AMERICAN PSYCHOLOGICAL ASSOCIATION

Journal of Educational Psychology

ISSN: 0022-0663

https://doi.org/10.1037/edu0001055

Interpolated Prequestioning Increases On-Task Thoughts, Dampens Off-Task Thoughts, and Facilitates Learning During Video Lectures

Steven C. Pan[1], Elisya Tung[1], Luke Chiang[1], Alyssa Indrajaya[1], and Michelle E. Kaku[2]

[1] Department of Psychology, National University of Singapore

[2] Tokyo Metropolitan Board of Education, JET Programme, Tokyo, Japan

OPEN DATA OPEN MATERIALS PREREGISTERED

Interpolated prequestioning is a learning technique in which practice questions targeting yet-to-be-learned information (i.e., prequestions) are interspersed throughout a set of learning materials. As learners view those materials, they make guesses for each question, then have the opportunity to discover the correct answers in subsequently viewed materials. Emerging evidence suggests that this technique can cause learners to pay closer attention during lectures, yielding better learning outcomes. That possibility has yet to be investigated, however, relative to a competitive nontesting technique, using categorical experience sampling methods, and without the contaminating influence of other, nonlecture learning activities. In each of three experiments, undergraduate students (Experiment 1) or online adult learners (Experiments 2–3) watched a prerecorded video lecture with prequestions or statements containing learning objectives interpolated throughout, responded to categorical thought probes, and then took a criterial test assessing memory and comprehension of the lecture content. Across experiments, regardless of question format, interpolated prequestioning yielded higher rates of lecture-related on-task thoughts than studying interpolated statements. Prequestioning also improved test scores significantly in two of the three experiments, which mediation analyses attributed to higher rates of on-task thoughts. Furthermore, although rates of off-task thinking were relatively low, there were indications that prequestioning dampened comprehension-related off-task and task-unrelated thoughts. Together, these findings suggest that interpolated prequestioning can strengthen focused attention and facilitate learning from lectures.

Educational Impact and Implications Statement

This study examined whether guessing the answers to practice questions that target upcoming content—known as *prequestions*—at multiple points throughout a video lecture can help learners maintain attention during the lecture and, as a result, improve learning. The results showed that doing so did indeed enhance attention and learning, primarily by increasing the proportion of thoughts that learners directed toward the lecture itself. Hence, incorporating prequestions throughout a lecture can be an effective approach for promoting focus and enhancing learning outcomes.

Keywords: prequestioning, pretesting, attention, mind wandering, interpolated testing

Supplemental materials: https://doi.org/10.1037/edu0001055.supp

Kalypso Iordanou served as action editor.

Steven C. Pan https://orcid.org/0000-0001-9080-5651
Elisya Tung https://orcid.org/0009-0006-8199-0588
Luke Chiang https://orcid.org/0009-0009-7577-3722
Alyssa Indrajaya https://orcid.org/0000-0003-1719-2777
Michelle E. Kaku https://orcid.org/0000-0001-9190-3181
Elisya Tung is now at the Ministry of Education, Singapore.

Materials, data, and analysis code are available at the Open Science Framework and can be accessed at https://osf.io/ph2em. All experiments were preregistered (Experiment 1: https://aspredicted.org/cs3s-d27c.pdf; Experiment 2: https://aspredicted.org/nbd2-npzf.pdf; Experiment 3: https://aspredicted.org/r72z-6ky8.pdf). The authors declare no conflict of interest. This research was supported by a National University of Singapore Faculty of Arts & Social Sciences grant to Steven C. Pan. Steven C. Pan presented portions of this research at Ruhr-Universität Bochum and the Würzburg International Autumn School on Lasting Learning in October 2023. Thanks to Faria Sana for useful suggestions and to Judith Schweppe for helpful comments on a draft of this article.

Steven C. Pan served as lead for funding acquisition, supervision, visualization, writing–original draft, and writing–review and editing, contributed equally to investigation and methodology, and served in a supporting role for formal analysis and project administration. Elisya Tung

continued

If you ask a student in the midst of a lecture what they are thinking about, their responses might range from "the lecture content" or "ideas related to the lecture" to "daydreams" or "something else." Such variability is expected: In everyday life, our thoughts tend to oscillate between the external environment and internal musings, as well as between current tasks and unrelated matters (Smallwood & Schooler, 2006). Among students, this phenomenon—known as *mind wandering* or the diversion of attention to *task-unrelated* or *off-task thoughts*—can occur 30% of the time during lectures and other learning activities (Wong et al., 2022). Although off-task thinking can aid creative pursuits (Mooneyham & Schooler, 2013), in educational settings it is often associated with reduced learning outcomes and lower academic performance (Hollis & Was, 2016; Schacter & Szpunar, 2015). It may pose even greater risks during unsupervised learning, such as the viewing of online videos (Pan, Sana, Schmitt, et al., 2020).

Given the negative link between off-task thoughts and learning outcomes—which a recent meta-analysis of 71 empirical studies quantified as $r = -.27$ (Wong et al., 2022)—researchers have investigated ways to improve students' attentional focus during lectures and other activities (e.g., Kane et al., 2017; Klimovich & Richter, 2024; Knopps & Wissman, 2023; Martin et al., 2018; Peterson & Wissman, 2020; Randall et al., 2022). A promising approach is to engage in practice testing periodically throughout a lecture, or *interpolated testing* (e.g., Chan et al., 2025; Jing et al., 2016). For example, Szpunar et al. (2013) showed that answering recall questions after each of four segments of a 21-min statistics lecture reduced students' self-reported mind wandering. In a similar fashion, the present study investigated the effects of an analogous technique, *interpolated prequestioning*, on learners' thought patterns and learning during video lectures.

Prequestioning is a learning technique wherein learners attempt practice questions about material not yet learned (prequestions). They then have the opportunity to learn the correct answers (and in the case of interpolated prequestioning, learners alternate between prequestions and opportunities to learn the correct answers). For example, before a lecture about Saturn, a student might attempt the prequestion, "How many moons does Saturn have?", then learn the correct answer ("274 moons") in a subsequent lecture. Compared with conditions lacking prequestions, doing so can enhance learning—as reflected in improved performance on subsequent criterial tests—for such materials as text passages (e.g., Richland et al., 2009), short videos (e.g., Carpenter & Toftness, 2017), and lectures (e.g., Sana & Carpenter, 2023). This *prequestioning effect* (Pan & Carpenter, 2023) can be specific (limited to prequestioned information) or general (spreading to other content). A recent meta-analysis of 76 empirical studies found that the specific effects of prequestioning (Hedges's $g =$ .54), but not the general effects ($g = .04$), were statistically significant (St. Hilaire et al., 2024).

Prequestioning differs from other instructional approaches that sometimes use practice questions. For instance, with self-explanation, learners commonly explain their thought processes while learning new material (Dunlosky et al., 2013), whereas with elaborative interrogation, learners answer "Why?"-type questions about what they are learning. Retrieval practice, meanwhile, involves using practice questions to recall information after formal learning has occurred (e.g., Imundo et al., 2025). Unlike all these approaches, prequestioning occurs before formal learning, thus requiring learners to rely entirely on prior knowledge and intuition to generate guesses. Such guessing also distinguishes prequestioning from relevance instructions, a class of instructional approaches that inform learners about the relative importance of different aspects of to-be-learned texts via reading goals, elaborative interrogation, or presenting questions at various points in the learning process (McCrudden & Schraw, 2007), as well as many types of advance organizers, which highlight aspects of to-be-learned text using organizational cues such as bold headings, summaries, and other visual displays (Lorch, 1989; Lorch et al., 1995; McCrudden & Rapp, 2017). Some prequestioning techniques do, however, overlap with relevance instructions that involve answering prereading questions (e.g., McCrudden et al., 2005) and studies in the adjunct questions literature that expose learners to practice questions before sections of text (e.g., Rothkopf & Billington, 1979; for reviews, see Anderson & Biddle, 1975; Hamaker, 1986). Relatively few such studies, however, have addressed learning from videos, presented practice questions exclusively before formal instruction, or employed the same experimental controls and analytic approaches as in modern prequestioning research (for discussion, see Pan & Carpenter, 2023). With respect to video-based learning, content is often presented at a predetermined speed that may not correspond to the learner's intended pace, and the capacity to glance between different sections of content and selectively revisit information is more limited. Given these characteristics, the capacity of prequestioning to increase attentiveness may be uniquely effective in video contexts, potentially producing benefits beyond those of other instructional techniques (Carpenter & Toftness, 2017).

