaustive, and individual readers might dispute some of the predictions attributed to the accounts here. To ensure that our predictions were driven purely by our interpretation of the existing literature, we had preregistered them on the Open Science Framework (OSF) before conducting any moderator analyses, and in cases where new predictions have arisen following the preregistration, we labeled the relevant analyses as exploratory.

A consideration of the mechanisms underlying the *general effect* makes this theory development even more challenging than for the *specific effect*, perhaps because there is little consensus on whether the general PE is even a reliable phenomenon (for a discussion of this point, see James & Storm, 2019). If there is no general effect (or a very weak one), it can be difficult to generate predictions. As such, the predictions outlined for the general effect are mostly based on the experimental designs and hypotheses provided by previous researchers interested in the general PE. For organizational and exposition purposes, these predictions were then fit to the two precursor accounts, which we believe to be the best approach at this nascent stage of theoretical development for the general PE. An important goal of this review was to encourage the field to progress beyond empirical demonstrations of the PE and to more forcefully consider its mechanisms.

Below, we first introduce the main assumptions and background of each account; we then discuss its predictions for relevant moderator variables for the specific PE and general PE, respectively.

Memory precursor account

The memory precursor account suggests that answering prequestions directs learners' attention to the tested material during the study phase because learners *remember* the prequestions (Frase, 1967; Hamaker, 1986; Hamilton, 1985; Hannafin & Hughes, 1986; St. Hilaire & Carpenter, 2020a). When participants encounter information relevant to the prequestions, the familiarity of that information provides a framework through which participants can incorporate or update their knowledge, thereby enhancing learning. Similar to this idea, when considering the potential benefits of answering prequestions on learning, Richland et al. (2009) suggested that creatively producing an answer to a prequestion may "create fertile ground for later encoding of the answer" (p. 252). This concept might be similar to the idea that people with greater existing knowledge about a topic tend to produce better new learning than those with poorer existing knowledge (Chan et al., 2017; Witherby & Carpenter, 2022). Conversely, if learners cannot retain the prequestions, their attention would not be biased toward the tested details during the study phase. St. Hilaire and Carpenter (2020a) asked participants to take notes while watching a video lecture. Critically, they wanted to determine whether the magnitude of the PE was larger for tested material that was *included* in participants' notes than for tested material that was *omitted* from participants' notes, with the assumption being that participants would be more likely to take notes for the material they found relevant to the prequestions *that they remember* than the ones that they do not remember. Across two experiments, St. Hilaire and Carpenter (2020a) found evidence supporting this prediction. Specifically, the PE was contingent on learners discovering and writing down information relevant to the prequestions. Performance on the tested material that participants did *not* write down was similar to that of the control participants, and if participants could not find the answer to a prequestion during the study phase, their performance was not significantly different from chance (St. Hilaire & Carpenter, 2020a, Experiment 4). Pan et al. (2020) found that participants' ability to recall the prequestions positively predicts the size of the PE, which suggests that remembering the prequestions is a key factor underlying the PE. Together, these data suggest that answering prequestions might direct participants' attention to the tested material because learners remember the prequestions during the encoding phase.

Predictions for the specific effect

Based on the memory precursor account, the magnitude of the specific effect should be dependent on memory strength for the prequestions. Specifically, one can expect a greater specific effect of prequestions under situations in which learners are better able to remember the prequestions, such as when (1) learners are better able to encode the prequestions, (2) the risk of forgetting the prequestions is minimized, or (3) the answers to the prequestions are familiar and easy to find. These principles and their manifestations as moderator variables are displayed in Table 1.

With regard to the first assumption above, a relevant moderator is whether learners are required to guess the answer to each prequestion. Because guessing forces participants to produce a reasonable answer to a prequestion, one needs to access stored knowledge relevant to the question. This type of semantic retrieval (i.e., a form of elaboration; see Craik & Tulving, 1975) should promote memory of the prequestions to a greater degree than reading the prequestions alone. Based on this reasoning, we expected that guessing would enhance the specific effect. The second prediction specifies that memory for the prequestions should be positively related to the length of time that participants spend encoding each prequestion.

The subsequent predictions are rooted in the idea that one can expect better retention of prequestions (and a stronger specific effect) when the risk of forgetting (of the

Table 1 Predictions for the moderators

Moderator	Predictions for the Specific Effect		Predictions for the General Effect	
	Memory Precursor Account	Curiosity Precursor Account	Memory Precursor Account	Curiosity Precursor Account
Prequestion Guess	Larger PE for guess than read-only prequestions	Larger PE for guess than read-only prequestions		
Length of Time per Prequestion	Larger PE with more time per prequestion			
Number of Prequestions	Larger PE with fewer prequestions		Larger PE with more prequestions	
Interleaved Study	Larger PE for interleaved than massed study		Larger PE for massed than interleaved study	
Type of Learning Materials			Larger PE for audiovisual than text materials	
Relation between Tested and Nontested Material			Larger PE for related than unrelated retention questions	
Prequestion Test Format	Larger PE for multiple choice than short answer prequestions	Larger PE for short answer than multiple choice prequestions		Larger PE for short answer than multiple choice prequestions
Prequestion Knowledge Type		Larger PE for conceptual than factual prequestions		Larger PE for conceptual than factual prequestions
Prequestion Performance		Larger PE for worse prequestion performance		

PE = Prequestion Effect. *Match between Prequestions and Retention Questions* was originally pre-registered on OSF as a theoretically relevant moderator variable for the specific account (i.e., larger PE for matched vs. mismatched question formats). However, in hindsight, this moderator does not have clear relevance for either of the precursor accounts and we have therefore removed it. Moreover, delay between the pre-study phase and study phase was predicted to be a theoretically relevant moderator for the memory precursor account, but this moderator was not included in any analyses because not enough studies have implemented a delay.

prequestions) is minimized. To this end, asking fewer prequestions (as opposed to asking more prequestions) would make it easier for participants to retain the prequestions (Hamilton, 1985). Similarly, interleaved prequestion designs (as opposed to massed prequestions) minimize the number of prequestions that participants need to retain at any one time. For example, it should be easier for participants to remember two prequestions before each of eight interleaved study blocks than needing to remember 16 prequestions before one massed study block (with the interleaved design reducing both set size and retention interval, see Pan et al., 2020). Participants should also be less likely to forget the prequestions when the delay between the pre-study phase and study phase is reduced; therefore, the magnitude of the specific effect would be negatively correlated with the duration of this delay. However, after collecting the data, it became clear that this variable could not serve as a moderator because almost no studies have implemented a delay (cf. Little, 2021; Salmon et al., 1977).

The last assumption of the memory precursor account is that being familiarized with the prequestioned concepts would facilitate the identification of relevant concepts. By this reasoning, multiple choice prequestions should produce a stronger PE than recall-based prequestions, because in the former case, participants are exposed to the correct answers among incorrect foils during the pre-study phase (Butler, 2018). If participants encode the stems of multiple choice prequestions along with their answer options (see Little & Bjork, 2011, 2016), then the familiarity of the answer afforded by multiple choice prequestions should promote its discovery compared with short answer prequestions.

Predictions for the general effect

The memory precursor account specifies that learners' attention is directed toward the tested material because they remember the prequestions. Although perhaps counterintuitive, there are reasons to believe that having *weaker* memories for the prequestions might benefit learning of the nontested material and lead to a larger general PE. When memory for the prequestions is *weakened*, participants might attend more broadly during the study phase rather than directly seek out materials that answer the prequestions, thereby enhancing encoding of the nontested material relative to participants in the control condition. This is not to say that participants should have *no* memory for the prequestions (as no memory would likely match the attentional processes of the control participants), but *weakened* memory for the prequestions would require participants to place greater reliance on *gist* rather than *verbatim* memory for the prequestions (Reyna & Brainerd, 1995). Below, we illustrate how this principle might manifest across several moderators for the general effect, and these predictions are tallied in Table 1.

Several predictions can be formulated based upon the logic explained above. First, the number of prequestions should negatively relate to the magnitude of the general effect. Asking participants a greater number of prequestions should weaken memory for each prequestion during the study phase (Hamilton, 1985). If so, participants might be more likely to attend to material that is conceptually related to the prequestions, thereby increasing the likelihood of a general PE. Second, because interleaved presentation of the prequestions should strengthen their memory relative to massed presentation, we predict that the latter would produce a stronger general PE. Finally, longer delays between the pre-study and study phases would attenuate memory for the prequestions, forcing participants to rely on gist rather than verbatim memory during the study phase (see Reyna & Brainerd, 1995).

Participants might be more likely to rely on gist memory when learning from audiovisual materials than from text-based materials. Due to the serial presentation nature of audiovisual materials (unlike text materials which are presented all at once), learners cannot scan for keywords. Thus, participants might be particularly likely to incidentally glean the answers to the nontested material while learning from audiovisual study materials (see Carpenter & Toftness, 2017).

Finally, we hypothesized that the *relation* between the tested and the nontested material is relevant for the general effect. Specifically, information related to the prequestions should be more likely to remind participants of the prequestions than information unrelated to the prequestions. This reminding should promote integration of the related, nontested material with memory of the prequestions (Chan, 2009; Chan et al., 2006; Hintzman, 2011; Jacoby et al., 2015). Therefore, we expected a greater general effect when the retention questions probe for material that is related to the tested material (see Little & Bjork, 2011, 2016).

Curiosity precursor account

The curiosity precursor account states that prequestions make learners *curious* to learn the correct answers, and this increased curiosity directs attention to the learning material during the study phase (Berlyne, 1954a, b, 1962; Bull & Dizney, 1973; Patrick, 1976; see also Metcalfe & Finn, 2011; Potts et al., 2019; Wade & Kidd, 2019). As we discussed earlier, the prequestion literature has largely been empirically driven, and relatively little attention has been paid to the cognitive mechanisms underlying the PE. Compared with the memory precursor account, which itself lacks specificity, the curiosity precursor account has received even less scrutiny. In fact, researchers often use the term "curiosity" in a somewhat colloquial manner, with little exploration of what curiosity constitutes, and taking curiosity as a facilitator of learning as a given (for a review of research on curiosity and learning, see Grossnickle, 2016; Gruber et al., 2016). Here, we attempt to provide a preliminary framework through which to consider how encountering prequestions might pique learners' curiosity, and how curiosity might promote attentive encoding of the study material.

A central hypothesis in the curiosity literature is that novelty, prediction errors, and the detection of knowledge gaps drive curiosity (Gottlieb et al., 2013; Marvin & Shohamy, 2016), which stimulates information-seeking or exploratory behaviors in learners to fill those gaps. Applying these ideas to the PE, one interpretation is that reading (and attempting to answer) prequestions helps participants to discover their knowledge gaps. This realization in turn elicits curiosity both about the specific information being queried and perhaps the overall topic being covered (Knobloch et al., 2004). When participants are then provided with an opportunity to learn something about that topic, they actively seek information in the learning materials that might help fill their knowledge gaps (Litman et al., 2005), which is manifested as the PE.

For example, in addition to prequestions, Berlyne (1954b) asked participants to indicate for which items they most wanted to learn the answers. Participants showed a larger PE for items that they were curious about than those they were not. Although this result is consistent with the curiosity precursor account, the evidence is correlational: it is possible that the prequestions that participants marked as being curious to learn were qualitatively different from those that they did not mark. Bull and Dizney (1973), however, provided empirical evidence consistent with a curiosity precursor account. Specifically, they *varied* the degree of curiosity induced by each prequestion through a *conceptual conflict* manipulation (see Berlyne, 1962; Engel & Randall, 2009). Specifically, Bull and Dizney (1973) compared low conflict (i.e., low curiosity; e.g., "Why did the Balinese send their daughters to be educated as school teachers?") and high conflict prequestions (i.e., high curiosity; e.g., "If teachers are generally viewed as middle class, why was it the Balinese of high caste who sent their daughters to be educated?") to a control condition. Critically, Bull and Dizney (1973) demonstrated a PE for high conflict but not low conflict prequestions. Thus, attempting to answer prequestions might drive learners to seek out answers to these questions because of the curiosity that they elicit.

Loewenstein (1994) has argued that curiosity arises due to the gaps in one's knowledge. Participants will occasionally have some relevant prior knowledge before the study phase begins, but performance on the prequestions varies widely. In some studies, prior knowledge does not differ from chance (e.g., Little & Bjork, 2011, 2016), whereas in other studies, participants have shown high degrees of prior knowledge (e.g., Anderson, 1978). Because prequestions force participants to confront what they do not know, it is possible that participants will experience greater curiosity when they are metacognitively *aware* of their knowledge gaps.

Predictions for the specific effect

If learners focus their attention on the tested material because they are curious to learn the answers, then making participants more curious should enhance the magnitude of the specific effect. Below, we consider moderators that are relevant to this notion, and the predictions are tabulated in the rightmost column of Table 1.

Recent research has suggested that committing prediction errors are essential to eliciting curiosity-dependent learning (Gruber et al., 2014; Marvin & Shohamy, 2016; Oudeyer et al., 2016). Because guessing forces participants to confront what they do and do not know, requiring participants to guess the answers to prequestions should make them more curious about the prequestion—and more likely to demonstrate a PE—than if they were simply presented with the prequestion but were not required to produce a response.

The type of prequestion asked during the pre-study phase could also influence curiosity. Specifically, short answer prequestions that require generation might be more curiosity-inducing than selection-based prequestions such as multiple choice. Because short answer prequestions require participants to rely on their own intuitions to generate an answer (and experience a prediction error), it should elicit greater curiosity than multiple choice prequestions.

Similarly, the specific effect might be larger for conceptual than factual prequestions. Unlike factual prequestions, which assess specific facts or details (e.g., the *name* of the largest geyser), conceptual prequestions (e.g., *why* a geyser erupts) might be more likely to activate prior knowledge related to the tested concepts (even if participants do not know the answer; see Hausman & Rhodes, 2018), thereby encouraging exploratory behavior during the study phase.

Finally, although participants do not receive feedback during the pre-study phase, they should have reasonably good insight into their prequestion performance. If so, participants might be more motivated to discover the answers to the prequestions when they score worse on the pretest than when they do well (see Loewenstein, 1994). However, it remains unclear *how much* prior knowledge is sufficient to enhance learning outcomes. For example, does poor performance on the pretest (i.e., 0% accuracy) lead to more curiosity than moderate pretest performance, or should learners have at least *some* degree of prior knowledge (but not know all of the answers) for curiosity to be piqued by prequestions?

Predictions for the general effect

In the following, we outlined moderator variables that we believe would influence the likelihood/magnitude of the general PE from the perspective of the curiosity precursor account, and a list of these predictions can be seen on the right side of Table 1. Once again, given the underspecificity of the curiosity precursor account, readers might not always agree with our predictions. We fully acknowledge this possibility, but we also believe that these predictions are logically sound and can serve as critical tests of the curiosity hypothesis at the present state of the literature.

In general, the notion that governed our predictions is that manipulations that help broaden learners' attention during the study phase should increase the likelihood of a general effect. First, short answer prequestions, as opposed to multiple choice prequestions, typically require more extensive comprehension and the consideration of multiple concepts to answer, which might help elicit greater curiosity to a broader set of ideas, thereby promoting the learning of materials beyond just the tested concepts. Indeed, St. Hilaire et al. (2019) demonstrated a general PE when they asked participants to generate responses to short answer prequestions (e.g., "What is the difference between mechanical brakes and hydraulic brakes?"). Although their short answer prequestions were specifically written to encourage the comprehension of disparate concepts (e.g., understanding of mechanical brakes *and* hydraulic brakes), this result might generalize to other short answer prequestions, such that one might expect a greater general PE for short answer than multiple-choice prequestions.

By similar logic, conceptual prequestions are believed to enhance curiosity because they are more likely to activate prior knowledge than factual prequestions, and conceptual questions might also encourage processing of the *nontested material*. For example, Hausman and Rhodes (2018) asked participants to answer conceptual prequestions such as "What is the consequence of higher levels of carbon dioxide in the atmosphere?", which require several degrees of inference. Specifically, more carbon dioxide raises temperatures, higher temperatures lead to warming, warming melts glaciers, and melted glaciers elevate the sea level. Because conceptual prequestions often require elaboration and consideration of the relations across disparate concepts, they might trigger curiosity-based exploratory behaviors and encourage learners to seek out answers to the nontested material more than factual prequestions.