Prior Research on Prequestioning and Thought Patterns During Lectures

To date, research on prequestioning and attentional focus during lectures has been limited. Pan, Sana, Schmitt, et al. (2020) conducted two experiments where undergraduate students watched a 26-min video on signal detection theory, divided into four

served as lead for data curation and investigation and served in a supporting role for formal analysis, methodology, project administration, validation, writing–original draft, and writing–review and editing. Luke Chiang served in a supporting role for data curation, investigation, methodology, and writing–review and editing. Alyssa Indrajaya served as lead for formal analysis and served in a supporting role for investigation, validation, and writing–review and editing. Michelle E. Kaku served in a supporting role for data curation, investigation, methodology, supervision, and writing–review and editing.

The data are available at https://osf.io/3ka5g/.
The experiment materials are available at https://osf.io/x495e/.
The preregistered design is accessible at https://aspredicted.org/cs3s-d27c.pdf (Experiment 1), https://aspredicted.org/nbd2-npzf.pdf (Experiment 2), https://aspredicted.org/r72z-6ky8.pdf (Experiment 3).

Correspondence concerning this article should be addressed to Steven C. Pan, Department of Psychology, National University of Singapore, 9 Arts Link, Singapore City 117572, Singapore. Email: scp@nus.edu.sg

segments. Before each segment, the students either answered four multiple-choice prequestions (interpolated prequestioning) or solved eight algebra problems. After each segment, they rated their attention retrospectively on a 0–100 scale (Weinstein, 2018), and a criterial test occurred after the lecture. The authors hypothesized that interpolated prequestioning would increase self-reported attention across the lecture, as was observed (Cohen's $d = 0.39$ and 0.74 in Experiments 1 and 2). Prequestioning also yielded specific effects on criterial test performance in Experiment 1 and general effects in Experiment 2.

Although the results of Pan, Sana, Schmitt, et al. (2020) seem promising, several limitations (most acknowledged by the authors) complicate interpretation. First, as in prior work (e.g., Szpunar et al., 2013), the researchers compared prequestioning to a control activity unrelated to the lecture, leaving unclear whether it enhanced attention, the control activity reduced it, or both. Second, other researchers have criticized retrospective numerical probes as vulnerable to memory or aggregation errors and response biases (Kane et al., 2021). Third, the probes did not capture the specific nature of participants' thoughts, whether about the lecture or other content (cf. Kane et al., 2017), leaving unaddressed the extent to which prequestioning influences the distribution of different thought types during lectures.

A study by Welhaf et al. (2022) raises further questions. In a single experiment, undergraduate students watched a 52-min statistics lecture split into 31 brief segments. Beforehand, the students answered 10 multiple-choice prequestions either drawn from lecture content ("matched") or about general statistics ("mismatched"). After every five to seven segments, they also attempted six recall questions or studied those questions and answers together (interpolated retrieval practice or interpolated studying). A total of 15 categorical thought probes were administered between segments, and a criterial test followed the entire lecture. There were no significant differences between matched and mismatched conditions on task-unrelated thoughts, other thought types, or criterial test performance.

Several limitations, however, qualify these null results. First, no benefits of prequestioning or retrieval practice occurred on the criterial test. Although the authors attributed this null effect to participants not taking notes, numerous studies have shown retrieval practice and prequestioning effects without note-taking (Chan et al., 2018; Rowland, 2014; St. Hilaire et al., 2024). Second, the researchers combined prequestioning with retrieval practice or studying questions and answers, both of which can affect learning (e.g., Kang et al., 2007; see also Gupta et al., 2024); potentially contaminating any prequestioning effects. Third, the lack of a non-prequestioning control—a prerequisite in most practice testing research (see Carrier & Pashler, 1992; Kornell et al., 2012; Rowland, 2014; and others) limits conclusions, as even those answering general statistics questions engaged in prequestioning. Fourth, the researchers did not interpolate the prequestions, raising the possibility that their effects faded early. Although Pan, Sana, Schmitt, et al. (2020) found similar effects for interpolated and noninterpolated prequestions, it remains unclear whether a single, noninterpolated "dose" of prequestions at the start of a lecture is effective for sessions lasting approximately an hour.

Given these mixed results, further research on prequestioning and thought patterns during lectures is clearly needed. Such research arguably needs to (a) compare prequestioning versus an appropriate non-prequestioning control (and potentially one that is more competitive with prequestioning than in prior work), (b) use well-validated thought probes (e.g., the categorical experience sampling approach endorsed by Kane et al., 2021), (c) avoid contaminating effects of other learning activities besides the lecture, and (d) consider implementation factors (e.g., interpolating prequestions). The present study addressed all of these considerations.

How Prequestioning Might Impact On-Task and Off-Task Thoughts

Prominent theoretical accounts of the prequestioning effect (for reviews, see Kornell & Vaughn, 2016; Mera et al., 2022; Metcalfe, 2017) imply that prequestioning might impact the quality and distribution of thought patterns during lectures. Each account attributes distinct cognitive mechanisms to prequestioning and can be evaluated within a common three-stage framework (Pan & Carpenter, 2023). In that framework, the prequestioning effect arises from attempting to answer a prequestion (the first stage), learning the correct answer (the second stage), and/or on a subsequent learning assessment (the third stage).

Attention-direction or *attention-focusing* theories propose that prequestioning increases attention during the second stage as learners seek answers (Little & Bjork, 2016; Pan, Sana, Schmitt, et al., 2020; St. Hilaire et al., 2024). These accounts resemble explanations for the benefits of advance organizers (McCrudden & Rapp, 2017), as well as the goal-focusing model of relevance instructions (McCrudden & Schraw, 2007). Supporting evidence includes the finding that learners better retain information encountered up to the point where a prequestion is answered, with diminished learning thereafter (Sana & Carpenter, 2023), which suggests that attention drops once answer search ceases. Moreover, eye-tracking and reading time studies show that learners focus more on prequestioned content (Lewis & Mensink, 2012; Yang et al., 2021). Crucially, the prequestioning effect depends on learners discovering correct answers (St. Hilaire & Carpenter, 2020). If attentional processes drive this effect, then prequestioning may shift thought patterns, potentially increasing the proportion of on-task thinking as learners seek out correct answers.

A related *error correction* account posits that answering a prequestion incorrectly triggers an error signal, which increases attention, cognitive control, or other mechanisms to detect and correct mistakes (Kang et al., 2011; Metcalfe, 2017; see also Seabrooke et al., 2022). Drawing on neuroscientific perspectives (e.g., Butterfield & Mangels, 2003), this account suggests that prequestioning can promote more focused, on-task thinking. Alternatively, prequestioning may heighten *curiosity* or *interest* in learning materials (Berlyne, 1966; Geller et al., 2017; Potts et al., 2019). Supporting evidence includes the finding that prequestioning-like activities increase curiosity ratings (Potts et al., 2019) and that prequestions introducing conflict or puzzles, sparking curiosity, are more effective (Bull & Dizney, 1973; cf. Hollins et al., 2023). If so, then curiosity may occupy a greater share of learners' thoughts after prequestioning, potentially limiting off-task thinking.

Other theoretical accounts offer different or no clear predictions about how prequestioning affects thought patterns during lectures. For instance, one theory suggests that mediator words generated during prequestioning-like activities aid later recall (Kornell et al., 2009; Metcalfe & Huelser, 2020), a possibility with unclear attentional consequences. Another account proposes that prequestioning activates relevant mental models (Mayer et al., 2002), which some researchers interpret as leaving attention unaffected (e.g., Motz et al., 2024). If so, then prequestioning might not influence the proportion of on- versus off-task thoughts.

Another important consideration concerns the relationship between incorrect responding, negative emotions, and perceived difficulty. Students experience frustration and dissatisfaction when they make errors (Pan, Sana, Samani, et al., 2020), which often occurs during prequestioning, and prequestioning itself can be perceived as relatively difficult (Bjork et al., 2015). Relatedly, off-task thoughts may increase when learners engage with difficult material or experience negative emotions (Mills et al., 2013; Smallwood et al., 2009; cf. Poerio et al., 2013). Hence, any increases in on-task thoughts that stem from prequestioning might be offset by corresponding increases in off-task thoughts, or more worrisomely, a net increase in off-task thinking may occur.

The Present Study

Across three experiments, we investigated the effects of interpolated prequestioning on learners' thought patterns and learning outcomes during video lectures. Responding to calls for more robust psychological research (Makel & Plucker, 2014; Nosek et al., 2022), these experiments, which were all preregistered, tested the reproducibility of findings reported by Pan, Sana, Schmitt, et al. (2020) and used the same video materials with minor modification. Moreover, beyond replication, they addressed, for the first time, all of the aforementioned conditions needed for more conclusive evidence.

To begin, we compared interpolated prequestioning with a control activity that is both relevant to the lecture content and potentially competitive: studying *learning objectives* (i.e., statements outlining what students should be expected to learn as a result of instruction; Osueke et al., 2018). Learning objectives are commonly used pedagogical tools that, according to at least one study (Yu et al., 2025; cf. Sana et al., 2020; Schweppe et al., 2025), can be competitive with prequestioning. This comparison evaluated whether prequestioning improves thought patterns beyond outperforming a weak and irrelevant control condition. In addition, to examine how interpolated prequestioning affects distinct thought types—lecture-related on-task, lecture-related off-task, comprehension-related off-task, and task-unrelated—we employed categorical thought probes reflecting current best practices in experience sampling research (Welhaf et al., 2022). Moreover, to isolate effects of prequestioning, participants engaged only with the lecture and their assigned activity, with no other learning tasks. Furthermore, we implemented prequestioning in interpolated fashion to better distribute its effects across the lecture.

We addressed several other methodological and theoretical issues. We examined lecture viewing in an in-person, supervised setting (Experiment 1) as well as in unsupervised online environments (Experiments 2 and 3). We explored variations in prequestion formats (multiple choice in Experiments 1–2 vs. short answer with embedded answer options in Experiment 3), in the portions of each lecture targeted by prequestions (Experiments 1 and 3 vs. 2), and on criterial tests assessing specific and general effects of prequestioning. In addition, to capture potential affective and metacognitive consequences of prequestioning, we measured learners' emotional states during the lecture (Experiment 1), perceptions of difficulty (Experiments 2 and 3), and the ability to discover correct answers to prequestions (St. Hilaire & Carpenter, 2020). A summary of key design features across experiments is presented in Table 1.

This study had several plausible outcomes. First, consistent with major theoretical accounts and our preregistered predictions, interpolated prequestioning might increase on-task thoughts while reducing off-task thoughts. These changes would be reflected in a greater proportion of lecture-related on-task thoughts and corresponding reductions in lecture-related off-task, comprehension-related off-task, and task-unrelated thoughts, although such effects could also occur independently. Alternatively, prequestioning might have no effect or even decrease on-task thoughts and increase off-task thoughts. Importantly, we expected that for prequestioning to meaningfully influence learning outcomes, it would need to alter thought patterns. We examined this precondition, informed by findings from Pan, Sana, Schmitt, et al. (2020) and Welhaf et al. (2022), across all three experiments.

Experiment 1

The first experiment took place in an in-person setting with undergraduate students randomly assigned to a prequestioning group or a statements group. Both groups watched a 22-min statistics lecture divided into eight segments. Before each segment, the prequestioning group answered two questions about upcoming content, whereas the statements group studied learning objectives related to the segment. Thought probes assessed participants' mental focus. After the lecture, all participants completed a 5-min delayed criterial test.

Method

Participants

One hundred and ten undergraduate students at a public university in Singapore participated in exchange for partial course credit or a payment of $7.50 (Singapore dollars). An a priori power analysis using the *Superpower* package in R (Lakens & Caldwell, 2021) determined that 50 participants per group would be necessary to detect an effect size of Cohen's $d = 0.74$, the reported advantage of the interpolated [prequestioning] group over the control group on attention probes in Pan, Sana, Schmitt, et al. (2020), Experiment 2, via an independent-samples t test with 80% power and an α level of .05. For the present purposes, we were interested in detecting a potential increase in task-related thoughts or a

Table 1
Experimental Design Summary

Experiment	Setting and sample	Prequestion format	Prequestioned content
1	In-person, undergraduate students	Multiple choice	Beginning and end of each video segment
2	Online, adult learners	Multiple choice	End of each video segment
3	Online, adult learners	Short answer with embedded options	Beginning and end of each video segment

decrease in task-unrelated thoughts in the interpolated prequestioning group versus the interpolated statements group.

Considering the possibility of smaller effects relative to a more competitive control condition, we preregistered a larger target of 100 participants per group. To that end, we posted numerous participation slots over two semesters (February–October 2023) through the university's psychology pool and paid recruitment portals. Students with no prior familiarity with signal detection theory were eligible to participate. Due to limited sign-ups and logistical constraints, however, we did not reach the preregistered number by the end of the academic term. Accordingly, to enable more efficient data collection, for Experiments 2 and 3, we switched to online and nonuniversity samples (ultimately, across all three experiments, we obtained a combined sample size of over 150 participants per group).

Ethics approval for the entire study preceded data collection. All participants were treated according to the Declaration of Helsinki, with informed consent given beforehand. We excluded seven participants who failed instruction comprehension checks; the final sample included 103 participants (*interpolated prequestioning* group, $n = 52$, *interpolated statements* group, $n = 51$), averaging 20.9 years old ($SD = 3.1$) and 72% female. About 78% were of Chinese ancestry, 8% Indian, 2% Malay, and 13% other ethnicities.

Design

Experiment 1 employed a 2 (Training Group: Interpolated Prequestioning vs. Interpolated Statements; between-subjects) × 2 (Question Type: Tested vs. New; within-subjects) mixed design.

Transparency and Openness

This article reports how the sample size was determined, all data exclusions, all manipulations, and all measures for each experiment (Simmons et al., 2012). This study follows JARS (Appelbaum et al., 2018). All data, analysis code, and research materials are available at https://osf.io/ph2em/. Design and analysis plans for Experiment 1 were preregistered at https://aspredicted.org/cs3s-d27c.pdf.

Materials

The materials consisted of a prerecorded video lecture, 28 questions used as prequestions and/or criterial test questions, eight sets of statements, a thought probe, and two supplementary probes (emotion and metacognitive). The materials further included a series of instructions and clarifications.

Video Lecture. Participants watched a video lecture on signal detection theory previously used in Pan, Sana, Schmitt, et al. (2020), Sana and Carpenter (2023), and Toftness et al. (2018). The video was edited into eight consecutive segments (enabling the inclusion of more frequent interpolated activities (with fewer prequestions per segment) than Pan, Sana, Schmitt, et al. (2020), and due to the need to ensure smooth transitions, had an overall 22-min length that was approximately 4 min shorter than the equivalent versions used in prior research. Originally from a psychology course, the lecture featured slides with white backgrounds, black text, images, and an instructor's voiceover (audio only). It introduced signal detection theory, covering its history, key concepts (e.g., "criterion level," "*d*′"), real-life applications, and key procedures. The lecture was divided into eight segments lasting 2–3.5 min each and edited to ensure smooth transitions. Hosted on YouTube and embedded in the experiment, it was displayed centrally in a 16:9 frame, occupying about three quarters of the browser screen.

Prequestions and Statements. Between two and four multiple-choice questions were developed for each video segment. Each question targeted a fact, concept, or procedure drawn from the lecture and resembled questions previously used by Toftness et al. (2018) and Pan, Sana, Schmitt, et al. (2020). All questions consisted of a question stem (e.g., "According to signal detection theory, which of the following factors can influence decisions?") and five answer options, of which four options—for example, (a) *cognitive biases,* (b) *past experiences,* (c) *sensitivity,* (d) *intelligence*—were designed to be competitive with one another (Little et al., 2012). The fifth option, "I don't feel comfortable guessing", was included to help assuage potential participant anxiety about incorrect responding and was reminiscent of nonresponse options that have been repeatedly used in other survey research (e.g., "I don't want to answer"; Montagni et al., 2019). Eight questions included visuals (e.g., tables, figures) that participants had to consult to choose their answer.

The interpolated prequestioning group received exactly two prequestions per segment—one targeting content from the first half and one from the end—totaling 16 questions. Of these questions, 12 appeared again on the criterial test as *tested questions*, alongside 12 novel items labeled *new questions*, yielding 24 criterial test questions. Both groups received the same criterial test; for analysis consistency, the *tested* versus *new* question distinction was applied to both groups. In the interpolated statements group, however, as participants did not receive any questions during the video segments, none of the test questions had been seen beforehand. Nevertheless, as described next, many of the questions aligned, at least in part, with content targeted by the learning objectives.