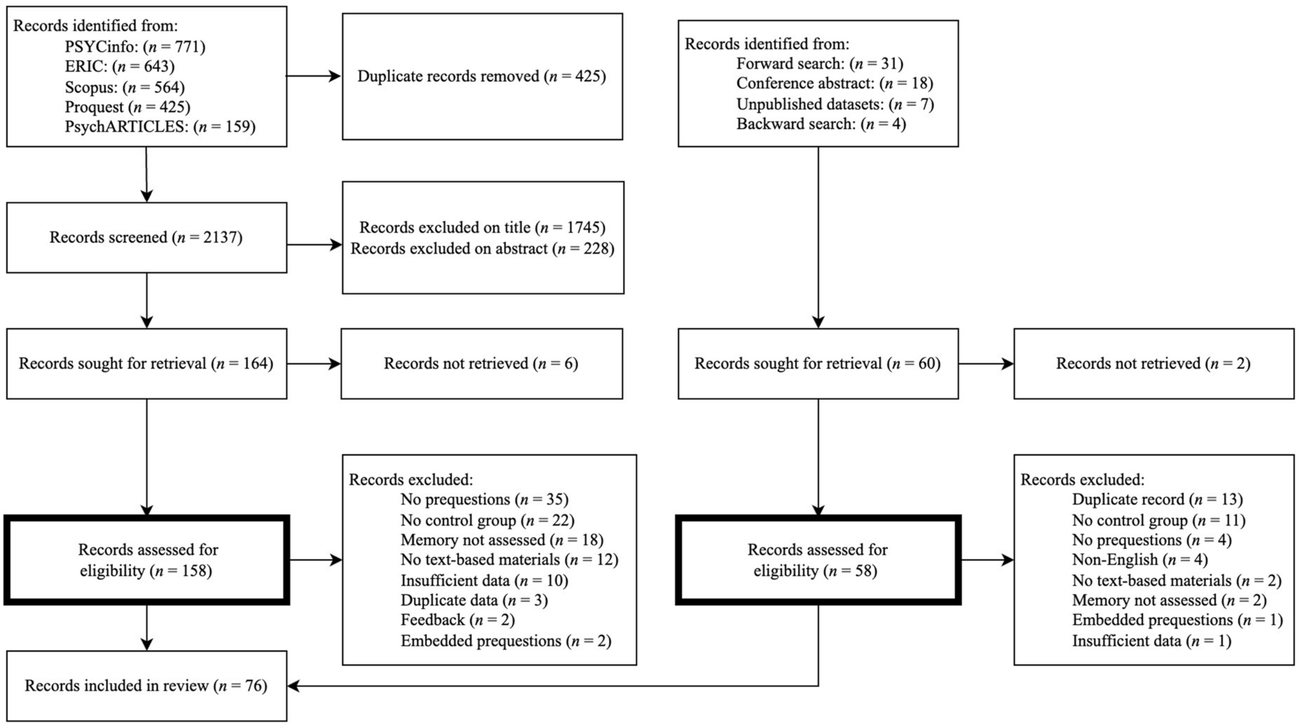


Fig. 3 PRISMA diagram outlining search process

The current study

The most important goal of this paper was to establish the overall meta-analytic effect sizes for both the specific and the general PEs. A secondary goal was to evaluate the evidence for the memory and curiosity precursor accounts. Finally, a quantitative summary of the literature provides the opportunity to identify areas of research that deserve further exploration.

Method

Data availability

All methods and predictions underlying this meta-analysis were pre-registered on the Open Science Framework (OSF; https://osf.io/69jqx/?view_only=2c4e3eb2b74943dd93eeccfdb4af69a6). Moderator variables and predictions that were included *after* the data were analyzed are provided in a revision document that was uploaded to OSF. All Boolean search strings and the data file are available online.

Literature search

Literature searches were conducted in two rounds. The initial literature search was conducted in March 2021 to establish the breadth of the literature, and the final literature search was completed in August 2021. Literature search was carried out using PsycINFO, the Education Resources Information Center (ERIC), Scopus, ProQuest Dissertations and Theses, and PsycArticles databases using Boolean search strings (see https://osf.io/69jqx/?view_only=2c4e3eb2b74943dd93eeccfdb4af69a6). Specific search terms included "prequestion*," "pretesting," "advance* question*," "adjunct* question*," "question* position*," and "embed* question*." Abstract books from three relevant research conferences (i.e., The Annual Meeting of the Psychonomic Society, The Society for Applied Research in Memory and Cognition [SARMAC], and The American Educational Research Association [AERA]) were sourced for other records. Backward literature searches were then conducted for all surviving records using the References section of each paper, and forward literature searches were conducted using Google Scholar to find other potentially relevant but missed records. Finally, nine research groups who have recently published on the PE were contacted to inquire about other in-press or unpublished datasets. All but two research groups responded to the email request and five groups provided additional data. Figure 3 shows a PRISMA diagram outlining this search process, exclusion decisions, and the number of records included in the review. The first author examined all records and triaged obvious ones for exclusions. Following the initial submission of the manuscript, the three authors reexamined all excluded entries in the "records assessed for eligibility" stage in the PRISMA diagram (highlighted by a thick border) to ensure that exclusions were applied properly to these 142 records. Two of these (Hollen, 1970; Rauscher, 1978) were deemed retainable through extensive discussions amongst the three authors. Consequently, exclusion agreement was at 98.6%. During this review process, we also identified one extra paper through the forward search process (Jersild, 1929).

Inclusion and exclusion criteria

Studies were selected for inclusion in the meta-analysis if participants in at least one experimental condition received prequestions before the study phase. If participants *did* receive

prequestions, then studies were further evaluated on eight criteria to establish their eligibility.

1. The magnitude of the PE is computed by comparing performance on the retention test for participants in a prequestion condition to those in a control condition. Without a control condition, there is no baseline from which to compute a PE. Thus, studies that did not include a no-prequestions control condition were excluded (e.g., Al Rasheed, 2014; Berlyne, 1966; de Lima & Jaeger, 2020; Diederich et al., 2018; Fowler & Lamberg, 1979; Geller et al., 2017; Morasky & Willcox, 1970; Noakes, 1969; Pan & Sana, 2021; Sana et al., 2020a, b; Wilhite, 1983). A control condition is defined as situations in which participants studied the target material without having completed other forms of activities that might otherwise affect their performance (e.g., answering postquestions, answering interspersed adjunct questions, preexposing to a subset of the study material via learning objective statements). Therefore, if a study compared participants who answered prequestions with participants who answered postquestions, but there was no pure control condition under which participants simply studied the material, the study would be excluded. A similar idea applied to studies that produced within-subjects estimates of the prequestion effect. Here, performance in the control condition must not be influenced by participants having answered prequestions. For example, if participants studied a single passage and received prequestions, and the criterial test featured both tested and nontested items from that same passage, performance on the nontested items would not satisfy our criterion for a control condition, because answering prequestions on a portion of the passage might influence how participants encode the remainder of the passage—this was the purpose for investigating whether there is a general prequestion effect. For a within-subjects study to be included in our meta-analysis, participants must have studied two unrelated passages and encountered prequestions on only one passage (with a read-only passage serving as control).
2. Studies were excluded if performance on a retention test was not provided. This might seem like an obvious exclusion; however, a handful of researchers have explored the PE *not* to determine how prequestions influence subsequent learning, but to investigate how prequestions affect *encoding* processes during the study phase (e.g., eye movement effects; Morasky, 1972), perceptions regarding the utility of prequestions in the classroom (e.g., Brodzinski, 1983; Pan et al., 2020; Witham & Linehan, 1995), and how prequestions can be used to measure students' prior knowledge (e.g., Cook, 1984; Simkins & Allen, 2000).
3. As described earlier in this proposal, the prequestion effect is distinct from the pretesting effect in that the former requires participants to *discover* the answers to prequestions among irrelevant material, whereas the latter does not (e.g., Grimaldi & Karpicke, 2012; Hays et al., 2013; Knight et al., 2012; Vaughn et al., 2017). A meta-analysis of the pretesting effect has already been reported elsewhere (Chan et al., 2018).
4. An important aim of this meta-analysis was to examine the extent to which answering prequestions can boost memory performance and inform educational practice; thus, we focused our analyses on studies that use content-rich learning materials similar to those typically used in education (e.g., prose passages, video lectures). We therefore excluded studies that used paired associates (e.g., Feiman, 1974; Watts, 1974), images (e.g., Hartley et al., 1970; MacLachlan & Jalan, 1985), nonverbal tasks (e.g., Raven's Progressive Matrices; Bauer, 1976), and materials or tasks that did not speak directly to memory performance such as mathematical computations (e.g., Dalton & Goodrum, 1991; Sana et al., 2020a, b).
5. Among prequestion studies that *have* used content-rich materials, we excluded the studies that provided participants with corrective feedback during the pre-study phase (Latimier et al., 2019; White, 1981), because feedback removes the need for participants to discover the answers to the prequestions later. If a study had *manipulated* feedback, we retained data for the condition(s) in which participants did not receive feedback (e.g., Little & Bjork, 2016; Patrick, 1976).
6. Studies were excluded if prequestions were *embedded* during the study phase rather than provided independently during a pre-study phase (Hill, 1986; Leonard, 1987; Sinnott & Alderman, 1977; Washburne, 1929). Embedded prequestions are provided to participants *during* the study phase, meaning participants do not need to retrieve the prequestions from long-term memory. Addi-

tionally, because embedded prequestions are introduced *during* the study phase, participants might interact with embedded prequestions differently than prequestions that were provided during an independent pre-study phase.

7. Studies not available in English were excluded from the meta-analysis to avoid translation errors (Furumoto and Matsumi, 2009; Shinogaya, 2011; Tanaka & Miyatani, 2015; van Aalst, 2021).
8. In a handful of instances, a dataset has been published in more than one format (e.g., Carpenter et al., 2018; Little, 2011; Little & Bjork, 2011, 2016; Rahman, 2017; Swenson & Kulhavy, 1973, 1974). To ensure that there were no duplicates in the meta-analysis, each dataset was included only once. In each case, records from the highest level of peer review were retained (e.g., a peer-reviewed journal article superseded a thesis).
9. Researchers must have provided sufficient data to compute effect size estimates for the specific effect and/or the general effect. Whenever possible, effect size estimates were computed using descriptive data (e.g., *M*, *SD*, *SE*, *N*). For cases in which descriptive data were not provided in tabular or in-text formats, relevant descriptive information was extracted from sufficiently detailed graphical displays using raster-based measurements. However, when descriptive data were unavailable, inferential statistics (e.g., *t*, *F*) were transformed into effect size estimates when possible. To minimize the number of excluded datasets in this review, we attempted to contact as many authors as possible to request missing data.

The final set of studies in the meta-analysis included a total of 76 records comprising 42 journal articles, 24 unpublished theses/dissertations, six unpublished conference presentations, and four unpublished datasets. Altogether, data from 7,902 participants contributed to the analyses for the specific effect and 7,705 participants contributed to the general effect.

Coding process

All records were coded by two knowledgeable independent raters. The first author served as the primary coder and trained the second coder, who was a senior-level undergraduate student. Both raters extracted details (e.g., authors, publication year, record type) germane to the moderator variables and effect sizes. Agreement in coding was 96% and discrepancies were resolved through discussions.

Coding of theoretically relevant moderators

We first describe how we coded moderator variables that were relevant to the memory and curiosity precursor accounts. These predictions were described in the previous section and summarized in Table 1. Briefly, the variables of prequestion guess, interleaved study, prequestion test format, type of learning materials, relation between tested and nontested materials, and prequestion knowledge type were coded categorically. In contrast, the variables of length of time per prequestion, number of prequestions, and prequestion performance were coded continuously. Details about how we coded these moderators can be found in Supplemental Materials.

Coding of empirically relevant moderators

We also examined the influence of several variables that were not theoretically relevant for the attention-based account. These variables were explored as empirically relevant moderators because they were of particular methodological or practical interest, and they included publication year, research design (i.e., within- vs. between-subjects), participant age group (i.e., grade-school children vs. adults), control condition task (i.e., mock pretest vs. extended study time vs. study-only), pre-study phase timing (i.e., experimenter-paced vs. self-paced), study phase timing (i.e., experimenter-paced vs. self-paced), reading speed (i.e., duration of study phase/length of study materials), retention interval between study phase and retrieval phase (i.e., duration of retention interval in min and logarithmically transformed), retention test format (i.e., multiple-choice vs. short answer), the match between prequestions and retention questions (i.e., match vs. mismatch), and exclusion of accurate prequestion guesses (i.e., including vs. excluding accurately answered prequestions in data from the retention test). For the last moderator, an addition category of *undefined* was included for studies that did not specify whether they had retained the accurately answered prequestions, which indicates the degree of prior knowledge of the learning materials. To be clear, we believe that excluding the correctly answered prequestions is the only way to ensure that any benefit of prequestions can be attributed to learning during the experiment (and is not conflated with prior knowledge; see Richland et al., 2009). Details about how we coded these variables can be found in Supplemental Materials.

Effect size calculations

The two dependent variables of interest were the specific effect (i.e., performance on the tested material) and the general effect (i.e., performance on the nontested material). Specifically, effect sizes (i.e., Cohen's *d*s) were computed by calculating the standard mean difference between the prequestion condition and the control condition and dividing this difference by the pooled standard deviation. Because

Cohen's *d* can overestimate the size of the true effect when samples sizes are small, all Cohen's *d* effect sizes were converted to Hedge's *g* using the following equation:

$$g = d\left(1 - \frac{3}{4\,(df) - 1}\right) \quad (1)$$

where *d* is Cohen's *d* and *df* is $n_1 + n_2 - 2$ for the prequestion condition and the control condition.

For 13 records, there were insufficient data to calculate effect size estimates. Whenever possible, we imputed standard deviation (*SD*) using data provided in a given study (e.g., we use *SD* from the prequestion phase to substitute for the *SD* during the retrieval phase; Rauscher, 1978; Rickards et al., 1976). In all cases, the researchers provided means for the prequestion and control conditions but no estimates of variability. Because variability for these records could not be computed using inferential statistics (e.g., precise *t values* were not reported; *F* values conflated more than one experimental condition), the average standard deviation for all remaining studies was used to impute variability estimates. Different standard deviation estimates were used to impute missing variability data for performance in the control condition (*SD* = .17), performance on the tested material in the prequestion condition (*SD* = .19), and performance on the nontested material in the prequestion condition (*SD* = .18). Effect sizes that were computed using average variability estimates are marked with an asterisk in the forest plots depicted in Figs. 4 and 5.

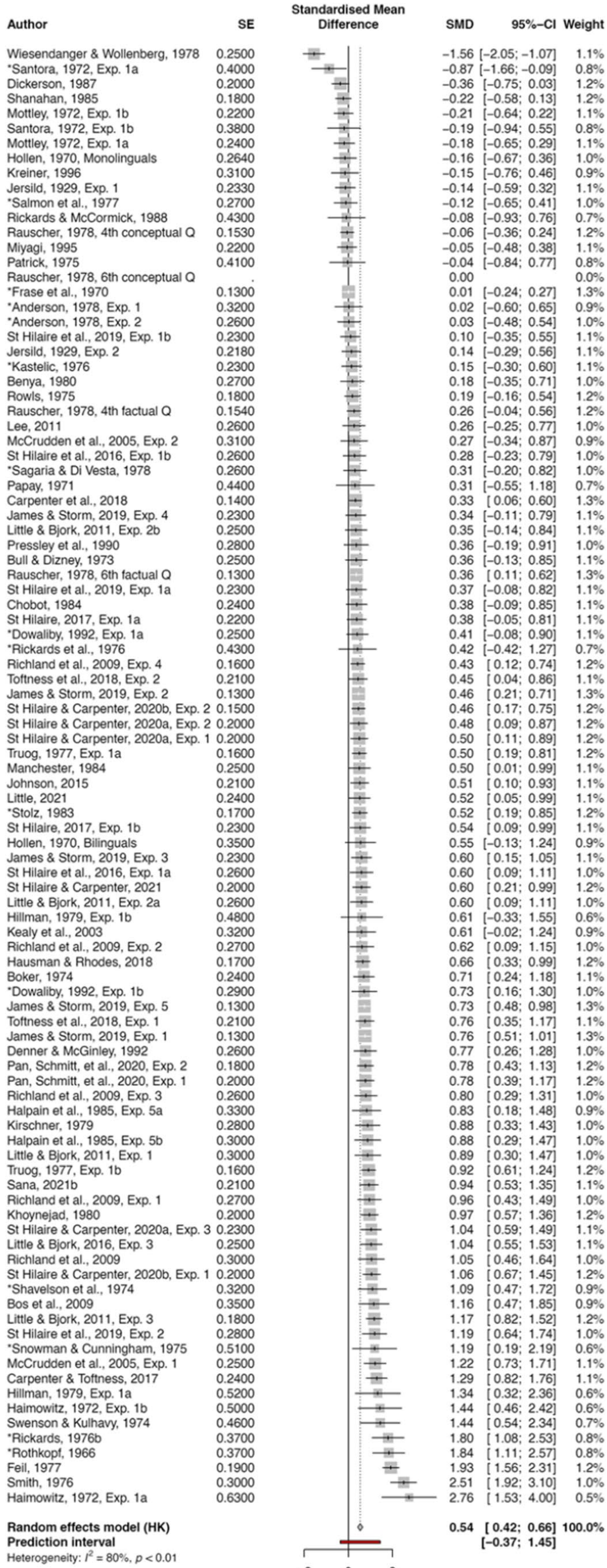
Different computations were used when calculating effect sizes for between-subjects versus within-subjects designs. Specifically, for within-subjects designs, one should use the correlation between the prequestion and control condition scores to adjust the magnitude of Cohen's *d*. However, none of the included studies provided correlation data regarding the relation between prequestion and control performance. As suggested by Cumming (2012), between-subjects effect size computations were used for within-subjects studies by assuming the correlation between the within-subjects conditions as .50.