One set of statements was developed for each video segment, for a total of eight sets, to be presented to the interpolated statements group. Each set consisted of two to three sentences that together presented a summary of the content and one or more learning objectives from the corresponding video segment (e.g., "In the next segment of the lecture, you will learn about how to measure discriminability. One can determine whether the ability to discriminate is high or low from a graph. You will also learn about ROC curves."). Thus, although the sets of statements did not exactly parallel the content covered by the prequestions, they previewed what participants would see in the video segment, highlighted specific outcomes that were achievable upon viewing that segment, and referenced content that would be assessed on the criterial test. For example, the learning objective mentioning ROC curves aligned with the criterial test question, "What is an ROC curve a plot of?" Together, each set of learning objectives targeted content from early and later portions of each respective segment.

At least three studies to date have investigated prequestions alongside learning objectives. Sana et al. (2020; Experiments 1–2) found that undergraduate students who studied learning objectives before reading neuroscience text passages exhibited better test performance relative to students who did not study the objectives, although multiple-choice prequestions yielded even higher performance. Yu et al. (2025; Experiment 2) found that preschool children exposed to learning objective statements before listening to science texts performed just as well as when they completed multiple-choice prequestions, but only for test questions that were not identical to the prequestions. Finally, Schweppe et al. (2025; Experiments 1–2) found that reading prequestions produced modest benefits over reading learning objectives for university students studying from a psychology textbook chapter. Together, these studies suggest that studying learning objective statements can promote learning and, in some cases, may be comparable to the benefits of prequestions.

Thought Probes. To assess participants' mental state and attention during the lecture, we employed categorical thought probes validated for construct validity, reliability, and accuracy (Kane et al., 2021). These probes appeared immediately after each odd-numbered segment (1, 3, 5, 7), asking participants to report their thoughts at that exact moment, for a total of four probes per lecture. This timing enabled consistent sampling throughout the session. We chose the probe format specifically to address methodological concerns raised about prior work.

The probes appeared in black colored text on a white background and began with a question stating, "What were you just thinking about?", just as in Kane et al. (2017), followed by seven categorical response options. We used the same response options as in Kane et al., with simplified wording for ease of understanding for participants from different cultural backgrounds. Participants chose the best fitting option by clicking a button next to the number for that option. The options were as follows:

1. Specifically, what was being discussed in the video at that time
2. The lecture topic, but not what was being presented in the video at the time
3. How well I'm understanding the video
4. Normal everyday thoughts about life
5. Daydreams
6. My current state of being
7. Other

Each probe presentation also included a note stating, "This is not an attention check, there is not necessarily a correct or incorrect answer to this question. We ask that you answer honestly and share your actual experience. We are interested in how people really think and feel." That note corresponded to instructional details presented at the outset of the experiment and addressed a methodological limitation of this probe type (see the Study Instructions section for details).

Supplementary Probes. For additional insights into mental states during the lecture, an emotion rating probe appeared after each even-numbered segment (four in total), asking participants to select how they felt at that moment from 12 options (e.g., "relaxed" to "unhappy"). At the end of the lecture, we also administered a metacognitive probe to the interpolated prequestioning group, asking participants to estimate the percentage of prequestions for which they had discovered the correct answers (St. Hilaire & Carpenter, 2020). We displayed both probe types similarly to the thought probes.

Study Instructions. At the outset of the experiment, we informed participants that they would watch a video lecture on signal detection theory and receive periodic probes. As in Kane et al. (2017), instructions emphasized that these probes would ask about their thoughts or feelings, that responses should be truthful, and that no answers were right or wrong. To address a methodological issue that, to the best of our knowledge, has not been reported in the literature and emerged during online pilot testing—namely that participants defaulted to the first option and reported near-perfect attention—we explicitly clarified that the probes were not attention checks. This issue is revisited in the General Discussion section.

In the interpolated prequestioning group, we informed participants that they would be shown multiple-choice questions before each video segment. They were "not expected to know the answers [and would] have to make a guess," and it was stated that incorrect answers would not be penalized. In addition, based on St. Hilaire and Carpenter's (2020) finding that the prequestioning effect depends critically on answer discovery and procedures used in prior prequestioning research, participants were told that each "video segment contains the answers to the questions, and you should discover them while watching the video." The same instructions were presented with each set of prequestions.

In the interpolated statements group, we informed participants that they would be shown a short text about each video segment before its presentation, that they were to read it carefully, and they should "do your best to understand it (even if you do not know much about the topic)." Each set of learning objectives contained implied instructions calling participants' attention to material that they would be expected to learn in the upcoming video segment (e.g., "In the next segment of the lecture, you will learn . . ."). We presented the same instructions with each set of statements.

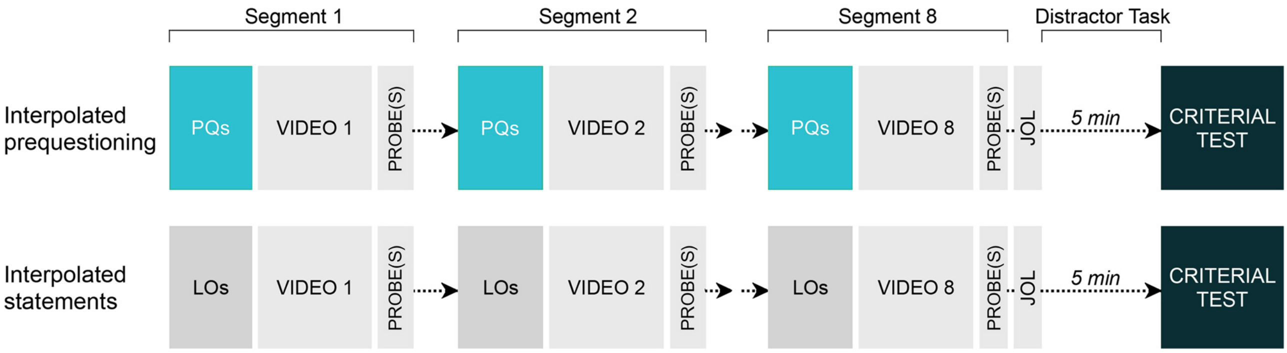
Procedure

Participants completed the 45-min experiment in a psychology lab with privacy-walled cubicles. Each cubicle contained either a Lenovo ThinkCentre M70t desktop computer with a 24-in. LCD monitor or a Dell OptiPlex 7040 desktop computer with a 20-in. monitor, along with over-the-ear wired headphones. The study was programmed using the open-source, PHP-based platform *Collector* (Garcia, 2015) and presented via Google Chrome. Up to four participants took part at a time, with an experimenter present and monitoring throughout.

All participants first gave informed consent, answered demographic questions, and put on headphones. The experimenter reminded them to follow instructions carefully and not take notes. During initial instructions, the experimenter monitored participants and answered any questions. To confirm understanding, participants completed multiple-choice comprehension checks (e.g., for the interpolated prequestioning group: "How can I find out the correct answers to the questions that will be asked?"; correct answer: "in the video lecture to follow"). As shown in Figure 1, participants then proceeded through the learning phase, a 5-min distractor task, and the criterial test.

Learning Phase. This phase included eight segments, each comprising a prevideo activity, a portion of the video lecture, and a postvideo probe (see Figure 1), all displayed in the browser. In the interpolated prequestioning group, the prevideo activity involved answering two prequestions about the upcoming segment, one question at a time. Each prequestion was shown for at least 5 s before a response was allowed. After answering, participants clicked to proceed to the next prequestion or, if both were completed, to the video. Prequestions were presented in chronological order based on when their answers appeared in the segment (first half vs. second half), with no feedback provided. In the interpolated statements group, the prevideo activity involved reading a

Figure 1
Experimental Design



Note. Participants watched an eight-segment lecture with a thought or supplementary probe after each segment. The *interpolated prequestioning* group answered prequestions (PQs) before each segment; the *interpolated statements* group studied learning objectives (LOs). After the final segment, participants completed a metacognitive probe (JOL), a 5-min distractor task, and a criterial test. In Experiment 1, only the interpolated prequestioning group received the metacognitive probe, and in Experiments 2–3, that group also answered a different metacognitive probe (answer discovery) after each segment. Segments 3–7 are omitted for brevity. See the online article for the color version of this figure.

set of statements about the upcoming segment, presented on one screen for at least 10 s.

After each prevideo activity, participants clicked a button to proceed to the next video segment. The segment played automatically without visible controls and continued uninterrupted. Immediately after, a thought probe (odd-numbered segments) or emotion probe (even-numbered segments) was presented, with unlimited response time. After all eight segments, participants in the interpolated prequestioning group completed the metacognitive probe before the distractor task, whereas the interpolated statements group proceeded directly to the distractor task.

Distractor Task. This task involved the game Tetris. The game was displayed within the same browser window for a period of 5 min. Participants were instructed to play the game throughout that period.

Criterial Test. After the distractor task, participants completed the criterial test, in which 24 multiple-choice questions were displayed in the browser window, one question at a time in random order determined anew per participant. Unlimited time was given to respond to each question, and no feedback was provided. Once participants had completed the test, they were debriefed and dismissed.

Scoring

All questions, including prequestions and criterial test questions, were computer-scored (1 point for correct answers and 0 for incorrect ones). Following Kane et al. (2017) and Welhaf et al. (2022), thought probes were scored as the proportion of reports for each category: Response 1 (*lecture-related on-task thoughts*), Response 2 (*lecture-related off-task thoughts*), Response 3 (*comprehension-related off-task thoughts*), and Responses 4–7 (*task-unrelated thoughts*).

Analysis Plan

We first computed descriptive statistics for criterial test performance and thought probe data, followed by formal analyses and mediation analyses integrating both data types. Null hypothesis significance testing was conducted at $\alpha = .05$, with 95% confidence intervals (CIs) reported where applicable (Cumming, 2014). When null hypothesis significance testing failed to reject the null, we conducted equivalence testing using a two one-sided test (TOST) approach with an equivalence margin of $\pm$ 0.5 standard deviations, representing the range of effects considered trivially small (Lakens, 2017). Equivalence testing evaluates whether an observed effect is small enough to be considered practically negligible: Using the TOST procedure, two one-sided *t* tests assess whether the effect exceeds the lower bound or falls below the upper bound of the equivalence margin; if both tests are significant (i.e., the 90% CI is fully within the bounds), then there is evidence that the effect is practically negligible.

We preregistered the use of analysis of variance (ANOVA). The mediation and equivalence analyses were exploratory, as were additional analyses of supplementary probe data in the online supplemental materials. Subsequent experiments followed the same analytic sequence, with preregistrations that progressively specified these steps in more detail. Data were analyzed using R, Version 4.4.0 (R Core Team, 2021) and the packages afex (Singmann, 2022), DescTools (Signorell, 2021), effectsize (Ben-Shachar et al., 2024), emmeans (Lenth, 2024), lme4 (Bates et al., 2015), lsr (Navarro, 2015), PROCESS (Igartua & Hayes, 2021), and TOSTER (Caldwell, 2023).

Results

Learning Phase

Prequestion performance in the interpolated prequestioning group averaged $M = 0.35$ ($SD = 0.12$), consistent with prior findings (Pan & Carpenter, 2023; St. Hilaire et al., 2024) and confirming low prior knowledge, as expected. In the interpolated statements group, participants viewed each statement for an average of 11.9 s ($SD = 1.3$), whereas in the prequestioning group, participants took an average of 18.6 s ($SD = 7.7$) per prequestion.

Criterial Test

We performed a mixed-factors ANOVA on criterial test scores with factors of training group (interpolated prequestioning vs.

interpolated statements) and question type (tested questions vs. new questions). In that analysis, there were main effects of training group, $F(1, 101) = 17.08$, $p < .001$, $\eta_p^2 = .14$, as well as question type, $F(1, 101) = 23.38$, $p < .001$, $\eta_p^2 = .19$. The main effects correspond to the criterial test results, detailed by segment in Table 2 (upper rows) and collapsed across segments in Figure 2, in which there were indications of prequestioning effects and better overall performance on tested versus new questions. There was also a significant interaction between training group and question type, $F(1, 101) = 12.35$, $p < .001$, $\eta_p^2 = .11$.

To explore those patterns, we performed independent-samples *t* tests for tested and new questions separately. The interpolated prequestioning group outperformed the interpolated statements group significantly on tested questions ($M = 0.72$, $SD = 0.15$ vs. $M = 0.53$, $SD = 0.20$), $t(101) = 5.58$, $p < .001$, $d = 1.10$, 95% CI [0.12, 0.26]. The groups did not differ significantly in performance on new questions ($M = 0.74$, $SD = 0.17$ vs. $M = 0.68$, $SD = 0.20$), $t(101) = 1.69$, $p = .095$, $d = 0.33$, 95% CI [–0.011, 0.14]. The divergent results by question type explain the significant interaction. A TOST equivalence test on the results for new questions, however, found that the 90% CI for the mean difference [0.001, 0.12] did not fall within the equivalence margin [−0.094, 0.094], indicating that the performance difference favoring the interpolated prequestioning group cannot be considered statistically equivalent to zero.

Notably, mean test performance was at least numerically higher in the interpolated prequestioning group for tested questions involving every lecture segment (see Table 2). Overall, the magnitude of the prequestion effect on tested questions is generally consistent with, although somewhat larger than, typical findings in the prequestioning literature. The effect for new questions, which was not statistically significant but still potentially meaningful, is generally consistent with the finding that prequestioning effects for information that is not directly prequestioned is often minimal or nonexistent (St. Hilaire et al., 2024).

Thought Probes

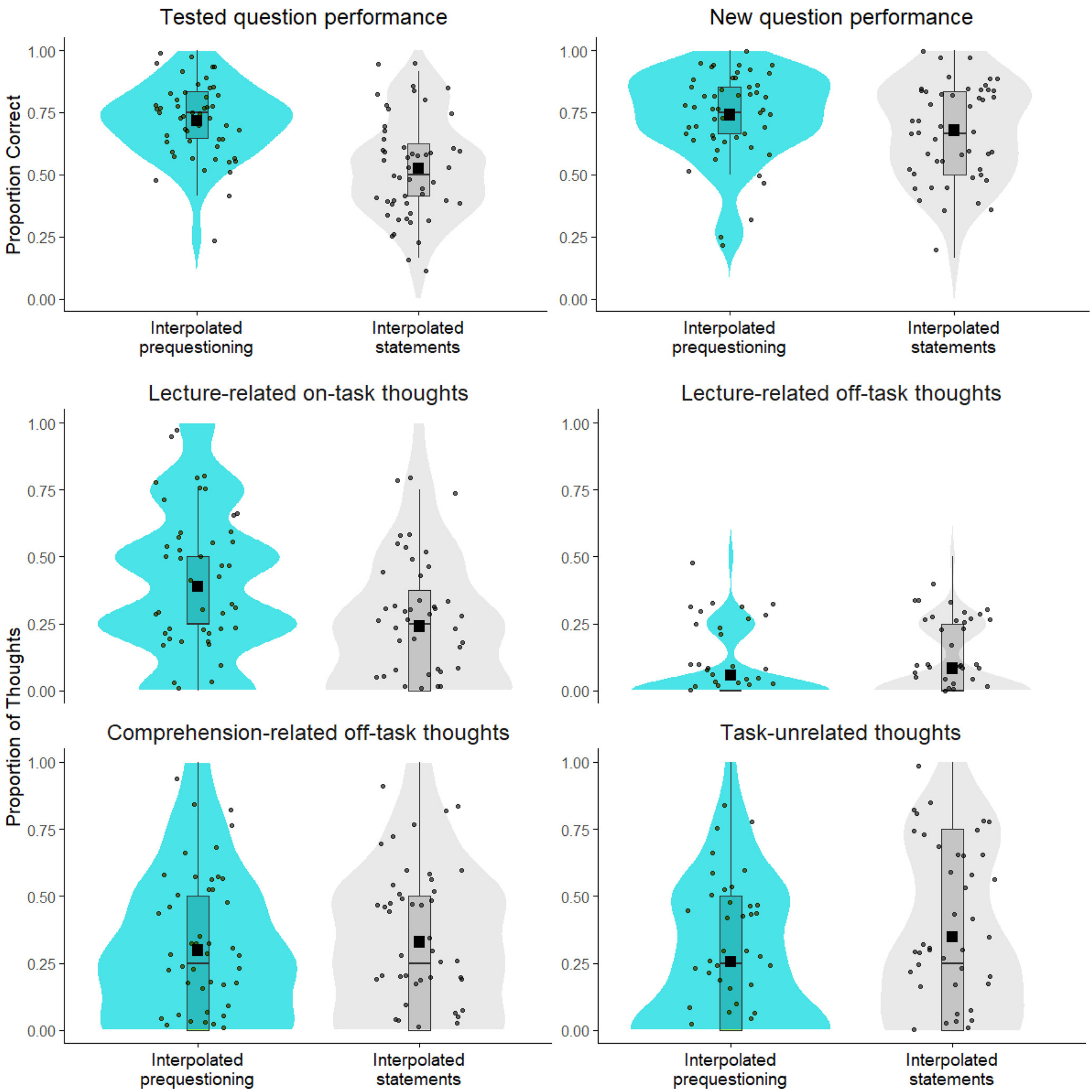
Thought probe results, separated by thought type and probe number, are presented in Table 3 (upper rows). To analyze these results, we conducted a mixed-factors ANOVA for each thought type, with training group (interpolated prequestioning vs. interpolated statements) and probe number (1 vs. 2 vs. 3 vs. 4) as factors. Probe number was included to capture possible group differences in thought patterns across the lecture duration, which would manifest as a significant interaction between training group and probe number. As detailed in Table 4, a main effect of probe number was observed in some cases, indicating differences across probes (see Table 3), but no significant interactions were found. Next, we focused on the main effect of training group, reflecting patterns in data collapsed across probes as depicted in Figure 2 (that main effect is mathematically equivalent to performing a *t* test on the data collapsed across groups, as both methods compare the same group means).