A number of studies have compared multiple experimental prequestion conditions against a single control condition (e.g., factual prequestions versus conceptual prequestions versus one control condition). To ensure independence of effect sizes in the *overall* meta-analytic estimates, data from all experimental conditions were aggregated and compared against the control condition if there was only one control condition in a given study (i.e., only *one* effect size per study). However, in studies with multiple experimental conditions for which there were separate, *fully matched controls* (e.g., factual prequestions versus factual control, conceptual prequestions versus conceptual control), separate effect sizes were calculated for each independent comparison.

For *moderator* analyses, independence of effect sizes was not considered so long as the dependent effect size estimates contributed to *different* levels of a moderator variable. For example, in a study with two experimental conditions (e.g., factual prequestions versus conceptual prequestions) and a single control, separate effect sizes were computed for *both* levels of the moderator variable regardless of dependence, but only when that variable was being analyzed during moderator analysis (e.g., Prequestion Knowledge Type: factual prequestions vs. conceptual prequestions). All levels of dependent effect sizes were retained for moderator analyses because (1) moderator analyses pool data into *independent* effect size estimates, and (2) it is important to maximize the number of represented studies in all moderator analyses to increase statistical power (see Higgins et al., 2022).

To make this analytic decision concrete, consider this example from our meta-analysis. To preview our results, a total of 98 studies evaluated the specific effect (i.e., the number of records in the overall meta-analytic model). However, readers might notice that some moderator analyses, such as Prequestion Knowledge Type, have more effects than the overall model. We report a total of *111* effects in the Prequestion Knowledge Type moderator analysis (i.e., 13 *more* records than in the overall meta-analytic model), which might be confusing considering that not every study directly explored Prequestion Knowledge Type as part of their original experimental design (i.e., only 87 of the original 98 studies). The majority of studies that explored Prequestion Knowledge Type compared *either* verbatim *or* conceptual prequestions with a single control condition (*k* = 65). The remaining 22 records compared multiple Prequestion Knowledge Type conditions to a single control condition (e.g., verbatim prequestions vs. conceptual prequestions vs. a single control condition). To ensure that data from all of these experimental conditions were retained for the moderator analysis (i.e., the data from the verbatim prequestions *and* the data from the conceptual prequestions), separate effect sizes were computed for each level of the moderator. For 21 of these records, two effect sizes were computed for each study (i.e., a total of 42 effects). For one record, four effect sizes were computed (as researchers manipulated multiple levels of both conceptual and verbatim prequestions, see Halpain et al., 1985, Experiment 5). Because of this analytic approach, some moderator analyses had more effects (shown via degrees of freedom) than the overall model.

All analyses were conducted in R using the *meta* package (Balduzzi et al., 2019; Harrer et al., 2021). Random effects model was used for all analyses employing DerSimonian–Laird estimators (DerSimonian & Laird, 1986); Knapp-Hartung adjustments (Knapp & Hartung, 2003) were used to calculate confidence intervals around the overall effect size estimates.


Author SE Standardised Mean Difference SMD 95%-CI Weight
Wiesendanger & Wollenberg, 1978 0.2500 −1.56 [−2.05; −1.07] 1.1%
*Santora, 1972, Exp. 1a 0.4000 −0.87 [−1.66; −0.09] 0.8%
Dickerson, 1987 0.2000 −0.36 [−0.75; 0.03] 1.2%
Shanahan, 1985 0.1800 −0.22 [−0.58; 0.13] 1.2%
Mottley, 1972, Exp. 1b 0.2200 −0.21 [−0.64; 0.22] 1.1%
Santora, 1972, Exp. 1b 0.3800 −0.19 [−0.94; 0.55] 0.8%
Mottley, 1972, Exp. 1a 0.2400 −0.18 [−0.65; 0.29] 1.1%
Hollen, 1970, Monolinguals 0.2640 −0.16 [−0.67; 0.36] 1.0%
Kreiner, 1996 0.3100 −0.15 [−0.76; 0.46] 0.9%
Jersild, 1929, Exp. 1 0.2330 −0.14 [−0.59; 0.32] 1.1%
*Salmon et al., 1977 0.2700 −0.12 [−0.65; 0.41] 1.0%
Rickards & McCormick, 1988 0.4300 −0.08 [−0.93; 0.76] 0.7%
Rauscher, 1978, 4th conceptual Q 0.1530 −0.06 [−0.36; 0.24] 1.2%
Miyagi, 1995 0.2200 −0.05 [−0.48; 0.38] 1.1%
Patrick, 1975 0.4100 −0.04 [−0.84; 0.77] 0.8%
Rauscher, 1978, 6th conceptual Q . 0.00 0.0%
*Frase et al., 1970 0.1300 0.01 [−0.24; 0.27] 1.3%
*Anderson, 1978, Exp. 1 0.3200 0.02 [−0.60; 0.65] 0.9%
*Anderson, 1978, Exp. 2 0.2600 0.03 [−0.48; 0.54] 1.0%
St Hilaire et al., 2019, Exp. 1b 0.2300 0.10 [−0.35; 0.55] 1.1%
Jersild, 1929, Exp. 2 0.2180 0.14 [−0.29; 0.56] 1.1%
*Kastelic, 1976 0.2300 0.15 [−0.30; 0.60] 1.1%
Benya, 1980 0.2700 0.18 [−0.35; 0.71] 1.0%
Rowls, 1975 0.1800 0.19 [−0.16; 0.54] 1.2%
Rauscher, 1978, 4th factual Q 0.1540 0.26 [−0.04; 0.56] 1.2%
Lee, 2011 0.2600 0.26 [−0.25; 0.77] 1.0%
McCrudden et al., 2005, Exp. 2 0.3100 0.27 [−0.34; 0.87] 0.9%
St Hilaire et al., 2016, Exp. 1b 0.2600 0.28 [−0.23; 0.79] 1.0%
*Sagaria & Di Vesta, 1978 0.2600 0.31 [−0.20; 0.82] 1.0%
Papay, 1971 0.4400 0.31 [−0.55; 1.18] 0.7%
Carpenter et al., 2018 0.1400 0.33 [0.06; 0.60] 1.3%
James & Storm, 2019, Exp. 4 0.2300 0.34 [−0.11; 0.79] 1.1%
Little & Bjork, 2011, Exp. 2b 0.2500 0.35 [−0.14; 0.84] 1.1%
Pressley et al., 1990 0.2800 0.36 [−0.19; 0.91] 1.0%
Bull & Dizney, 1973 0.2500 0.36 [−0.13; 0.85] 1.1%
Rauscher, 1978, 6th factual Q 0.1300 0.36 [0.11; 0.62] 1.3%
St Hilaire et al., 2019, Exp. 1a 0.2300 0.37 [−0.08; 0.82] 1.1%
Chobot, 1984 0.2400 0.38 [−0.09; 0.85] 1.1%
St Hilaire, 2017, Exp. 1a 0.2200 0.38 [−0.05; 0.81] 1.1%
*Dowaliby, 1992, Exp. 1a 0.2500 0.41 [−0.08; 0.90] 1.1%
*Rickards et al., 1976 0.4300 0.42 [−0.42; 1.27] 0.7%
Richland et al., 2009, Exp. 4 0.1600 0.43 [0.12; 0.74] 1.2%
Toftness et al., 2018, Exp. 2 0.2100 0.45 [0.04; 0.86] 1.1%
James & Storm, 2019, Exp. 2 0.1300 0.46 [0.21; 0.71] 1.3%
St Hilaire & Carpenter, 2020b, Exp. 2 0.1500 0.46 [0.17; 0.75] 1.2%
St Hilaire & Carpenter, 2020a, Exp. 2 0.2000 0.48 [0.09; 0.87] 1.2%
St Hilaire & Carpenter, 2020a, Exp. 1 0.2000 0.50 [0.11; 0.89] 1.2%
Truog, 1977, Exp. 1a 0.1600 0.50 [0.19; 0.81] 1.2%
Manchester, 1984 0.2500 0.50 [0.01; 0.99] 1.1%
Johnson, 2015 0.2100 0.51 [0.10; 0.93] 1.1%
Little, 2021 0.2400 0.52 [0.05; 0.99] 1.1%
*Stolz, 1983 0.1700 0.52 [0.19; 0.85] 1.2%
St Hilaire, 2017, Exp. 1b 0.2300 0.54 [0.09; 0.99] 1.1%
Hollen, 1970, Bilinguals 0.3500 0.55 [−0.13; 1.24] 0.9%
James & Storm, 2019, Exp. 3 0.2300 0.60 [0.15; 1.05] 1.1%
St Hilaire et al., 2016, Exp. 1a 0.2600 0.60 [0.09; 1.11] 1.0%
St Hilaire & Carpenter, 2021 0.2000 0.60 [0.21; 0.99] 1.2%
Little & Bjork, 2011, Exp. 2a 0.2600 0.60 [0.09; 1.11] 1.0%
Hillman, 1979, Exp. 1b 0.4800 0.61 [−0.33; 1.55] 0.6%
Kealy et al., 2003 0.3200 0.61 [−0.02; 1.24] 0.9%
Richland et al., 2009, Exp. 2 0.2700 0.62 [0.09; 1.15] 1.0%
Hausman & Rhodes, 2018 0.1700 0.66 [0.33; 0.99] 1.2%
Boker, 1974 0.2400 0.71 [0.24; 1.18] 1.1%
*Dowaliby, 1992, Exp. 1b 0.2900 0.73 [0.16; 1.30] 1.0%
James & Storm, 2019, Exp. 5 0.1300 0.73 [0.48; 0.98] 1.3%
Toftness et al., 2018, Exp. 1 0.2100 0.76 [0.35; 1.17] 1.1%
James & Storm, 2019, Exp. 1 0.1300 0.76 [0.51; 1.01] 1.3%
Denner & McGinley, 1992 0.2600 0.77 [0.26; 1.28] 1.0%
Pan, Schmitt, et al., 2020, Exp. 2 0.1800 0.78 [0.43; 1.13] 1.2%
Pan, Schmitt, et al., 2020, Exp. 1 0.2000 0.78 [0.39; 1.17] 1.2%
Richland et al., 2009, Exp. 3 0.2600 0.80 [0.29; 1.31] 1.0%
Halpain et al., 1985, Exp. 5a 0.3300 0.83 [0.18; 1.48] 0.9%
Kirschner, 1979 0.2800 0.88 [0.33; 1.43] 1.0%
Halpain et al., 1985, Exp. 5b 0.3000 0.88 [0.29; 1.47] 1.0%
Little & Bjork, 2011, Exp. 1 0.3000 0.89 [0.30; 1.47] 1.0%
Truog, 1977, Exp. 1b 0.1600 0.92 [0.61; 1.24] 1.2%
Sana, 2021b 0.2100 0.94 [0.53; 1.35] 1.1%
Richland et al., 2009, Exp. 1 0.2700 0.96 [0.43; 1.49] 1.0%
Khoynejad, 1980 0.2000 0.97 [0.57; 1.36] 1.2%
St Hilaire & Carpenter, 2020a, Exp. 3 0.2300 1.04 [0.59; 1.49] 1.1%
Little & Bjork, 2016, Exp. 3 0.2500 1.04 [0.55; 1.53] 1.1%
Richland et al., 2009 0.3000 1.05 [0.46; 1.64] 1.0%
St Hilaire & Carpenter, 2020b, Exp. 1 0.2000 1.06 [0.67; 1.45] 1.2%
*Shavelson et al., 1974 0.3200 1.09 [0.47; 1.72] 0.9%
Bos et al., 2009 0.3500 1.16 [0.47; 1.85] 0.9%
Little & Bjork, 2011, Exp. 3 0.1800 1.17 [0.82; 1.52] 1.2%
St Hilaire et al., 2019, Exp. 2 0.2800 1.19 [0.64; 1.74] 1.0%
*Snowman & Cunningham, 1975 0.5100 1.19 [0.19; 2.19] 0.6%
McCrudden et al., 2005, Exp. 1 0.2500 1.22 [0.73; 1.71] 1.1%
Carpenter & Toftness, 2017 0.2400 1.29 [0.82; 1.76] 1.1%
Hillman, 1979, Exp. 1a 0.5200 1.34 [0.32; 2.36] 0.6%
Haimowitz, 1972, Exp. 1b 0.5000 1.44 [0.46; 2.42] 0.6%
Swenson & Kulhavy, 1974 0.4600 1.44 [0.54; 2.34] 0.7%
*Rickards, 1976b 0.3700 1.80 [1.08; 2.53] 0.8%
*Rothkopf, 1966 0.3700 1.84 [1.11; 2.57] 0.8%
Feil, 1977 0.1900 1.93 [1.56; 2.31] 1.2%
Smith, 1976 0.3000 2.51 [1.92; 3.10] 1.0%
Haimowitz, 1972, Exp. 1a 0.6300 2.76 [1.53; 4.00] 0.5%
Random effects model (HK) 0.54 [0.42; 0.66] 100.0%
Prediction interval [−0.37; 1.45]
Heterogeneity: I² = 80%, p < 0.01
−2 0 2

◂**Fig. 4** Forest plot for the specific effect. *Note*. Studies were ordered by their effect sizes. Studies marked with an asterisk had variability estimates imputed

Results

We present the results in three sections. First, the overall meta-analytic effect sizes for the specific and general PEs are provided with forest plots. All effect size estimates are accompanied by .95 confidence intervals in brackets. Second, publication bias is explored using multiple methods, although we caution that no publication bias analyses are foolproof. Third, we report results of the theoretically relevant and empirically relevant moderator analyses.

Overall effect size

The pooled effect size for the specific effect was $g = 0.54$ [0.42, 0.66], providing solid evidence that prequestions enhance learning of the tested material ($k = 97$), $p < .001$. A forest plot for the specific effect shows that most effect sizes were positive (85%), with only 15 studies demonstrating negative effects (see Fig. 4). There was significant heterogeneity in the data, $\tau^2 = .21$, $I^2 = 80\%$, $Q = 489.98$, $p < .001$, indicating that the size of the effect is likely influenced by moderators. In stark contrast to the specific effect, there was virtually *no* evidence for the general effect, $g = 0.04$ [−0.04, 0.11], as 54% of studies showed positive effects and 46% of studies showed either no effect or a negative effect ($k = 91$; see Fig. 5), $p = .349$. There was also substantial heterogeneity in the data, $\tau^2 = .07$, $I^2 = 60\%$, $Q = 225.92$, $p < .001$, indicating that some moderator variables might be associated with positive or negative general effects.

Publication bias

We used multiple methods to assess publication bias. First, all studies were categorized based on their publication status to determine whether the effects were larger for *published* (i.e., peer-reviewed journal articles) than *unpublished* studies.

For the specific effect, 53 studies were published in peer-reviewed journal articles and 45 studies were from unpublished sources (30 theses/dissertations, 12 conference presentations, and three unpublished datasets). For the general effect, 51 studies were published and 40 were unpublished (25 theses/dissertations, 11 conference presentations, and four unpublished datasets). For both the specific (published: $g = 0.59$; unpublished: $g = 0.49$) and general effects (published: $g = 0.06$; unpublished: $g = 0.003$), publication status did not significantly influence the magnitude of the PE, *p*s > .437. This provides preliminary evidence that publication bias, should it be present, is likely not a major source of influence in the data.

Funnel plots were inspected for publication bias. Because it is easier to publish significant effects than non-significant ones, asymmetrical funnel plots will typically have missing underpowered studies with smaller (or negative) effects. The funnel plots in Fig. 6 do not show asymmetry, as evidenced by results of the Egger's test (Egger et al., 1997) for both the specific and general effects, $t_{specific}(95) = 1.38$, $p = .172$, $t_{general}(89) = -0.05$, $p = .964$.

Lastly, the PET-PEESE technique was used to further assess for publication bias and, if necessary, provide adjusted effect size estimates (Stanley & Doucouliagos, 2014). For both PET and PEESE, the beta weights denote the *adjusted* effect size estimates. For the specific effect, the intercept of the PET test was significantly greater than 0 ($\beta = 0.37$, $p = .041$), so PEESE was used for the bias-adjusted effect ($\beta = 0.45$, $p < .001$). For the general effect, the intercept of the PET test was not significant ($\beta = 0.04$, $p = .714$), so PET was retained (for completeness, the results of PEESE for the general effect was: $\beta = 0.05$, $p = .402$). For both PEs, the adjusted effect sizes (i.e., $g = 0.45$ for the specific effect; $g = 0.04$ for the general effect) are only slightly deflated compared to the original effect size estimates, which, when interpreted alongside Egger's test, do not affect the interpretations of the overall meta-analytic effects.

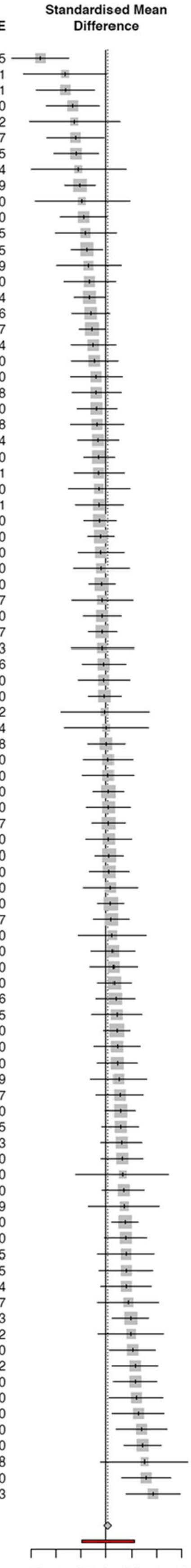
In sum, all tests of publication bias showed little risk of selective reporting of significant results for both the specific and the general effects.

Moderator analyses

Results for all categorical moderator analyses are presented in Tables 2 and 3. Meta-regression analyses for continuous moderators are presented in Tables 4 and 5 For all moderator analyses, the number of effects (k) for each level of the moderator is provided alongside the Q_B statistic.

Theoretically relevant moderator analyses

For the *specific prequestion effect*, prequestion guessing produced a significant effect, whereas prequestion knowledge type and prequestion performance produced marginal effects (Tables 2 and 4). All other moderators were not significant. Both the memory and curiosity precursor accounts correctly predicted a greater specific effect when participants were asked to guess the answer to the prequestions during the pre-study phase ($g = 0.65$) than when they were not ($g = 0.22$), $p < .001$. Indeed, amongst all of the moderator variables, prequestion guessing had the strongest influence on the PE. In contrast, prequestion knowledge type moderator produced a pattern that was contrary to that predicted by the



Author	SE	SMD	95%-CI	Weight
*Sagaria & Di Vesta, 1978	0.2895	−1.31	[−1.88; −0.75]	0.9%
Patrick, 1975	0.4251	−0.82	[−1.65; 0.01]	0.5%
Peeck, 1970	0.2991	−0.81	[−1.40; −0.23]	0.9%
Hollen, 1970, Monolinguals	0.2710	−0.66	[−1.19; −0.13]	1.0%
Haimowitz, 1972, Exp. 1b	0.4582	−0.63	[−1.53; 0.27]	0.5%
Strollo, 1972	0.2937	−0.61	[−1.18; −0.03]	0.9%
Smith, 1976	0.2285	−0.60	[−1.04; −0.15]	1.1%
Snowman & Cunningham, 1975	0.4804	−0.55	[−1.50; 0.39]	0.5%
Truog, 1977, Exp. 1a	0.1559	−0.52	[−0.82; −0.21]	1.5%
Hillman, 1979, Exp. 1a	0.4800	−0.48	[−1.42; 0.46]	0.5%
Boker, 1974	0.2400	−0.44	[−0.92; 0.03]	1.1%
*Shavelson et al., 1974	0.3075	−0.41	[−1.01; 0.19]	0.8%
Feil, 1977	0.1595	−0.38	[−0.69; −0.07]	1.5%
Karjala, 1984, Exp. 1b	0.3269	−0.35	[−0.99; 0.29]	0.8%
Richland et al., 2009, Exp. 1	0.2600	−0.33	[−0.84; 0.18]	1.0%
Truog, 1977, Exp. 1b	0.1544	−0.33	[−0.63; −0.03]	1.5%
Memory, 1981	0.1946	−0.30	[−0.69; 0.08]	1.3%
*Frase et al., 1970	0.1297	−0.28	[−0.53; −0.02]	1.6%
Miyagi, 1995	0.2244	−0.26	[−0.70; 0.18]	1.1%
McCrudden et al., 2005, Exp. 1	0.2350	−0.23	[−0.69; 0.23]	1.1%
St Hilaire et al., 2016, Exp. 1b	0.2600	−0.20	[−0.71; 0.31]	1.0%
*Dowaliby, 1992, Exp. 1a	0.2468	−0.20	[−0.68; 0.29]	1.1%
St Hilaire & Carpenter, 2020a, Exp. 1	0.2000	−0.19	[−0.58; 0.21]	1.3%
Smith, 1996	0.2678	−0.18	[−0.71; 0.34]	1.0%
Johnson, 2015	0.2054	−0.16	[−0.56; 0.24]	1.2%
Rauscher, 1978, 4th conceptual Q	0.1540	−0.14	[−0.45; 0.16]	1.5%
Bull & Dizney, 1973	0.2521	−0.14	[−0.64; 0.35]	1.0%
Kreiner, 1996	0.3100	−0.14	[−0.75; 0.47]	0.8%
Chobot, 1984	0.2411	−0.14	[−0.61; 0.33]	1.1%
Richland et al., 2009, Exp. 4	0.1600	−0.13	[−0.44; 0.19]	1.5%
James & Storm, 2019, Exp. 2	0.1300	−0.10	[−0.36; 0.15]	1.6%
St Hilaire et al., 2019, Exp. 1b	0.2300	−0.10	[−0.55; 0.35]	1.1%
Richland et al., 2009	0.2800	−0.10	[−0.65; 0.45]	0.9%
James & Storm, 2019, Exp. 1	0.1300	−0.09	[−0.34; 0.17]	1.6%
*Rothkopf, 1966	0.3087	−0.08	[−0.68; 0.53]	0.8%
St Hilaire & Carpenter, 2021	0.1900	−0.08	[−0.45; 0.30]	1.3%
Carpenter et al., 2018	0.1427	−0.07	[−0.35; 0.21]	1.5%
*Rickards, 1976b	0.3163	−0.07	[−0.69; 0.55]	0.8%
St Hilaire, 2017, Exp. 1a	0.2196	−0.04	[−0.47; 0.39]	1.2%
Richland et al., 2009, Exp. 2	0.2600	−0.04	[−0.55; 0.47]	1.0%
*Stolz, 1983	0.1640	−0.03	[−0.35; 0.29]	1.4%
Haimowitz, 1972, Exp. 1a	0.4472	−0.02	[−0.90; 0.85]	0.5%
Rickards & McCormick, 1988	0.4264	0.00	[−0.83; 0.84]	0.5%
Khoynejad, 1980	0.1858	0.00	[−0.36; 0.37]	1.3%
Little & Bjork, 2011, Exp. 2b	0.2500	0.04	[−0.45; 0.53]	1.0%
St Hilaire et al., 2016, Exp. 1a	0.2600	0.04	[−0.47; 0.55]	1.0%
Rauscher, 1978, 4th factual Q	0.1530	0.04	[−0.26; 0.34]	1.5%
James & Storm, 2019, Exp. 4	0.2200	0.04	[−0.39; 0.48]	1.2%
Little & Bjork, 2011, Exp. 3	0.1667	0.05	[−0.28; 0.37]	1.4%
St Hilaire et al., 2019, Exp. 1a	0.2300	0.05	[−0.40; 0.50]	1.1%
Janelli & Lipnevich, 2021	0.1400	0.06	[−0.22; 0.33]	1.5%
St Hilaire & Carpenter, 2020a, Exp. 2	0.2000	0.06	[−0.33; 0.45]	1.3%
Pressley et al., 1990	0.2740	0.08	[−0.45; 0.62]	0.9%
James & Storm, 2019, Exp. 5	0.1300	0.09	[−0.17; 0.34]	1.6%
Shanahan, 1985	0.1787	0.10	[−0.25; 0.45]	1.4%
Hollen, 1970, Bilinguals	0.3440	0.12	[−0.56; 0.79]	0.7%
James & Storm, 2019, Exp. 3	0.2200	0.13	[−0.30; 0.56]	1.2%
Little, 2021	0.2400	0.15	[−0.32; 0.62]	1.1%
Hausman & Rhodes, 2018	0.1700	0.16	[−0.17; 0.49]	1.4%
Pan, Schmitt, et al., 2020, Exp. 1	0.1956	0.19	[−0.20; 0.57]	1.3%
Yang et al., 2021	0.2535	0.21	[−0.29; 0.71]	1.0%
Rauscher, 1978, 6th conceptual Q	0.1290	0.21	[−0.04; 0.46]	1.6%
Felker & Dapra, 1975	0.2300	0.21	[−0.24; 0.67]	1.1%
Toftness et al., 2018, Exp. 1	0.2000	0.22	[−0.17; 0.61]	1.3%
Little & Bjork, 2011, Exp. 1	0.2839	0.25	[−0.31; 0.80]	0.9%
Little & Bjork, 2016, Exp. 3	0.2367	0.27	[−0.20; 0.73]	1.1%
St Hilaire & Carpenter, 2020b, Exp. 2	0.1500	0.28	[−0.02; 0.57]	1.5%
Rowls, 1975	0.1835	0.28	[−0.08; 0.64]	1.3%
Memory, 1983	0.2053	0.30	[−0.10; 0.70]	1.2%
St Hilaire & Carpenter, 2020a, Exp. 3	0.2100	0.31	[−0.10; 0.72]	1.2%
Hillman, 1979, Exp. 1b	0.4700	0.32	[−0.60; 1.24]	0.5%
Sana, 2021b	0.2100	0.34	[−0.07; 0.75]	1.2%
Wilkerson, 1982	0.3569	0.35	[−0.35; 1.05]	0.7%
Rauscher, 1978, 6th factual Q	0.1300	0.37	[0.11; 0.62]	1.6%
Toftness et al., 2018, Exp. 2	0.2100	0.38	[−0.03; 0.80]	1.2%
*Dowaliby, 1992, Exp. 1b	0.2855	0.39	[−0.17; 0.95]	0.9%
Kirschner, 1979	0.2705	0.39	[−0.14; 0.92]	1.0%
Little & Bjork, 2011, Exp. 2a	0.2524	0.39	[−0.10; 0.89]	1.0%
McCrudden et al., 2005, Exp. 2	0.3087	0.43	[−0.17; 1.04]	0.8%
Pan, Schmitt, et al., 2020, Exp. 2	0.1803	0.48	[0.13; 0.83]	1.4%
Karjala, 1984, Exp. 1a	0.3292	0.49	[−0.16; 1.13]	0.8%
Sana, 2021a	0.2300	0.52	[0.07; 0.97]	1.1%
St Hilaire, 2017, Exp. 1b	0.2282	0.57	[0.13; 1.02]	1.1%
Carpenter & Toftness, 2017	0.2200	0.57	[0.14; 1.01]	1.2%
St Hilaire et al., 2019, Exp. 2	0.2700	0.60	[0.07; 1.12]	1.0%
Richland et al., 2009, Exp. 3	0.2600	0.64	[0.13; 1.15]	1.0%
Denner & McGinley, 1992	0.2600	0.70	[0.19; 1.21]	1.0%
St Hilaire & Carpenter, 2020b, Exp. 1	0.1900	0.72	[0.35; 1.09]	1.3%
Swenson & Kulhavy, 1974	0.4398	0.76	[−0.10; 1.62]	0.5%
Osman & Hannafin, 1994	0.2500	0.79	[0.30; 1.28]	1.0%
Lee, 2011	0.2763	0.93	[0.38; 1.47]	0.9%
Random effects model (HK)		**0.04**	**[−0.04; 0.11]**	**100.0%**
Prediction interval			**[−0.49; 0.56]**	

Heterogeneity: $I^2 = 60\%$, $p < 0.01$

◀**Fig. 5** Forest plot for the general effect. *Note*. Studies were ordered by their effect sizes. Studies marked with an asterisk had variability estimates imputed

curiosity precursor account, such that a marginally larger specific effect was observed for *factual prequestions* (g = 0.58) than for *conceptual prequestions* (g = 0.28), p = .073. Lastly, the curiosity precursor account correctly predicted a larger specific PE when participants have lower prior knowledge, as demonstrated by a negative association between the PE and prequestion test performance (see Fig. 7a), β = −0.49, p = .094. Of note is that there were some outliers in the data (~80% prequestion performance). In both cases, participants were adult experts in their fields studying new information in their area of expertise (Anderson, 1978). To avoid range restriction in the data, outliers were retained in all analyses. This was an especially important consideration because fewer than 30% of the studies have reported prequestion performance (primarily those published within the past 15 years). Due to the limited reporting of and the presence of outliers in this moderator variable, these data should be interpreted with caution.

For the *general prequestion effect*, interleaving, relation between tested and nontested materials, type of learning materials, and length of time per prequestion were significant moderators. None of the other moderators were significant (Tables 3 and 5). Notably, three of these moderators produced results that were consistent with the predictions emerging from the memory precursor account. Specifically, prequestions produced a larger general benefit for massed study (g = 0.11) than for interleaved study (g = −0.07), p = .036, for nontested retention questions that were related to the prequestions (g = 0.17) than those unrelated to the prequestions (g = 0.01), p = .029, and for audiovisual (g = 0.17) than text *materials* (g = −0.005), p = .017.

Empirically relevant moderator analyses

Amongst the empirically relevant moderators, five produced significant results, including publication year, participant age group, study phase timing, reading speed, and retention test format. Publication year had a marginally significant influence on the size of the specific effect (β = 0.01; see Fig. 8a), p = .058. However, somewhat expectedly, the general effect has *increased* significantly over time (β = 0.01; see Fig. 8b), p < .001. Speculation regarding why the size of the general effect has increased (whereas the specific effect has not) will be offered in the Discussion. The age group moderator showed that adult participants (g = 0.62) tended to produce a large specific effect than grade school children (g = 0.22), p = .020. However, no differences in participant age were observed for the general effect (*adults:* g = 0.02; *grade school children*: g = 0.11), p = .403. We caution against drawing strong conclusions from these results because of the borderline p value and the limited number of effect sizes (k = 20 for the specific effect and k = 18 for the general effect) in the grade school category.

Interestingly, the magnitude of both the specific and general effects was larger when participants received an *experimenter-paced study phase* as opposed to a *self-paced study phase* (ps < .05). For experimenter-paced study phase, reading speed (words per minute) was then regressed onto the overall effect size estimates. For the specific effect, there was no relation between *reading speed* and the magnitude of the PE (β = 0.002; see Fig. 9a), p = .282.

However, for the general effect, there was a significant negative relation between *reading speed* and the general PE (β = −0.004; see Fig. 9b), p = .005. In other words, a greater general effect of prequestions was observed when participants had more time to study the material. The size of the PE was not hypothesized to differ depending on study phase length, so it is important not to overinterpret these data; however, we consider why study phase timing and study rate might influence the magnitude of the PE in the Discussion.

Discussion

Research has consistently shown that prequestions offer *specific* benefits for the initially tested material; the results of this meta-analysis provide compelling evidence supporting this conclusion (g = 0.54). However, evidence for whether prequestions additionally offer *general* benefits to the nontested material has been more equivocal (see James & Storm, 2019). Regarding the studies that were included in the current analysis, nearly half of them demonstrated a negative general effect and the remaining demonstrated a positive general effect. It is not surprising, then, that the magnitude of the general effect was virtually nonexistent (g = 0.04). In what follows, we discuss results of the moderator analyses in the context of the memory and curiosity precursor accounts. To facilitate discussions, we summarized the impressions from the moderator analyses in Table 6.

Theoretical evaluations

There is considerable consensus that prequestions enhance learning because they selectively orient learners' attention during the study phase to the tested material (Frase, 1967; Hannafin & Hughes, 1986; Hartley & Davies, 1976). However, what has remained unclear are the mechanisms underscoring this attentional shift. Here, we proposed two noncompetitive precursor explanations: a memory precursor account and a curiosity precursor account. It is important to reiterate that these accounts are unlikely to be exclusive or exhaustive. Rather, we believe that they are complementary,

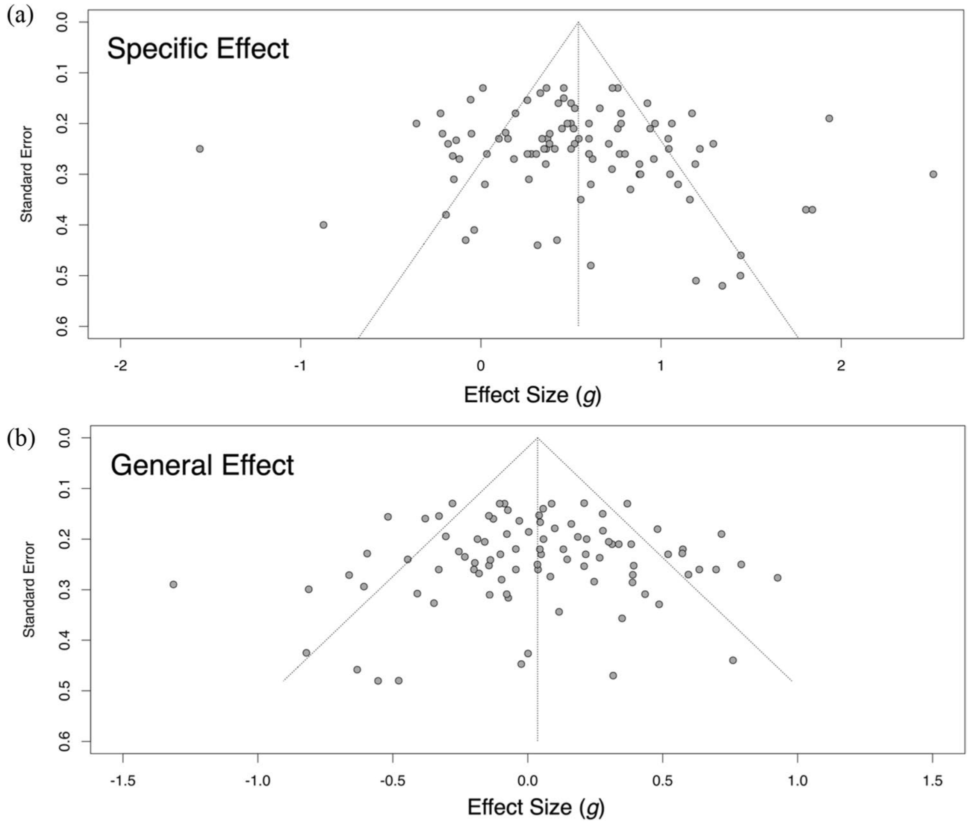


Fig. 6 Funnel plots for the specific effect (**a**) and the general effect (**b**)

and we hope that by specifying how memory and curiosity can enable the attention-shifting influences of prequestions, and by outlining specific predictions of these precursor accounts, the present review will spur further theoretically oriented studies.