Lecture-Related On-Task Thoughts. The interpolated prequestioning group exhibited substantially higher rates of lecture-related on-task thoughts than the interpolated statements group ($M = 0.39$, $SD = 0.29$ vs. $M = 0.24$, $SD = 0.25$). That result reflects a 63% increase in on-task thoughts in the prequestioning

Table 2
Mean Proportion (SD) of Criterial Test Performance

Experiment	Question type	Group	Overall	Lecture segment 1	2	3	4	5	6	7	8
1	Tested	Prequestioning	0.72 (0.15)	0.98 (0.14)	0.67 (0.33)	0.79 (0.41)	0.89 (0.21)	0.73 (0.45)	0.68 (0.31)	0.70 (0.33)	0.19 (0.40)
		Statements	0.53 (0.20)	0.65 (0.48)	0.54 (0.40)	0.63 (0.49)	0.74 (0.38)	0.35 (0.48)	0.48 (0.35)	0.50 (0.37)	0.18 (0.39)
	New	Prequestioning	0.74 (0.17)	0.63 (0.34)	0.92 (0.27)	0.96 (0.19)	0.83 (0.28)	0.60 (0.37)	0.67 (0.47)	0.75 (0.36)	0.73 (0.45)
		Statements	0.68 (0.20)	0.56 (0.42)	0.84 (0.37)	0.92 (0.27)	0.75 (0.29)	0.61 (0.42)	0.76 (0.43)	0.56 (0.37)	0.65 (0.48)
2	Tested	Prequestioning	0.57 (0.21)	0.53 (0.50)	0.56 (0.38)	0.51 (0.37)	0.51 (0.51)	0.59 (0.50)	0.55 (0.38)	0.64 (0.37)	0.68 (0.47)
		Statements	0.51 (0.20)	0.42 (0.50)	0.58 (0.37)	0.54 (0.39)	0.40 (0.50)	0.54 (0.50)	0.47 (0.38)	0.48 (0.32)	0.58 (0.50)
	New	Prequestioning	0.53 (0.21)	0.52 (0.35)	0.51 (0.51)	0.42 (0.50)	0.61 (0.39)	0.49 (0.51)	0.56 (0.34)	0.50 (0.34)	0.62 (0.49)
		Statements	0.50 (0.17)	0.56 (0.31)	0.50 (0.51)	0.58 (0.50)	0.50 (0.30)	0.40 (0.50)	0.48 (0.38)	0.46 (0.33)	0.56 (0.50)
3	Tested	Prequestioning	0.72 (0.19)	0.85 (0.36)	0.82 (0.24)	0.73 (0.32)	0.84 (0.37)	0.70 (0.46)	0.63 (0.35)	0.55 (0.39)	0.77 (0.43)
		Statements	0.55 (0.18)	0.60 (0.49)	0.63 (0.25)	0.68 (0.36)	0.56 (0.50)	0.62 (0.49)	0.46 (0.32)	0.42 (0.36)	0.49 (0.50)
	New	Prequestioning	0.54 (0.21)	0.64 (0.35)	0.56 (0.50)	0.63 (0.49)	0.60 (0.32)	0.48 (0.50)	0.39 (0.33)	0.43 (0.34)	0.69 (0.47)
		Statements	0.54 (0.23)	0.49 (0.40)	0.47 (0.50)	0.68 (0.47)	0.65 (0.39)	0.44 (0.50)	0.47 (0.38)	0.45 (0.34)	0.77 (0.42)

Figure 2
Experiment 1 Criterial Test Performance and Thought Probe Results



Note. See the online article for the color version of this figure.

group, which was statistically significant (p = .0058) with an effect size of d = 0.56. That effect size is similar to that observed in Pan, Sana, Schmitt, et al. (2020; Experiment 2) for the case of attention paid to the same video lecture.

Lecture-Related Off-Task Thoughts. Rates of lecture-related off-task thoughts were generally low and did not significantly differ between the interpolated prequestioning and interpolated statements groups (M = 0.057, SD = 0.12 vs. M = 0.083, SD = 0.13), with an effect size of d = –0.21 and no significant effect in the ANOVA results (Table 4). A TOST equivalence test, however, showed that the 90% CI for the mean difference [–.066, .015] fell just outside the equivalence margin [–0.062, 0.062], suggesting that the numerically higher rate of lecture-related off-task thought in the interpolated statements group cannot be considered statistically equivalent to zero.

Table 3
Mean Proportion (SD) of Thought Probe Responses

Experiment	Thought type	Group	Overall	Probe 1	Probe 2	Probe 3	Probe 4
1	Lecture-related on-task	Prequestioning	0.39 (0.29)	0.42 (0.50)	0.48 (0.50)	0.29 (0.46)	0.37 (0.49)
		Statements	0.24 (0.25)	0.20 (0.40)	0.41 (0.50)	0.22 (0.42)	0.14 (0.35)
	Lecture-related off-task	Prequestioning	0.06 (0.12)	0.08 (0.27)	0.04 (0.19)	0.04 (0.19)	0.08 (0.27)
		Statements	0.08 (0.13)	0.06 (0.24)	0.02 (0.14)	0.16 (0.37)	0.10 (0.30)
	Comprehension-related off-task	Prequestioning	0.30 (0.28)	0.33 (0.47)	0.33 (0.47)	0.27 (0.45)	0.27 (0.45)
		Statements	0.33 (0.28)	0.47 (0.50)	0.41 (0.50)	0.18 (0.39)	0.26 (0.44)
	Task-unrelated	Prequestioning	0.25 (0.27)	0.17 (0.38)	0.15 (0.36)	0.40 (0.50)	0.29 (0.46)
		Statements	0.35 (0.32)	0.28 (0.45)	0.16 (0.37)	0.45 (0.50)	0.51 (0.50)
2	Lecture-related on-task	Prequestioning	0.58 (0.37)	0.64 (0.48)	0.60 (0.49)	0.47 (0.50)	0.60 (0.49)
		Statements	0.43 (0.34)	0.36 (0.49)	0.54 (0.50)	0.44 (0.50)	0.38 (0.49)
	Lecture-related off-task	Prequestioning	0.08 (0.15)	0.06 (0.23)	0.04 (0.19)	0.11 (0.32)	0.09 (0.30)
		Statements	0.07 (0.14)	0.02 (0.14)	0.06 (0.24)	0.08 (0.27)	0.10 (0.30)
	Comprehension-related off-task	Prequestioning	0.22 (0.33)	0.17 (0.38)	0.26 (0.45)	0.28 (0.46)	0.17 (0.38)
		Statements	0.33 (0.29)	0.50 (0.51)	0.28 (0.45)	0.28 (0.45)	0.26 (0.44)
	Task-unrelated	Prequestioning	0.12 (0.24)	0.13 (0.34)	0.09 (0.30)	0.13 (0.34)	0.13 (0.34)
		Statements	0.18 (0.26)	0.12 (0.33)	0.12 (0.33)	0.20 (0.40)	0.26 (0.44)
3	Lecture-related on-task	Prequestioning	0.68 (0.35)	0.77 (0.43)	0.71 (0.46)	0.56 (0.50)	0.70 (0.46)
		Statements	0.50 (0.36)	0.60 (0.49)	0.55 (0.50)	0.40 (0.49)	0.45 (0.50)
	Lecture-related off-task	Prequestioning	0.02 (0.09)	0.00 (0.00)	0.01 (0.12)	0.05 (0.23)	0.03 (0.16)
		Statements	0.06 (0.14)	0.04 (0.19)	0.06 (0.25)	0.08 (0.27)	0.08 (0.27)
	Comprehension-related off-task	Prequestioning	0.22 (0.29)	0.22 (0.42)	0.22 (0.42)	0.29 (0.46)	0.15 (0.36)
		Statements	0.31 (0.28)	0.31 (0.47)	0.28 (0.45)	0.35 (0.48)	0.30 (0.46)
	Task-unrelated	Prequestioning	0.07 (0.18)	0.01 (0.12)	0.05 (0.23)	0.10 (0.30)	0.12 (0.33)
		Statements	0.13 (0.23)	0.05 (0.22)	0.10 (0.31)	0.18 (0.39)	0.18 (0.39)