Memory precursor account

The memory precursor account specifies that prequestions direct attention to the tested material because learners remember the prequestions (see St. Hilaire & Carpenter, 2020a). When learners encounter information during the study phase that reminds them of these prequestions, they can more readily integrate the answer with the prequestions, thereby enhancing learning. The logic parallels other elaborative encoding accounts, such as those in levels of processing (Craik & Tulving, 1975). Remembering the prequestions can also enhance new learning by providing learners with an existing framework through which to comprehend the study material (Kaplan & Rothkopf, 1974), which might also facilitate learning.

Overall, the memory precursor account received very little support regarding the *specific effect*, as only one of the six moderators produced results consistent with its predictions (i.e., guess prequestions led to a larger specific PE than read-only prequestions; see Table 6). Data from four moderators did not provide support for the memory precursor account—with all showing a null effect (although none showed a significant effect in the opposite direction). In hindsight, these null results might not be entirely surprising because manipulations of the number of prequestions (James & Storm, 2019), interleaving (Pan et al., 2020), and prequestion test format (Little & Bjork, 2016) have shown no significant effects on the magnitude of the specific effect in individual investigations.

Contrary to the specific effect, the memory precursor account received moderate support for its predictions regarding *the general effect* on three of the four tested predictions: the general effect was larger for massed study than interleaved study, for audiovisual than text materials, and for related than unrelated retention questions. Although these data are favorable to the memory precursor account, they must also be considered in the context of rather small general effects. For all three of these moderators, the magnitude of the general effect in the condition that was predicted to promote the PE did not exceed $g = 0.20$. Furthermore, there was little evidence that the size of the general effect was related to the number of prequestions asked, and not

Table 2 Results of the categorical moderator analyses for the specific effect

Categorical Moderators	Level	*k*	*g*	*CI*	Q_B	*p*
Overall Model		97	0.54	[0.42, 0.66]	489.98	< .001
[MC]Prequestion Guess					11.05	< .001
	Guess	70	0.65	[0.53, 0.77]		
	Read-Only	29	0.22	[-0.02, 0.45]		
[M]Interleaved Study					0.44	.509
	Massed	49	0.50	[0.28, 0.71]		
	Interleaved	49	0.57	[0.47, 0.68]		
[MC]Prequestion Test Format					0.31	.578
	Multiple Choice	42	0.51	[0.36, 0.67]		
	Short Answer	57	0.58	[0.41, 0.75]		
[C]Prequestion Knowledge Type					3.22	.073
	Factual	88	0.58	[0.45, 0.72]		
	Conceptual	22	0.28	[-0.03, 0.60]		
Type of Learning Materials					1.94	.163
	Text	77	0.51	[0.37, 0.65]		
	Audiovisual	21	0.66	[0.49, 0.83]		
Relation between Tested and Nontested Material					0.61	.435
	Related	12	0.63	[0.40, 0.86]		
	Unrelated	77	0.54	[0.40, 0.67]		
Research Design					0.08	.774
	Between-Subjects	85	0.54	[0.40, 0.67]		
	Within-Subjects	12	0.57	[0.34, 0.80]		
Participant Age					5.45	.020
	Adults	77	0.62	[0.50, 0.74]		
	Grade School Children	20	0.22	[-0.11, 0.56]		
Control Condition Task					1.95	.378
	Mock Pretest	5	0.66	[-0.19, 1.52]		
	Extended Study Time	17	0.66	[0.48, 0.83]		
	Study-Only	75	0.51	[0.36, 0.65]		
Pre-Study Phase Timing					1.71	.190
	Experimenter-Paced	36	0.63	[0.48, 0.77]		
	Self-Paced	60	0.48	[0.31, 0.65]		
Study Phase Timing					3.85	.049
	Experimenter-Paced	50	0.65	[0.54, 0.76]		
	Self-Paced	45	0.42	[0.20, 0.63]		
Retention Test Format					4.65	.031
	Multiple Choice	47	0.42	[0.27, 0.57]		
	Short Answer	50	0.67	[0.49, 0.85]		
Match between Prequestions and Retention Questions					0.00	.950
	Match	74	0.54	[0.40, 0.68]		
	Mismatch	23	0.55	[0.32, 0.77]		
Exclusion of Accurate Prequestion Guesses					4.58	.101
	Include	31	0.60	[0.50, 0.71]		
	Exclude	6	0.83	[0.49, 1.17]		
	Undefined	60	0.48	[0.30, 0.67]		

Superscripts in front of the moderator name indicates a theoretically relevant variable. M denotes that the moderator is relevant for the memory precursor account, and C denotes that the moderator is relevant for the curiosity precursor account

Table 3 Results of the categorical moderator analyses for the general effect

Categorical Moderators	Level	*k*	*g*	*CI*	Q_B	*p*
Overall Model		91	0.04	[-0.04, 0.11]	225.92	.349
Prequestion Guess					1.76	.185
	Guess	71	0.06	[-0.02, 0.14]		
	Read-Only	23	-0.07	[-0.28, 0.13]		
[M]Interleaved Study					4.41	.036
	Massed	48	0.11	[0.02, 0.20]		
	Interleaved	44	-0.06	[-0.18, 0.07]		
[C]Prequestion Test Format					0.01	.913
	Multiple Choice	40	0.04	[-0.07, 0.14]		
	Short Answer	53	0.04	[-0.06, 0.15]		
[C]Prequestion Knowledge Type					0.28	.600
	Factual	77	0.03	[-0.06, 0.11]		
	Conceptual	22	0.07	[-0.10, 0.25]		
[M]Type of Learning Materials					5.69	.017
	Text	70	-0.005	[-0.10, 0.09]		
	Audiovisual	22	0.17	[0.05, 0.29]		
[M]Relation Between Tested and Nontested Material					4.78	.029
	Related	15	0.17	[0.05, 0.30]		
	Unrelated	72	0.01	[-0.08, 0.10]		
Research Design					0.99	.319
	Between-Subjects	78	0.02	[-0.06, 0.11]		
	Within-Subjects	13	0.09	[-0.02, 0.20]		
Participant Age					0.70	.403
	Adults	73	0.02	[-0.06, 0.10]		
	Grade School Children	18	0.11	[-0.11, 0.33]		
Control Condition Task					2.28	.320
	Mock Pretest	5	0.20	[-0.22, 0.61]		
	Extended Study Time	17	0.10	[-0.04, 0.25]		
	Study-Only	69	0.01	[-0.08, 0.10]		
Pre-Study Phase Timing					0.00	.944
	Experimenter-Paced	34	0.04	[-0.08, 0.16]		
	Self-Paced	56	0.03	[-0.07, 0.14]		
Study Phase Timing					4.12	.042
	Experimenter-Paced	50	0.11	[0.01, 0.21]		
	Self-Paced	39	-0.05	[-0.17, 0.08]		
Retention Test Format					0.07	.791
	Multiple Choice	45	0.03	[-0.09, 0.14]		
	Short Answer	46	0.05	[-0.06, 0.16]		
Match between Prequestions and Retention Questions					2.69	.101
	Match	66	0.003	[-0.09, 0.09]		
	Mismatch	25	0.14	[-0.003, 0.28]		
Exclusion of Accurate Prequestion Guesses					5.21	.074
	Include	31	0.13	[0.04, 0.21]		
	Exclude	6	0.10	[-0.32, 0.52]		
	Undefined	54	-0.04	[-0.16, 0.08]		

Table 4 Results of the continuous moderator analyses for the specific effect

Continuous Moderators	*k*	β	*CI*	Q_B	*p*
[M]Length of Time Per Prequestion	34	-0.003	[-0.01, 0.01]	72.06	.481
[M]Number of Prequestions	97	-0.004	[-0.02, 0.01]	482.43	.672
[C]Prequestion Performance	40	-0.49	[-1.06, 0.09]	79.17	.095
Publication Year	97	0.01	[-0.0002, 0.01]	452.58	.060
Reading Speed	25	0.002	[-0.001, 0.005]	53.41	.282
Retention Interval Between Study Phase and Retrieval Phase	96	-0.04	[-0.14, 0.06]	483.71	.412

Table 5 Results of the continuous moderator analyses for the general effect

Continuous Moderators	*k*	β	*CI*	Q_B	*p*
Length of Time Per Prequestion	34	0.01	[0.001, 0.01]	53.56	.020
[M]Number of Prequestions	93	-0.002	[-0.02, 0.01]	227.46	.724
Prequestion Performance	37	0.22	[-0.35, 0.79]	56.87	.441
Publication Year	91	0.008	[0.004, 0.01]	195.99	< .001
Reading Speed	24	-0.004	[-0.006, -0.001]	36.18	.005
Retention Interval Between Study Phase and Retrieval Phase	84	-0.04	[-0.11, 0.03]	217.32	.263

enough studies have manipulated the length of the pretest-study delay to explore this moderator ($k = 1$).

Curiosity precursor account

According to the *curiosity precursor account*, prequestions make participants more curious to learn the answers to the prequestions, and this enhanced curiosity motivates learners to seek out the answers to the prequestions, encode the material with greater interest, and thus promote attentive encoding during the study phase (see Berlyne, 1954a, b). The curiosity precursor account makes slightly different predictions for the specific and general effects.

Regarding the *specific effect*, the present data are weakly in favor of the curiosity precursor account, as only one of the original predictions was supported (i.e., guess prequestions corresponded to a larger PE than read-only prequestions, see Table 6). However, it is worth noting that there was also a marginal, predicted relation between prequestion performance and the specific PE ($p = .094$), such that studies with lower prequestion performance were associated with a larger specific effect (see Fig. 7a). Because relatively few studies have reported prequestion performance data ($k = 40$), we believe that this variable deserves further empirical scrutiny, and researchers are advised to always report prequestion

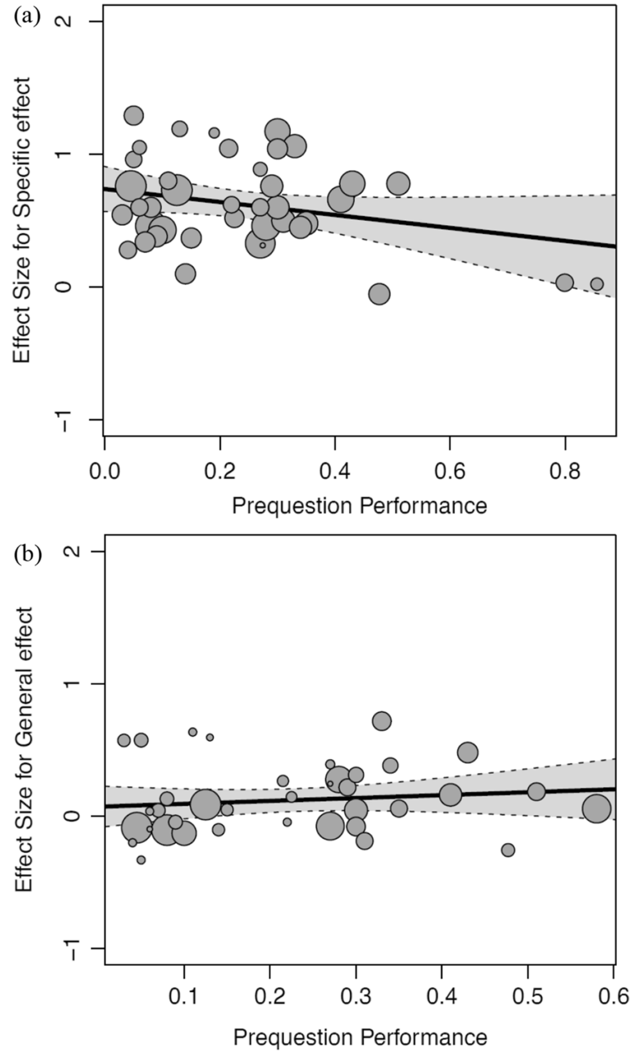


Fig. 7. Bubble plots showing the relation between prequestion performance and the specific effect (**a**) and the general effect (**b**)

performance. Prequestion knowledge type also produced a marginally significant effect ($p = .073$), but the direction of this effect was opposite to the prediction emerging from the curiosity precursor account: the magnitude of the specific PE was *smaller* for conceptual ($g = 0.28$) than for factual prequestions ($g = 0.58$). Because relatively few studies have explored the PE with conceptual prequestions ($k = 22$), more research is needed to establish the reliability of this finding. Finally, there was no evidence to suggest that the PE was larger for short answer than for multiple choice prequestions.

To our knowledge, only Berlyne (1954b) has asked participants to provide curiosity ratings during the pre-study phase. Unfortunately, Berlyne's data were not included in the meta-analysis due to a missing control condition. Because the only moderator that supported the curiosity precursor account also supported the memory precursor account (i.e., the PE was larger for guessing than read-only prequestions

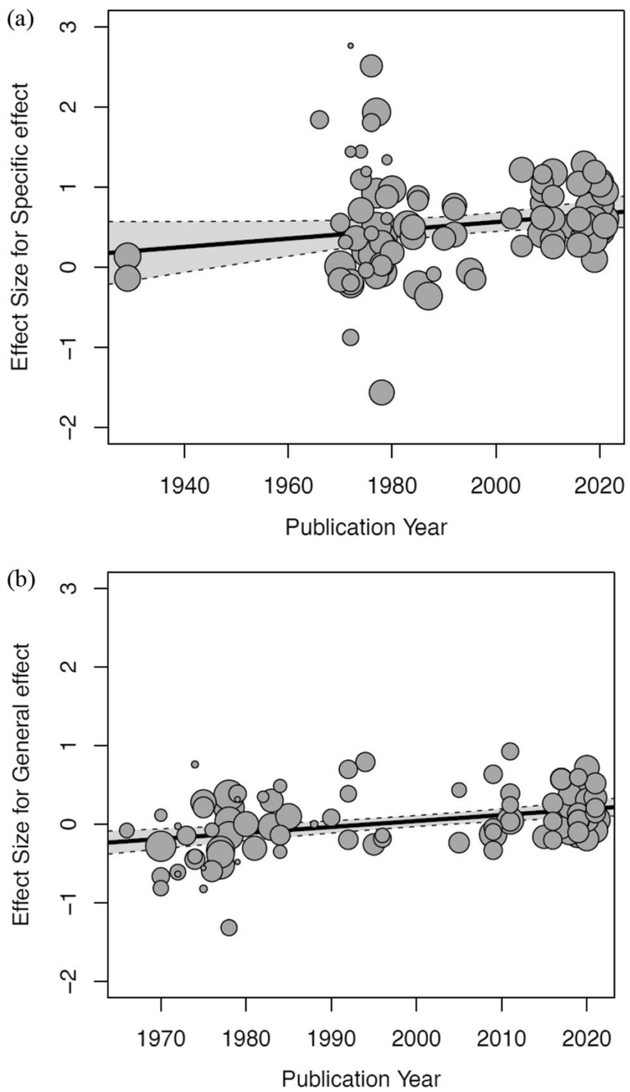


Fig. 8. Bubble plots showing the relation between publication year and the specific effect (**a**) and the general effect (**b**)

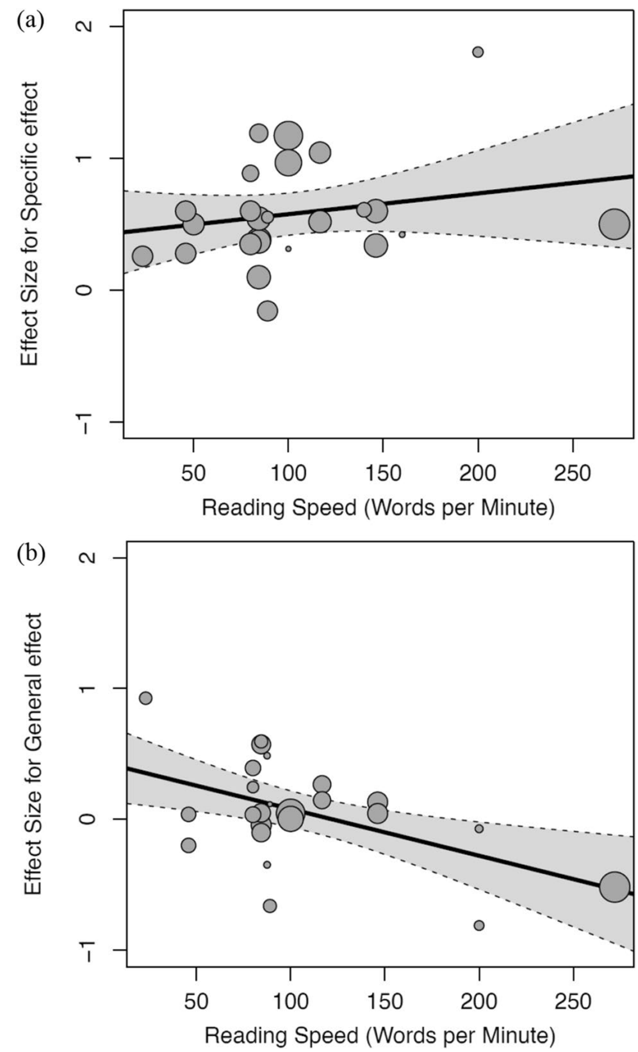


Fig. 9 Bubble plots showing the relation between reading speed and the specific effect (**a**) and the general effect (**b**)

for both accounts), it is hard to know if guessing the answers to prequestions directly influences curiosity or if guessing benefits learning for other reasons (i.e., guessing activates participants' semantic network; see Craik & Tulving, 1975). This is especially true considering the difficulty in assessing curiosity without asking prequestions (e.g., in a control condition). Crucially, as we have alluded to in the Introduction, the curiosity precursor account has received scant attention, and it requires further specification.

Contrary to our predictions, neither prequestion test format nor prequestion knowledge type affected the magnitude of the *general effect*. In fact, there was no general PE regardless of the type of questions or test format used. As outlined above for the specific effect, the curiosity precursor account still lacks specification, so few predictions could be made for this account. Additionally, few studies have directly manipulated curiosity or collected curiosity ratings during the pre-study phase, so further research is necessary to evaluate whether there is a meaningful relation between curiosity and the magnitude of the general effect.

Summary evaluations for the attention-based account of the PE

Overall, our moderator analyses provided considerable support for the memory precursor account, at least as it pertains to the general effect. For the specific effect, however, the only moderator analysis that supported the memory precursor account (i.e., guess prequestions versus read-only prequestions) supported the curiosity precursor account. Based on the present results, we tentatively conclude that the mechanistic precursors underlying the PE might be *different* for the specific and general effects—it is possible that

Table 6 Support for each theory based on moderator analyses

	Memory Precursor Account		Curiosity Precursor Account	
	Specific PE	General PE	Specific PE	General PE
Prequestion Guess	✓		✓	
Length of Time per Prequestion	N			
Number of Prequestions	N	N		
Interleaved Study	N	✓		
Prequestion Test Format	N		N	N
Type of Learning Materials		✓		
Relation between Tested and Nontested Material		✓		
Prequestion Knowledge Type			X	N
Prequestion Performance			N*	

✓ = Supporting evidence; N= Null evidence; X = Opposite evidence at $p = .073$; N* = Null evidence at $p = .094$.

familiarity with the prequestions fuels the general effect (if one exists) and that some other mechanism, be it curiosity or a yet-unspecified process, is responsible for the specific effect. Future researchers are encouraged to continue exploring the mechanisms underlying the PE and should consider supporting evidence for the specific effect separately from the evidence supporting the general effect.

Problems with the attention-based account

A goal of this paper was to provide *testable* predictions of the attention-based account, but the predictions produced here might occasionally be considered open to interpretation. The literature is primarily composed of experiment-specific explanations describing the pattern of results of individual studies. Indeed, the vagueness of existing explanations makes theory testing difficult.

A prime example is that *contrasting* patterns of results have sometimes been used to provide support for *the same account*. Eye-tracking, which can provide compelling evidence for *how* participants orient their attention during encoding, has provided inconsistent data across studies. For example, using eye-tracking data, Rothkopf and Billington (1979) showed that participants who received prequestions spent *more* time processing the prequestion-relevant text than control participants, possibly because participants perceived the tested material as being particularly important (see also Lewis & Mensink, 2012). This result was taken as providing support for the idea that answering prequestions promotes learning by increasing learners' attention toward the tested material. However, other researchers have found the *opposite* pattern of results, such that prequestions *shortened* the time that participants spent processing the tested material compared to control participants (Cerdán et al., 2009; Little & Bjork, 2016; McCrudden et al., 2005). In these cases, researchers have argued that prequestions make learning more *efficient* because the tested material was *primed* during the pre-study phase. Together, these data suggest that prequestions *do* influence how participants encode the tested material (i.e., prequestions can *slow* or *speed* the processing of tested materials); but in our opinion, that contradictory results have been interpreted to support *the same hypothesis* is deeply problematic. Proponents of the attention-based account should make explicit and preferably preregistered predictions so that opposite results cannot *both* support the attention-based account; otherwise, the account becomes unfalsifiable.

Beyond the specific effect (which appears to be quite robust), the mechanisms underlying the *general effect* should be regarded with more caution. Following the results of this meta-analysis, it remains unclear whether prequestions benefit the learning of nontested material at all. At present, the theory development underscoring the general effect is primarily based on a handful of studies that have successfully demonstrated a positive general effect (e.g., Carpenter & Toftness, 2017; St. Hilaire & Carpenter, 2020a). Although the two precursor accounts were used to fit the meta-analytic data to both the specific and general effects, it is unlikely that these accounts are exhaustive, and, as outlined above, it is possible that the specific and general effects will depend on different theoretical mechanisms. We hope that by outlining the possible mechanisms with testable predictions, our study will encourage researchers to move beyond simple demonstrations of the PE and towards more theoretical evaluations.

Empirical moderators

Although they were not theoretically relevant to the attention-based account, a few empirically relevant variables moderated the PE. Studies with an experimenter-paced *study* phase produced a larger PE than those with a self-paced study phase. Initially, this pattern seemed contradictory to what might be expected (as self-paced study phases do not restrict how long participants can review the study materials). However, it is possible that providing a fixed-duration study phase had the opposite effect on learning: Instead of functioning as the *maximum* amount of time that participants could study (i.e., learners have *difficulty* reviewing all the material before the study time expires), an experimenter-paced study phases might behave more like a *minimum* (i.e., participants were forced to study *longer* than they might have otherwise). Students' self-regulated study decisions are based on metacognitive indicators that are sometimes error-prone (Bjork et al., 2013), and some researchers have shown that study time is positively related to student grades (Doumen et al., 2014). Thus, it is possible that experimenter-paced study phases led to better learning than self-paced study phases because they encouraged participants to study longer than they would have otherwise. However, because researchers using self-paced study phases generally have not measured or reported the duration with which participants spent studying the material, we did not have the data to assess this hypothesis. As such, we advise researchers to report how much time participants spend during the study phase for both self-paced and experimenter-paced study phases.

As described previously, the specific PE was reduced for conceptual prequestions ($g = .28$) compared with factual prequestions ($g = .58$), although the difference was not significant, $p = .073$. For some readers, this might be sufficient evidence that prequestions do not meaningfully support student learning outcomes because prequestions preferentially support the acquisition of specific facts over the ability to answer more conceptually-rich and ecologically valid questions. However, such a conceptualization would require one to ignore the fact that answering conceptual prequestions still benefited learning relative to control. Moreover, a recent meta-analysis on the transfer of learning showed that an important precondition to successful transfer of knowledge is having sufficient background knowledge of the to-be-transferred concepts (see Pan & Rickard, 2018). Therefore, factual prequestions could be incorporated into the classroom to help students better learn the necessary background knowledge, which would in turn better prepare students to engage in the conceptual or inferential transfer.

Finally, a somewhat curious finding showed that the magnitude of the *general effect* was positively associated with publication year, a pattern that was absent for the specific effect. Indeed, studies conducted before 2006[2] showed a small negative general PE, $g = -0.09$, $p = .118$, $k = 49$, whereas those conducted since 2006 showed the opposite, $g = 0.15$, $p < .001$, $k = 42$. Given the myriad differences between studies conducted before and during the modern era (e.g., paper-and-pencil vs. computer administration), it is not clear what exactly contributes to the rise in the general effect over time. But one possibility is that researchers' interest in the PE has shifted over time from demonstrations of the specific effect (which has been shown in dozens of papers) toward discovering when and whether prequestions might lead to a general benefit. Recent manipulations of the PE have started to incorporate design characteristics that have been hypothesized to produce a positive general effect (e.g., asking nontested retention questions that are related to the prequestions; Little & Bjork, 2016). Therefore, an increase in the magnitude of the general effect over time might reflect how research interest in the PE has evolved rather than being a genuine cause for concern. As we noted earlier, because the overall effect size of the general PE is essentially zero, the present data suggest that prequestions do not benefit learning as a whole but rather learning of the queried material specifically. The positive trend toward observing a positive general PE makes this conclusion less definitive.

Future directions

Although considerable research has examined the learning benefits of prequestions, a number of questions remain. First, *what happens to the PE when there is a delay between the pre-study and study phases?* Despite its theoretical relevance for the specific and general effects, only two known studies have manipulated the delay between the pre-study and study phases (Little, 2021; Salmon et al., 1977), and we had to dropped this variable from our analyses.

Second, *does the magnitude of the PE depend on prequestion performance?* A growing body of research is starting to explore the role of *prior knowledge* on subsequent learning (see Witherby & Carpenter, 2022); it is possible that varying degrees of prior knowledge (i.e., higher versus lower prequestion performance) could differentially contribute to the magnitude of the PE. Because of its direct relevance to the PE, and particularly the curiosity precursor account, we stress that researchers should always report prequestion performance and describe whether the correctly answered prequestions were included or excluded at the time of retrieval. Moreover, because the unit of variation in a moderator analysis is studies rather than individuals, a meta-analysis might not be ideal to address this question.

Third, *do motivational factors influence the magnitude of the PE?* Oftentimes, a prequestion condition is compared to a no-prequestions control condition with little regard to the instructions that control participants receive. After all, vague study instructions in a low-incentive research laboratory setting might not encourage the control participants to engage in

[2] We selected 2006 as the year to split the data given that the present era of prequestion research arguably coincided with the advent of interest in educationally relevant memory phenomena, which occurred in the mid-2000s.

the experimental task to a similar degree as those participants who completed a pretest before the study phase. It is possible that the magnitude of the PE will be attenuated when control participants are incentivized or provided with motivating before the study phase (Geller et al., 2017). Researchers should also consider higher-stakes environments as well, such as the classroom, which could affect student motivation in a different manner than in the research laboratory (see Duchastel, 1983). Although the PE has direct implications for education, few studies have explored the benefits of prequestions in the classroom (Carpenter et al., 2018; Hollen, 1970; Jersild, 1929) or online learning. Efforts should be taken to determine whether prequestion interventions support student learning outcomes in ecologically valid environments with real incentives (see Diederich et al., 2018; Grover et al., 2009).

Fourth, *how might individual and contextual differences affect the PE?* Most researchers have explored the magnitude of the PE with college undergraduates, and it is clear that the specific effect is robust in this population. However, substantially fewer studies have established the reliability of the effect with younger or older populations. The data presented in this meta-analysis were inconclusive as the confidence intervals surrounding the PE for children were quite large, and only a handful of studies have assessed the magnitude of the PE with older adults (so that it was inadvisable to place these studies into a separate level in the moderator). Although the magnitude of the specific effect was numerically larger for adults than for grade schoolers, this effect was not robust and deserved additional attention. Further, despite the purported role that attention plays in the PE, very little research has investigated whether the benefits of prequestions depend on learners' attentional control processes or working memory capacity (de la Riva & Ryan, 2015; Kaakinen et al., 2003; Zimmerman, 1990). This is a particularly important topic given that online testing is becoming increasingly common despite the fact that students often show poorer learning outcomes in online classes than in-person classes (Chan & Ahn, 2023), possibly because of the myriad distractions they face in online learning environments (Francis et al., 2019). There has similarly been very little research exploring the PE with participants of different races and ethnicities, or whether gender plays a role in the effect. In fact, researchers have rarely reported demographic information about their participants. It is possible that the effect is robust against participant heterogeneity, but this remains to be determined. Researchers are encouraged to examine the PE with diverse populations and to report relevant participant demographic information (see American Psychological Association, 2019).

Conclusion

The results of this meta-analysis showed that the PE is robust but largely specific to the tested material. When new questions were asked at retrieval (i.e., the general effect), the prequestion benefit was virtually nil—although this conclusion is somewhat tentative given that the general effect has been increasing over time. Results from moderator analyses provided modest support for the widely postulated, but severely underspecified, attention-based explanations of prequestions. In this paper, we attempted to address how encountering prequestions might promote attentive learning by first proposing and then testing the *memory* and *curiosity* precursor accounts. Overall, the memory precursor account received moderate support from the moderator analyses, particularly for the general effect. In contrast, the curiosity precursor account—which clearly needs further development—received only weak support from the present data. It remains to be determined whether one of these accounts can be refined and developed enough to provide satisfactory explanations for *both* the specific and general effects, or if the two accounts are best used in tandem. It is of course possible that a yet unproposed account might be required to fully address the PE. Lastly, although meta-analyses have the advantage of aggregating data at a large scale, they cannot replace carefully devised empirical work for critical theory testing. We hope that this review can serve as a springboard to help steer further theoretically motivated work in the literature.

Supplementary information The online version contains supplementary material available at https://doi.org/10.3758/s13423-023-02353-8.

Author note This paper is based on the dissertation of Kyle St. Hilaire at Iowa State University. Kyle St. Hilaire is now at Chegg, Inc. This work is partly supported by the National Science Foundation Science of Learning and Augmented Intelligence Grant 2017333.

Data Availability All data and analysis scripts for this study was included on the OSF page listed in the Method section of this paper.

Declarations

Conflict of interest We have no conflict of interest to report.

References

Note: References marked with an asterisk were included in the meta-analysis.

Ahn, D., & Chan, J. C. K. (in press). Does testing potentiate new learning because it enables learners to use better strategies? *Journal of Experimental Psychology: Learning, Memory, and Cognition*.