Comprehension-Related Off-Task Thoughts. Rates of comprehension-related off-task thoughts did not significantly differ between the interpolated prequestioning and interpolated statements groups (M = 0.30, SD = 0.28 vs. M = 0.33, SD = 0.28), with an effect size of d = –0.11 and no significant effect in the ANOVA results (Table 4). A TOST equivalence test showed that the 90% CI for the mean difference [–0.122, 0.061] fell within the equivalence margin [–0.14, 0.14], indicating practical equivalence.

Task-Unrelated Thoughts. Rates of task-unrelated thoughts were lower in the interpolated prequestioning group (M = 0.26, SD = 0.27 vs. M = 0.35, SD = 0.32), corresponding to an effect size

Table 4
Outcomes of Omnibus Analysis of Variance of Thought Probe Responses in Experiment 1

Analysis	df	F	p	η_p^2
Lecture-related on-task thoughts: 2 (Training Group: Interpolated Prequestioning vs. Interpolated Statements) × 4 (Probe Number: 1 vs. 2 vs. 3 vs. 4)				
Main effect of training group	1, 101	7.94	.0058**	.07
Main effect of probe number	3, 303	4.84	.0026**	.05
Interaction	3, 303	1.18	.32	.01
Lecture-related off-task thoughts: 2 (Training Group: Interpolated Prequestioning vs. Interpolated Statements) × 4 (Probe Number: 1 vs. 2 vs. 3 vs. 4)				
Main effect of training group	1, 101	1.11	.29	.01
Main effect of probe number	3, 303	1.42	.24	.01
Interaction	3, 303	1.62	.19	.02
Comprehension-related off-task thoughts: 2 (Training Group: Interpolated Prequestioning vs. Interpolated Statements) × 4 (Probe Number: 1 vs. 2 vs. 3 vs. 4)				
Main effect of training group	1, 101	0.30	.58	.003
Main effect of probe number	3, 303	4.11	.0070**	.04
Interaction	3, 303	1.60	.19	.02
Task-unrelated thoughts: 2 (Training Group: Interpolated Prequestioning vs. Interpolated Statements) × 4 (Probe Number: 1 vs. 2 vs. 3 vs. 4)				
Main effect of training group	1, 101	2.56	.11	.02
Main effect of probe number	3, 303	12.37	<.001***	.11
Interaction	3, 303	1.57	.20	.02

** p < .01. *** p < .001.

of d = –0.32, but no significant effect was found in the ANOVA results (Table 4). A TOST equivalence test, however, showed that the 90% CI for the mean difference [–0.19, 0.004] exceeded the equivalence margin [–0.15, 0.15], suggesting that the 29% reduction in task-unrelated thoughts in the interpolated prequestioning group cannot be considered statistically equivalent to zero.

Mediation Analyses

To further elucidate the effect of interpolated prequestioning on criterial test performance—and specifically whether lecture-related on-task thoughts, for which the group differences were most pronounced, acted as a mediating variable—we conducted two mediation analyses. The first analysis focused on overall criterial test performance, whereas the second analysis examined performance on tested questions only. These analyses used the *PROCESS* macro for R to estimate indirect effects by calculating 95% CIs with 5,000 bootstraps. The results are depicted in Figure 3. Additional exploratory mediation analyses involving time spent on prevideo activities are reported in the online supplemental materials.

On-Task Thoughts and Overall Criterial Test Performance. In the analysis of effects on overall criterial test performance, the total effect of interpolated prequestioning (β = 0.13, $p < .001$; 95% CI = [0.066, 0.19]) was smaller upon inclusion of the mediator, lecture-related on-task thoughts. The direct effect was significant (β = 0.095, p = .002; 95% CI [0.036, 0.15]), and the indirect effect was significant as well (β = 0.031, 95% CI [0.0080, 0.061]). These findings suggest that interpolated prequestioning encouraged and increased lecture-related on-task thoughts, which in turn improved learning from the lecture and criterial test performance.

On-Task Thoughts and Performance on Tested Questions. In the analysis of performance on tested questions, the total effect of interpolated prequestioning (β = 0.19, $p < .001$; 95% CI [0.12, 0.26]) decreased with the inclusion of the mediator, lecture-related on-task thoughts. Both the direct effect (β = 0.17, $p < .001$; 95% CI [0.10, 0.24]) and the indirect effect (β = 0.022; 95% CI [0.0011, 0.050]) were significant. These results suggest that interpolated prequestioning increased on-task thoughts, which in turn enhanced performance on tested questions.

Supplementary Probes

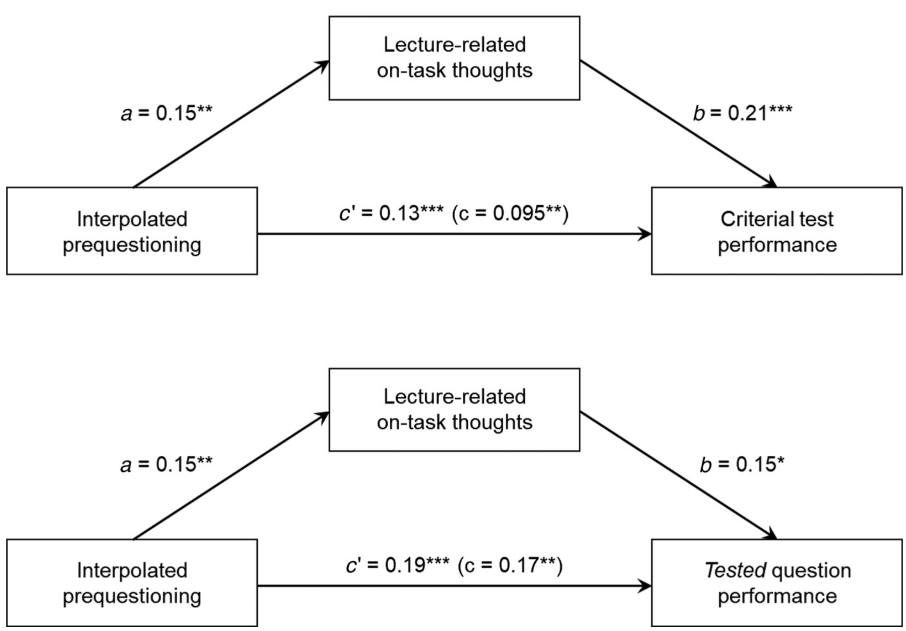
The average rate of answer discovery in the interpolated prequestioning group, as measured via the metacognitive probe administered at the end of the entire lecture, was relatively high (M = 73%, SD = 21%). That pattern suggests compliance with instructions and relative ease at finding the answers. The overall proportion of negative emotions across video segments, as measured using the emotion probes, was roughly comparable between the interpolated prequestioning (M = 0.50, SD = 0.50) and interpolated statements (M = 0.56, SD = 0.50) groups. Hence, prequestioning did not yield more negative emotions than reading statements (cf. Pan, Sana, Samani, et al., 2020). Additional supplementary analyses of the answer discovery and emotion probe data are presented in the online supplementary materials.