Ahn, D., & Chan, J. C. K. (2022). Does testing enhance new learning because it insulates against proactive interference? *Memory & Cognition, 50*(8), 1664–1682. https://doi.org/10.3758/s13421-022-01273-7

Al Rasheed, H. S. S. (2014). Examining the effectiveness of pre-reading strategies on Saudi EFL college students' reading

comprehension. *English Language Teaching, 7*(11), 79–91. https://doi.org/10.5539/elt.v7n11p79
American Psychological Association. (2019). APA guidelines on race and ethnicity in psychology: Promoting responsiveness and equity. https://www.apa.org/about/policy/guidelines-race-ethnicity.pdf
*Anderson, E. J. (1978). The effects of combined learning cues on teacher learning outcomes. *Journal of Research in Science Teaching, 15*(3), 221–226.
Andrew, L. D. (1975). *The effects of three types of prequestions and type of posttest on student achievement* (Publication No. 302760731) [Doctoral dissertation, Brigham Young University]. ProQuest Dissertations & Theses Global.
Balduzzi, S., Rücker, S., & Schwarzer, G. (2019). How to perform a meta-analysis with R: A practical tutorial. *Evidence-Based Mental Health, 22*(4), 153–160.
Bauer, C. A. (1976). *The effects of a pretesting session on the scores obtained by adults on a nonverbal test of intelligence* (Publication No. 302775240) [Doctoral dissertation, Loyola University Chicago]. ProQuest Dissertations & Theses Global.
*Benya, R. A. (1980). *The effect of the placement of guiding questions for listening passages on the retention of factual material by third quarter college Spanish students* (Publication No. 303056025) [Doctoral dissertation, The Ohio State University]. ProQuest Dissertations & Theses Global.
Berlyne, D. E. (1954). A theory of human curiosity. *British Journal of Psychology, 45*, 180–191. https://doi.org/10.1111/j.2044-8295.1954.tb01243.x
Berlyne, D. E. (1954). An experimental study of human curiosity. *British Journal of Psychology, 45*(4), 256–265. https://doi.org/10.1111/j.2044-8295.1954.tb01253.x
Berlyne, D. E. (1962). Uncertainty and epistemic curiosity. *British Journal of Psychology, 53*(1), 27–34. https://doi.org/10.1111/j.2044-8295.1962.tb00811.x
Berlyne, D. E. (1966). Conditions of prequestioning and retention of meaningful material. *Journal of Educational Psychology, 57*(3), 128–132. https://doi.org/10.1037/h0023346
Bjork, R. A., Dunlosky, J., & Kornell, N. (2013). Self-regulated learning: Beliefs, techniques, and illusions. *Annual Review of Psychology, 64*, 417–444. https://doi.org/10.1146/annurev-psych-113011-143823
*Boker, J. R. (1974). Immediate and delayed retention effects of interspersing questions in written instructional passages. *Journal of Educational Psychology, 66*(1), 96–98. https://doi.org/10.1037/h0036072
*Bos, F., Terlouw, C., & Pilot, A. (2009). The effect of a pretest in an interactive, multimodal pretraining system for learning science concepts. *Educational Research and Evaluation, 15*(6), 571–590. https://doi.org/10.1080/13803610903458659
Brodzinski, J. D. (1983). *An investigation of the effects of questionnaire length, question position, and respondent motivation on item omission, response quantity, scale score, and scale standard deviation* (Publication No. 303187788) [Doctoral dissertation, Ohio University]. ProQuest Dissertations & Theses Global.
*Bull, S. G., & Dizney, H. F. (1973). Epistemic-curiosity-arousing prequestions: Their effect on long-term retention. *Journal of Educational Psychology, 65*(1), 45–49. https://doi.org/10.1037/h0034817
Butler, A. C. (2018). Multiple-choice testing in education: Are the best practices for assessment also good for learning? *Journal of Applied Research in Memory and Cognition, 7*(3), 323–331. https://doi.org/10.1016/j.jarmac.2018.07.002
*Carpenter, S. K., Rahman, S., & Perkins, K. (2018). The effects of prequestions on classroom learning. *Journal of Experimental Psychology: Applied, 24,* 34-42https://doi.org/10.1037/xap0000145
*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
Cerdán, R., Vidal-Abarca, E., Martínez, T., Gilabert, R., & Gil, L. (2009). Impact of question-answering tasks on search processes and reading comprehension. *Learning and Instruction, 19*(1), 13–27. https://doi.org/10.1016/j.learninstruc.2007.12.003
Chan, J. C. K. (2009). When does retrieval induce forgetting and when does it induce facilitation? Implications for retrieval inhibition, testing effect, and text processing. *Journal of Memory and Language, 61*(2), 153–170. https://doi.org/10.1016/j.jml.2009.04.004
Chan, J. C. K., & Ahn, D. (2023). Unproctored online exams provide meaningful assessment of student learning. In *Proceedings of the National Academy of Sciences of the United States of America, 120*(31), e2302020120, https://doi.org/10.1073/pnas.2302020120
Chan, J. C. K., McDermott, K. B., & Roediger, H. L. (2006). Retrieval-induced facilitation: Initially nontested material can benefit from prior testing of related material. *Journal of Experimental Psychology: General, 135*(4), 553–571. https://doi.org/10.1037/0096-3445.135.4.553
Chan, J. C. K., Manley, K. D., & Lang, K. (2017). Retrieval-enhanced suggestibility: A retrospective and a new investigation. *Journal of Applied Research in Memory and Cognition, 6*(3), 213–229. https://doi.org/10.1016/j.jarmac.2017.07.003
Chan, J. C. K., Meissner, C. A., & Davis, S. D. (2018). Retrieval potentiates new learning: A theoretical and meta-analytic review. *Psychological Bulletin, 144*, 1111–1146. https://doi.org/10.1037/bul0000166
Chan, J. C. K., Manley, K. D., & Ahn, D. (2020). Does retrieval potentiate new learning when retrieval stops but new learning continues? *Journal of Memory and Language, 115*, 104150. https://doi.org/10.1016/j.jml.2020.104150
Chan, J. C. K., O'Donnell, R., & Manley, K. D. (2022). Warning weakens retrieval-enhanced suggestibility only when it is given shortly after misinformation: The critical importance of timing. *Journal of Experimental Psychology: Applied, 28*, 694–716. https://doi.org/10.1037/xap0000394
*Chobot, M. C. (1984a). *The interactive effects of field dependence and adjunct questions on learning from prose by library/information science students* (Publication No. 303288000) [Doctoral dissertation, The Catholic University of America]. ProQuest Dissertations & Theses Global.
Cook, D. H. (1984). *The development and the evaluation of a diagnostic mathematics pretest for chemistry and of a program to strengthen mathematics proficiencies for chemistry students* (Publication No. 303315322) [Doctoral dissertation, Temple University]. ProQuest Dissertations and Theses Global.
Craik, F. I. M., & Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General, 104*(3), 268–294. https://doi.org/10.1037/0096-3445.104.3.268
Cumming, G. (2012). *Understanding the new statistics*. Routledge.
Dalton, D. W., & Goodrum, D. A. (1991). The effects of computer based pretesting strategies on learning and continuing motivation. *Journal of Research on Computing in Education, 24*(2), 204–213. https://doi.org/10.1080/08886504.1991.10782002
de la Riva, S., & Ryan, T. G. (2015). Effect of self-regulating behaviour on young children's academic success. *International Journal of Early Childhood Special Education, 7*(1), 69–96.
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

*Denner, P. R., & McGinley, W. J. (1992). Effects of prereading activities on junior high students' recall. *The Journal of Educational Research, 86*(1), 11–19.https://doi.org/10.1080/00220671.1992.9941822

DerSimonian, R., & Laird, N. (1986). Meta-analysis in clinical trials. *Controlled Clinical Trials, 7*(3), 177–188. https://doi.org/10.1016/0197-2456(86)90046-2

*Dickerson, C. T. (1988). *The effects of adjunct questions on learning from text inconsistent with prior knowledge* (Publication No. 303587896) [Doctoral dissertation, North Carolina State University]. ProQuest Dissertations & Theses Global.

Diederich, E., Thomas, L., Mahnken, J., & Lineberry, M. (2018). Pretest scores uniquely predict 1-year-delayed performance in a simulation-based mastery course for central line insertion. *Simulation in Healthcare: The Journal of the Society for Simulation in Healthcare, 13*(3), 163–167. https://doi.org/10.1097/SIH.0000000000000327

Doumen, S., Broeckmans, J., & Masui, C. (2014). The role of self-study time in freshmen's achievement. *Educational Psychology, 34*(3), 385–402. https://doi.org/10.1080/01443410.2013.785063

*Dowaliby, F. J. (1992). The effects of adjunct questions in prose for deaf and hearing students at different reading levels. *American Annals of the Deaf, 137*(4), 338–344. https://doi.org/10.1353/aad.2012.0426

Duchastel, P. C. (1983). Interpreting adjunct question research: Processes and ecological validity. *Human Learning: Journal of Practical Research & Applications, 2*(1), 1–5.

Duncan, N., & Harris, C. (1974). An index of effectiveness for criterion-referenced items used in pre-tests and posttests. *Programmed Learning and Educational Technology, 11*, 125–132.

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*, 4–58. https://doi.org/10.1177/1529100612453266

Egger, M., Davey Smith, G., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *British Medical Journal, 315*(7109), 629–634. https://doi.org/10.1136/bmj.315.7109.629

Engel, S., & Randall, K. (2009). How teachers respond to children's inquiry. *American Educational Research Journal, 46*, 183–202. https://doi.org/10.3102/0002831208323274

*Feil, P. H. (1977). *A quasi-experimental analysis of the effects of adjunct questions on retention* (Publication No. 302816658) [Doctoral dissertation, Indiana University]. ProQuest Dissertations & Theses Global.

Feiman, J. R. (1974). *Effect of type and location of questions and acoustic similarity on short-term retention of prose material* (Publication No. 302713605) [Doctoral dissertation, Case Western Reserve University]. ProQuest Dissertations & Theses Global.

*Felker, D. B., & Dapra, R. A. (1975). Effects of question type and question placement on problem-solving ability from prose material. *Journal of Educational Psychology, 67*(3), 380–384. https://doi.org/10.1037/h0076616

Fowler, E. D., & Lamberg, W. J. (1979). Effect of pre-questions on oral reading by elementary students. *Reading Improvement, 16*(1), 71–74.

Francis, M. K., Wormington, S. V., & Hulleman, C. (2019). The costs of online learning: Examining differences in motivation and academic outcomes in online and face-to-face community college developmental mathematics courses. *Frontiers in Psychology, 10*, 2054. https://doi.org/10.3389/fpsyg.2019.02054

Frase, L. T. (1967). Learning from prose material: Length of passage, knowledge of results, and position of questions. *Journal of Educational Psychology, 58*(5), 266–272. https://doi.org/10.1037/h0025028

*Frase, L. T., Patrick, E., & Schumer, H. (1970). Effect of question position and frequency upon learning from text under different levels of incentive. *Journal of Educational Psychology, 61*(1), 52–56. https://doi.org/10.1037/h0028823

Furumoto, Y., & Matsumi, N. (2009). The effects of pre-reading questions and reading time on the recall and comprehension of explanatory text: A comparison of junior high-school students and university students. *The Science of Reading, 52*(2), 49–59. https://doi.org/10.19011/sor.52.2_49

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*(42). https://doi.org/10.1186/s41235-017-0078-z

Gottlieb, J., Oudeyer, P.-Y., Lopes, M., & Baranes, A. (2013). Information-seeking, curiosity, and attention: Computational and neural mechanisms. *Trends in Cognitive Sciences, 17*(11), 585–593.

Grimaldi, P. J., & Karpicke, J. D. (2012). When and why do retrieval attempts enhance subsequent encoding? *Memory & Cognition, 40*, 505–513.

Grossnickle, E. M. (2016). Disentangling curiosity: Dimensionality, definitions, and distinctions from interest in educational contexts. *Educational Psychology Review, 28*, 23–60. https://doi.org/10.1007/s10648-014-9294-y

Grover, G., Heck, J., & Heck, N. (2009). Pretest in an introductory finance course: Value added? *Journal of Education for Business, 85*(2), 64–67. https://doi.org/10.1080/08832320903252405

Gruber, M. J., Gelman, B. D., & Ranganath, C. (2014). States of curiosity modulate hippocampus-dependent learning via the dopaminergic circuit. *Neuron, 84*, 486–496. https://doi.org/10.1016/j.neuron.2014.08.060

Gruber, M. J., Ritchey, M., Wang, S. F., Doss, M. K., & Ranganath, C. (2016). Post-learning hippocampal dynamics promote preferential retention of rewarding events. *Neuron, 89*(5), 1110–1120. https://doi.org/10.1016/j.neuron.2016.01.017

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

*Haimowitz, C. (1972). *Effect of adjunct questions on learning from prose at the junior college level by reading or listening with rate controlled* (Publication No. 302617333) [Doctoral dissertation, Rutgers University]. ProQuest Dissertations & Theses Global.

*Halpain, D. R., Glover, J. A., & Harvey, A. L. (1985). Differential effects of higher and lower order questions: Attention hypotheses. *Journal of Educational Psychology, 77*(6), 703–715. https://doi.org/10.1037/0022-0663.77.6.703

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. https://doi.org/10.2307/1170407

Hannafin, M. J., & Hughes, C. W. (1986). A framework for incorporating orienting activities in computer-based interactive video. *Instructional Science, 15*(3), 239–255. https://doi.org/10.1007/BF00139613

Harrer, M., Cuijpers, P., Furukawa, T. A., & Ebert, D. D. (2021). *Doing meta-analysis with R: A hands-on guide*. Chapman & Hall/CRC Press.

Hartley, J., & Davies, I. K. (1976). Preinstructional strategies: The role of pretests, behavioral objectives, overviews and advance organizers. *Review of Educational Research, 46*(2), 239–265. https://doi.org/10.3102/00346543046002239

Hartley, J., Holt, J., & Swain, F. (1970). The effects of pre-tests, interim tests, and age on post-test performance following self-instruction. *Programmed Learning and Educational Technology, 7*(4), 250–256. https://doi.org/10.1080/1355800700070402

*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*, 290–296.

Herman, W. E. (2010). *How much do students remember from an introductory psychology course? [Conference paper]*. Annual Conference on the Teaching of Psychology: Ideas & Innovations. https://files.eric.ed.gov/fulltext/ED509985.pdf

Higgins, J. P. T., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M. J., & Welch, V. A. (2022). *Cochrane Handbook for Systematic Reviews of Interventions*. Cochrane. https://www.training.cochrane.org/handbook

Hill, G. P. (1986). *The levels-of-question effect on intentional learning* (Publication No. 303367861) [Doctoral dissertation, Catholic University of America]. ProQuest Dissertations & Theses Global.

*Hillman, S. B. (1979). *Information processing as a function of question type and position* [Conference paper]. The Annual Meeting of the American Educational Research Association, San Francisco, CA, United States.

Hintzman, D. L. (2011). Research strategy in the study of memory: Fads, fallacies, and the search for the "Coordinates of Truth."*Perspectives on Psychological Science, 6*(3), 253–271. https://doi.org/10.1177/1745691611406924

*Hollen, T. (1970). *Interaction of individual abilities with the presence and position of adjunct questions in learning from prose materials* (Publication No. 7111554) [Doctoral dissertation, University of Texas at Austin]. ProQuest Dissertations & Theses Global.

Jacoby, L. L., Wahlheim, C. N., & Kelley, C. M. (2015). Memory consequences of looking back to notice change: Retroactive and proactive facilitation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 41*(5), 1282–1297. https://doi.org/10.1037/xlm0000123

*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

*Jersild, A. T. (1929). Examination as an aid to learning. *Journal of Educational Psychology, 20*, 602–609. https://doi.org/10.1037/h0070530

*Johnson, D. R. (2015). *The effects of pre-reading relevance instructions and individual interest on learning outcomes and curiosity* (Publication No. 1696060585) [Doctoral dissertation, Northern Illinois University]. ProQuest Dissertations & Theses Global.

Kaakinen, J. K., Hyönä, J., & Keenan, J. M. (2003). How prior knowledge, WMC, and relevance of information affect eye fixations in expository text. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29*(3), 447–457.

Kaplan, R., & Rothkopf, E. Z. (1974). Instructional objectives as directions to learners. *Journal of Educational Psychology, 66*, 448–456.

*Karjala, Y. (1984). *Objectives, questions, and memory for instruction* (Publication No. 303343225) [Doctoral dissertation, Arizona State University]. ProQuest Dissertations & Theses Global.

*Kastelic, T. G. (1977). *Effects of adjunct aids and passage complexity on reading behavior and comprehension* (Publication No. 302823652) [Doctoral dissertation, Pennsylvania State University]. ProQuest Dissertations & Theses Global.

*Kealy, W. A., Bakriwala, D. J., & Sheridan, P. B. (2003). When tactics collide: Counter effects between an adjunct map and prequestions. *Educational Technology Research and Development, 51*(2), 17–39. https://doi.org/10.1007/BF02504524

*Khoynejad, G. (1980). *The effects of behavioral objectives, prequestions, and a combination of both on intentional and incidental learning from written text by secondary school students* (Publication No. 303073359) [Doctoral dissertation, University of Texas, Austin]. ProQuest Dissertations & Theses Global.

*Kirschner, P., & Brink, H. v. d. (1979). *The effect of adjunct questions on learning from a videotape lesson*. (ERIC Document Reproduction Service No. ED 206264).

Knapp, G., & Hartung, J. (2003). Improved tests for a random effects meta-regression with a single covariate. *Statistics in Medicine, 22*(17), 2693–2710. https://doi.org/10.1002/sim.1482

Knight, J. B., Ball, B. H., Brewer, G. A., DeWitt, M. R., & Marsh, R. L. (2012). Testing unsuccessfully: A specification of the underlying mechanisms supporting its influence on retention. *Journal of Memory and Language, 66*, 731–746. https://doi.org/10.1016/j.jml.2011.12.008

Knobloch, S., Patzig, G., Mende, A. M., & Hastall, M. (2004). Affective news: Effects of discourse structure in narratives on suspense, curiosity, and enjoyment while reading news and novels. *Communication Research, 31*(3), 259–287. https://doi.org/10.1177/0093650203261517

Kornell, N., & Vaughn, K. E. (2016). How retrieval attempts affect learning: A review and synthesis. *Psychology of Learning and Motivation, 65*, 183–215. https://doi.org/10.1016/bs.plm.2016.03.003

*Kreiner, D. S. (1996). Effects of advance questions on reading comprehension. *Journal of General Psychology, 123*(4), 352–364. https://doi.org/10.1080/00221309.1996.9921287

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*, 15. https://doi.org/10.1038/s41539-019-0053-1

*Lee, S. Y. (2011). The effects of pre-reading questions on middle school students' reading comprehension. *Modern English Education, 12*(1), 58–72.

Leonard, W. H. (1987). Does the presentation style of questions inserted into text influence understanding and retention of science concepts? *Journal of Research in Science Teaching, 24*(1), 27–37. https://doi.org/10.1002/tea.3660240104

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

Litman, J., Hutchins, T., & Russon, R. (2005). Epistemic curiosity, feeling-of-knowing, and exploratory behaviour. *Cognition & Emotion, 19*(4), 559–582. https://doi.org/10.1080/02699930441000427

Little, J. L. (2011). *Optimizing multiple-choice tests as learning events* (Publication No. 919052703) [Doctoral dissertation, University of California, Los Angeles]. ProQuest Dissertations & Theses Global.

*Little, J. L. (2021).*The prequestion effect when retention interval is manipulated* [Unpublished raw data]. California State University.

*Little, J. L., & Bjork, E. L. (2011) Pretesting with multiple-choice questions facilitates learning. In L. Carlson, C. Hölscher, & T. Shipley, (eds.), In *Proceedings of the 33rd Annual Conference of the Cognitive Science Society* (pp. 294–299). https://cogsci.mindmodeling.org/2011/papers/0058/paper0058.pdf

Little, J. L., & Bjork, E. L. (2016). Multiple-choice pretesting potentiates learning of related information. *Memory and Cognition, 44*(7), 1085–1101. https://doi.org/10.3758/s13421-016-0621-z

Loewenstein, G. (1994). The psychology of curiosity: A review and reinterpretation. *Psychological Bulletin, 116*(1), 75–98. https://doi.org/10.1037/0033-2909.116.1.75

MacLachlan, J., & Jalan, P. (1985). The effect of pre-questions on advertising recall. *Journal of Advertising, 14*(1), 18–22. https://doi.org/10.1080/00913367.1985.10672925

*Manchester, N. E. (1984). *The effects of conceptual level and position of adjunct questions on recall measures of reading comprehension in beginning college French* (Publication No. 303321991) [Doctoral dissertation, The Ohio State University]. ProQuest Dissertations & Theses Global.

Manley, K. D., & Chan, J. C. K. (2019). Does retrieval enhance suggestibility because it increases perceived credibility of the postevent information? *Journal of Applied Research in Memory and Cognition, 8*(3), 355–366. https://doi.org/10.1016/j.jarmac.2019.06.001

Marvin, C. B., & Shohamy, D. (2016). Curiosity and reward: Valence predicts choice and information prediction errors enhance learning. *Journal of Experimental Psychology: General, 145*(3), 266–272. https://doi.org/10.1037/xge0000140

*McCrudden, M. T., Schraw, G., & Kambe, G. (2005). The effect of relevance instructions on reading time and learning. *Journal of Educational Psychology, 97*(1), 88–102. https://doi.org/10.1037/0022-0663.97.1.88

McDermott, K. B. (2021). Practicing retrieval facilitates learning. *Annual Review of Psychology, 72*, 609–633. https://doi.org/10.1146/annurev-psych-010419-051019

*Memory, D. M. (1981). An extended comparison of adjunct aids with low-average and good readers in the sixth grade. *The Journal of Educational Research, 74*(6), 405–410https://doi.org/10.1080/00220671.1981.10885340

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

Metcalfe, J. (2009). Metacognitive judgments and control of study. *Current Directions in Psychological Science, 18*(3), 159–163. https://doi.org/10.1111/j.1467-8721.2009.01628.x

Metcalfe, J., & Finn, B. (2011). People's hypercorrection of high-confidence errors: Did they know it all along? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37*(2), 437–448. https://doi.org/10.1037/a0021962

*Miyagi, P. S. (1995). *The effect of level and location of adjunct questions* (Publication No. 304235846) [Doctoral Dissertation, University of Texas, Austin]. ProQuest Dissertations & Theses Global.

Morasky, R. L. (1972). Eye movements as a function of adjunct question placement. *American Educational Research Journal, 9*(2), 251–261. https://doi.org/10.2307/1161686

Morasky, R. L., & Willcox, H. H. (1970). Time required to process information as a function of question placement. *American Educational Research Journal, 7*(4), 561–567. https://doi.org/10.2307/1161837

*Mottley, R. R. (1972). *The effect of prequestioning versus no prequestioning reading conditions upon subsequent comprehension scores* (Publication No. 302618032) [Doctoral dissertation, University of Georgia]. ProQuest Dissertations & Theses Global.

Noakes, A. M. (1969). *The effects of conditions of pre-questioning upon comprehension of fiction and non-fiction selections with fifth grade children* (Publication No. 302412966) [Doctoral dissertation, University of Delaware]. ProQuest Dissertations & Theses Global.

*Osman, M., & Hannafin, M. (1994). Effects of advance questioning and prior knowledge on science learning. *The Journal of Educational Research, 88*(1), 5–13.https://doi.org/10.1080/00220671.1994.9944829

Oudeyer, P.Y., Gottlieb, J., & Lopes, M. (2016). Intrinsic motivation, curiosity, and learning. In B. Studer & S. Knecht (eds.), *Motivation—Theory, neurobiology and applications: Progress in brain research* (vol. 229, pp. 257–284). Elsevier. https://doi.org/10.1016/bs.pbr.2016.05.005

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., & 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://doi.org/10.1037/xap0000345

*Pan, S. C., Schmitt, A. G., Bjork, E. L., & Sana, F. (2020). 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

*Papay, J. P. (1971). *An investigation of the effects of type, location, and distribution of orienting stimuli on the acquisition and retention of meaningful prose materials* (Publication No. 302494708) [Doctoral dissertation, Florida State University]. ProQuest Dissertations & Theses Global.

Pashler, H., Bain, P. M., Bottge, B. A., Graesser, A., Koedinger, K., McDaniel, M., & Metcalfe, J. (2007). *Organizing instruction and study to improve student learning. IES practice guide. NCER 2007-2004.* National Center for Education Research, Institute of Education Science, U.S. Department of Education. https://eric.ed.gov/?id=ED498555

*Patrick, E. M. (1976). *The effects of variations in the uncertainty of prequestions and information feedback on the short and long term retention of prose materials* (Publication No. 302822125) [Doctoral dissertation, University of Massachusetts]. ProQuest Dissertations & Theses Global.

*Peeck, J. (1970). Effect of prequestions on delayed retention of prose material. *Journal of Educational Psychology, 61*(3), 241–246. https://doi.org/10.1037/h0029104

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*, 27–35. https://doi.org/10.1016/0361-476X(90)90003-J

Rahman, S. (2017). *The effects of prequestions on classroom learning* (Publication No. 1984981819) [Master's thesis, Iowa State University]. ProQuest Dissertations & Theses Global.

*Rauscher, W. S. (1978). *An investigation into the facilitative effects of two kinds of adjunct questions on the comprehension of students at two grade levels during the reading of textual materials* [Doctoral dissertation, University of Iowa]. ProQuest Dissertations & Theses Global.

Reyna, V. F., & Brainerd, C. J. (1995). Fuzzy-trace theory: An interim synthesis. *Learning and Individual Differences, 7*, 1–75. https://doi.org/10.1016/1041-6080(95)90031-4

*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

Rickards, J. P. (1976a). 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. (1976b). Interaction of position and conceptual level of adjunct questions on immediate and delayed retention of text. *Journal of Educational Psychology, 68*, 210–217. https://doi.org/10.1037/0022-0663.68.2.210

*Rickards, J. P., Anderson, M. C., & McCormick, C. B. (1976). Processing effects of common-word and number questions inserted in reading materials. *The Journal of Educational Research, 69*(7), 274–277. https://doi.org/10.1080/00220671.1976.10884897

*Rickards, J. P., & McCormick, C. B. (1988). Effect of interspersed conceptual prequestions on note-taking in listening comprehension. *Journal of Educational Psychology, 80*(4), 592–594. https://doi.org/10.1037/0022-0663.80.4.592

*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. https://doi.org/10.3102/00028312003004241

Rothkopf, E. Z., & Billington, M. J. (1979). Goal-guided learning from text: Inferring a descriptive processing model from inspection times and eye movements. *Journal of Educational Psychology, 71*(3), 310–327. https://doi.org/10.1037/0022-0663.71.3.310

*Rowls, M. D. (1975). *The facilitative and interactive effects of adjunct questions on retention of eighth graders across three prose passages* (Publication No. 302719621) [Doctoral dissertation, Indiana University]. ProQuest Dissertations & Theses Global.

*Sagaria, S. D., & Di Vesta, F. J. (1978). Learner expectations induced by adjunct questions and the retrieval of intentional and incidental information. *Journal of Educational Psychology, 70*(3), 280–288. https://doi.org/10.1037/0022-0663.70.3.280

*Salmon, R., Hines, M. L., Benson, I., Newell, J. (1977, April). *The advanced organizer concept: Some methodological questions* [Conference paper]. American Educational Research Association Meeting. https://eric.ed.gov/?id=ED150183

*Sana, F. (2021a). *Prequestions with psychological concepts* [Unpublished raw data]. Athabasca University.

*Sana, F. (2021b). *The prequestion effect and the order of answer presentation* [Unpublished raw data]. Athabasca University.

Sana, F., Yan, V. X., Clark, C. M., Bjork, E. L., & Bjork, R. A. (2020a). Improving conceptual learning via pretests. *Journal of Experimental Psychology: Applied, 27*(2), 228–236. https://doi.org/10.1037/xap0000322

Sana, F., Forrin, N. D., Sharma, M., Dubljevic, T., Ho, P., Jalil, E., & Kim, J. A. (2020b). Optimizing the efficacy of learning objectives through pretests. *CBE—Life Sciences Education, 19*(3). https://doi.org/10.1187/cbe.19-11-0257

*Santora, O. M. (1972). *The effect of question position, pacing, and mode of presentation on learning from written prose in elementary school children* (Publication No. 302613885) [Doctoral dissertation, State University of New York at Albany]. ProQuest Dissertations & Theses Global.

*Shanahan, T. (1986).Predictions and the limiting effects of prequestions. *National Reading Conference Yearbook, 35*, 92–98.

*Shavelson, R. J., Berliner, D. C., Ravitch, M. M., & Loeding, D. (1974). Effects of position and type of question on learning from prose material: Interaction of treatments with individual differences. *Journal of Educational Psychology, 66*(1), 40–48.https://doi.org/10.1037/h0035813.

Shinogaya, K. (2011). Preparation for meaningful learning: Effects of a method involving answering pre-questions and judging confidence in the answers. *The Japanese Journal of Educational Psychology, 59*(3), 355–366. https://doi.org/10.5926/jjep.59.355

Simkins, S., & Allen, S. (2000). Pretesting students to improve teaching and learning. *International Advances in Economic Research, 6*(1), 100–112. https://doi.org/10.1007/BF02295755

Sinnott, L. T., & Alderman, D. L. (1977). *The effects of pre- and post-questions on learning from textual material in a CAI format* (Technical Report No. 4). Educational Testing Service.

*Smith, D. G. (1976). *Position of inserted questions and ability in learning from prose* (Publication No. 302823758) [Doctoral dissertation, Indiana University]. ProQuest Dissertations & Theses Global.

*Smith, B. A. (1996). Preadjunct questions as a learning strategy for older adults: Help or hindrance? *Educational Gerontology, 22*(7), 625–636. https://doi.org/10.1080/0360127960220701

*Snowman, J., & Cunningham, D. J. (1975). A comparison of pictorial and written adjunct aids in learning from text. *Journal of Educational Psychology*, *67*(2), 307–311. https://doi.org/10.1037/h0076934

*St. Hilaire, K. J., Carpenter, S. K. (2020a). Prequestions enhance learning, but only when they are remembered. *Journal of Experimental Psychology: Applied, 26,* 705–716. https://doi.org/10.1037/xap0000296.

*St. Hilaire, K. J., Carpenter, S. K. (2020b). *How question-type affects learning from prequestions versus learning objectives* [Conference poster]. The Annual Meeting of the Psychonomic Society, Virtual Conference.

*St. Hilaire, K. J., Carpenter, S. K. (2021). *Does noticing prequestions enhance learning?* [Unpublished raw data]. Iowa State University.

*St. Hilaire, K. J. (2017). *The pretesting effect: How question-type and structure building ability impact learning* (Publication No. 1954045888) [Master's Thesis, Wake Forest University]. ProQuest Dissertations & Theses Global.

*St. Hilaire, K. J., Blumenthal, S. A., Butkevits, N. M., Geiger, O. J., Fisenne, S. E., & Jennings, J. M. (2016, November). *The pretesting effect: A limited efficacy learning technique?* [Conference poster]. The Annual Meeting of the Psychonomic Society, Boston, MA, United States.

*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,* 1204–1213.https://doi.org/10.1080/09658211.2019.1641209.

Stanley, T. D., & Doucouliagos, H. (2014). Meta-regression approximations to reduce publication selection bias. *Research Synthesis Methods, 5*(1), 60–78. https://doi.org/10.1002/jrsm.1095

*Stolz, V. M. (1985). *The effects of the position and level of questions on the retention of prose material* (Publication No. 303174726) [Doctoral dissertation, University of Regina]. ProQuest Dissertations & Theses Global.

*Strollo, J. C. (1972). *The effects of varied positions of study guide questions on the immediate and delayed retention of a selected sample of handicapped learners* (Publication No. 302609622) [Doctoral dissertation, Pennsylvania State University]. ProQuest Dissertations & Theses Global.

Swenson, I., & Kulhavy, R. (1973, February). *Retention of item comprehension in prose as a function of question placement and pacing* [Conference paper]. Annual Meeting of the American Educational Research Association.

*Swenson, I., & Kulhavy, R. (1974). Adjunct questions and the comprehension of prose by children. *Journal of Educational Psychology, 66*(2), 212–215. https://doi.org/10.1037/h0036276

Tanaka, S., & Miyatani, M. (2015). Incorrect answer in pretest and memory fixation. *Learning System Research, 1*, 4–15. https://doi.org/10.15027/39105

*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 & Cognition, 7,* 370–378. https://doi.org/10.1016/j.jarmac.2018.06.003

*Truog, A. L. (1977). *Prose learning as a function of adjunct questions and individual differences* (Publication No. 302821962) [Doctoral dissertation, University of Wisconsin, Madison]. ProQuest Dissertations & Theses Global.

van Aalst, I. (2021). *The effect of prequestions in primary education* [Master's thesis, Open Universiteit].

Vaughn, K. E., Hausman, H., & Kornell, N. (2017). Retrieval attempts enhance learning regardless of time spent trying to retrieve. *Memory, 25*, 298–316.

Wade, S., & Kidd, C. (2019). The role of prior knowledge and curiosity in learning. *Psychological Bulletin & Review, 26*, 1377–1387. https://doi.org/10.3758/s13423-019-01598-6

Washburne, J. N. (1929). The use of questions in social science material. *The Journal of Educational Psychology, 20*, 321–359. https://doi.org/10.1037/h0075982

Watts, G. H. (1974). Effect of prequestions on control of attention in written instruction. *Australian Journal of Education, 18*(1), 79–85. https://doi.org/10.1177/000494417401800106

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

White, R. E. (1981). *The effects of organizational themes and adjunct question placements on children's prose learning: A developmental perspective* (Publication No. 303022239) [Doctoral dissertation, Northwestern University]. ProQuest Dissertations & Theses Global.

*Wiesendanger, K., & Wollenberg, J. (1978). Prequestioning inhibits third graders' reading comprehension. *The Reading Teacher, 31*(8), 892–895. https://doi.org/10.2307/20194671

Wilford, M. M., Chan, J. C. K., & Tuhn, S. J. (2014). Retrieval enhances eyewitness suggestibility to misinformation in free and cued recall. *Journal of Experimental Psychology: Applied, 20*(1), 81–93. https://doi.org/10.1037/xap0000001

Wilhite, S. C. (1983). Pre-passage questions: The influence of structural importance. *Journal of Educational Psychology, 75*(2), 234–244. https://doi.org/10.1037/0022-0663.75.2.234

Wilhite, S. C. (1984). Hierarchical importance of pre-passage questions: Effects on cued recall. *Journal of Reading Behavior, 16*(1), 41–59. https://doi.org/10.1080/10862968409547503

*Wilkerson, N. N. (1982). *Effects of two preinstructional strategies on cognitive learning of the Roy Adaptation Model of nursing* (Publication No. 303250442) [Doctoral dissertation, University of Texas, Austin]. ProQuest Dissertations & Theses Global.

Witham, J., & Linehan, P. (1995). *Pretesting: Is it a strategy used with inclusion?* [Conference paper]. Annual Meeting of the Council for Exceptional Children.

Witherby, A. E., & Carpenter, S. K. (2022). The rich-get-richer effect: Prior knowledge predicts new learning of domain-relevant information. *Journal of Experimental Psychology Learning, Memory, and Cognition, 48*(4), 483–498. https://doi.org/10.1037/xlm0000996

*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

Zimmerman, B. J. (1990). Self-regulating academic learning and achievement: The emergence of a social cognitive perspective. *Educational Psychology Review, 2*(2), 173–201. https://doi.org/10.1007/BF01322178