Discussion

Experiment 1 yielded several notable findings. First, interpolated prequestioning improved criterial test performance—specifically for prequestioned content—compared with reading interpolated

Figure 3

Experiment 1 Mediation Analyses



Note. Relationships between interpolated prequestioning, lecture-related on-task thoughts, and criterial test performance (upper panel) and tested question performance (lower panel).
* $p < .05$. ** $p < .01$. *** $p < .001$.

statements. This result reinforces prior evidence that prequestioning is more effective than other advance organizers (Sana et al., 2020) and suggests that prequestioning itself uniquely enhances learning beyond merely previewing content. Moreover, the strong prequestioning effect met the prerequisite for observing beneficial impacts on thought patterns during the lecture. Second, thought probe data showed that interpolated prequestioning significantly increased lecture-related on-task thoughts, indicating enhanced attention. This result conceptually replicates Pan, Sana, Schmitt, et al.'s (2020) findings using a different probe and with a more competitive control condition. In addition, there were indications of reductions in off-task thoughts, particularly task-unrelated thoughts, although those reductions were not statistically significant. Hence, interpolated prequestioning appears to primarily improve focused attention, with a possible secondary effect on off-task thinking. Third, lecture-related on-task thoughts mediated criterial test performance in the prequestioning group, consistent with attention-direction theories of prequestioning.

Experiment 2

The second experiment investigated whether the aforementioned patterns would replicate in an online setting with adult learners and with prequestions targeting different portions of each lecture segment. The shift in setting allowed for exploring interpolated prequestioning in remote learning conditions (Janelli & Lipnevich, 2021), which are increasingly prevalent and where learners' difficulties in maintaining attention may be even more pronounced. It featured redesigned prequestions that maintained the same format and style as in the prior experiment but targeted only materials presented in the final 30–60 s of each video segment. This approach drew inspiration from Sana and Carpenter (2023; Experiment 2), who found better learning of both tested and new content when prequestions focused on material presented toward the end, rather than the beginning, of a video lecture.

Method

The experimental design, target sample size, and analysis plan were preregistered at https://aspredicted.org/nbd2-npzf.pdf.

Participants

One hundred nineteen adults were recruited from Prolific Academic in January and February 2024, compensated \$6.67 (U.S. dollars) each. Prolific Academic is a reputable crowdsourcing platform for psychological research (Palan & Schitter, 2017). Participants were fluent English speakers from Australia, Canada, New Zealand, Singapore, the United Kingdom, or the United States, aged 21–40, with a platform approval rate of 95% or higher, and no prior training in signal detection theory. Before analysis, we excluded 15 participants for failing instruction checks, and we removed one participant for technical issues. The final sample, which met preregistered targets, included 103 participants (*interpolated prequestioning* group, $n = 53$, *interpolated statements* group, $n = 50$), averaging 31.9 years old ($SD = 5.2$) and 66% male. Ethnic composition was 6% Asian, 26% Black, 3% Hispanic/Latino, 57% White, and 8% other. Participants were 62% from the United Kingdom, 21% from the United States, and 17% from other eligible countries.

Design and Materials

Experiment 2 built on its predecessor with several key changes. The prequestions kept the style of Experiment 1 but focused solely on content from the end of each video segment, reflecting our interest in the effects of prequestions targeting information whose answers could only be learned by paying close attention in the latter portions of each segment. The learning objective statements remained the same. Three changes were made to the supplementary probes: (a) to explore potential group differences in perceived effort (Onan et al., 2022), the emotion probes were replaced with difficulty probes measuring perceived learning difficulty on a 5-point Likert scale (from *very easy* to *very difficult*); (b) the interpolated prequestioning group provided answer discovery ratings after each segment rather than a single retrospective rating, allowing more precise temporal measurement; and (c) all participants gave predictions of their future test performance after the lecture to gauge perceived learning.

Procedure

The procedure was adapted from Experiment 1 for an online setting. After signing up on Prolific Academic—with requirements to complete the study in a quiet, undisturbed location using a desktop or laptop with Google Chrome and working audio—participants were redirected to the experiment in their browser. We instructed them to mute all notifications, close other programs and tabs, avoid distractions, and refrain from taking notes or unrelated activities during the experiment. The rest of the procedure closely followed Experiment 1. Upon completion and debriefing, participants were redirected back to Prolific Academic for participation recording and payment processing.

Scoring and Analysis Plan

Scoring occurred in the same fashion as in Experiment 1. The analysis sequence followed the same approach as in the previous experiment but was documented in a more detailed preregistration.

Results

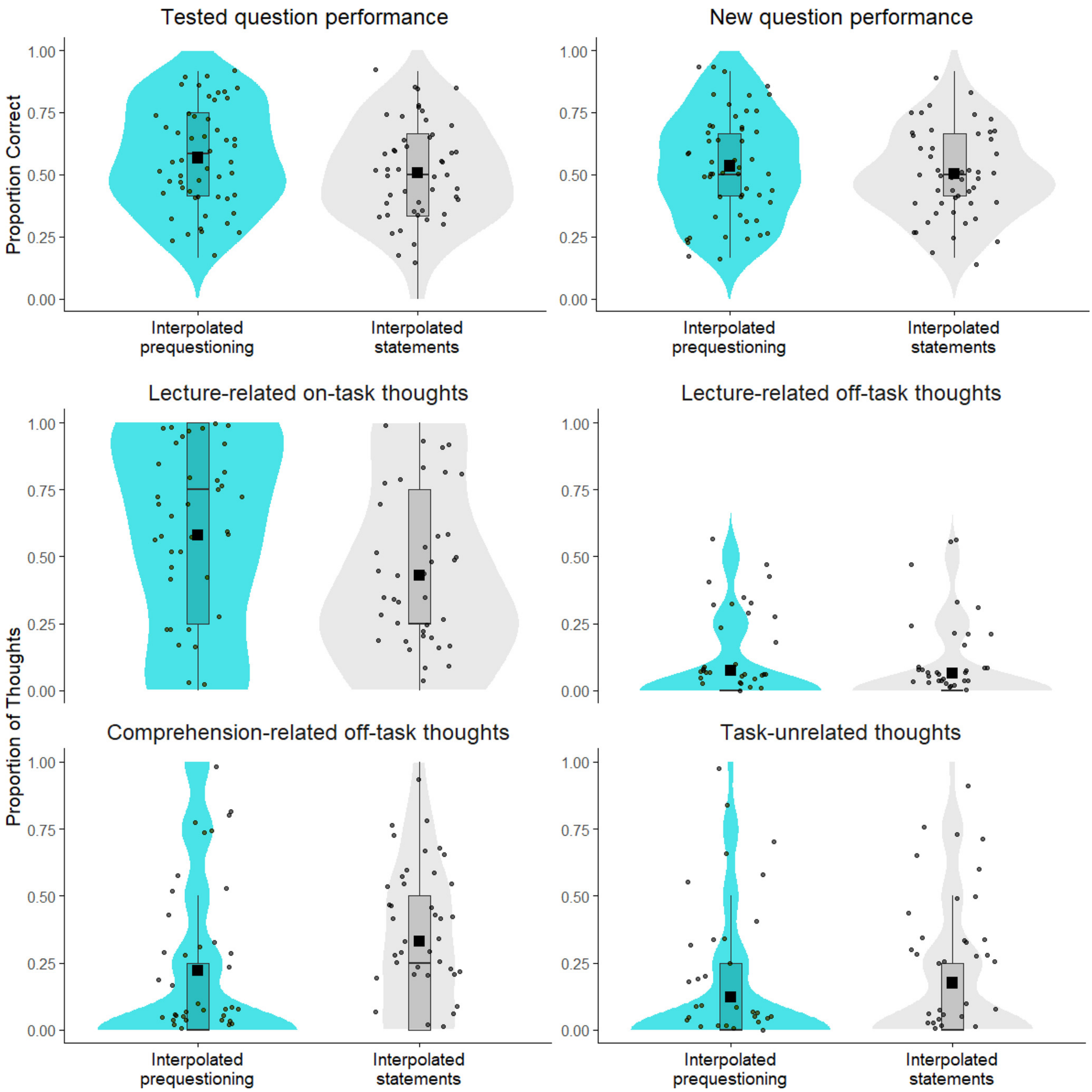
Learning Phase

Prequestion performance in the interpolated prequestioning group was $M = 0.36$ ($SD = 0.14$), which is nearly identical to that observed in Experiment 1. The average time spent viewing each statement in the interpolated statements group was 19.9 s ($SD = 15.2$), whereas the average time taken to respond to each prequestion in the interpolated prequestioning group was 25.0 s ($SD = 12.7$), which is analogous to patterns observed in Experiment 1.

Criterial Test

A mixed-factors ANOVA on criterial test scores with factors of training group and question type, analogous to that conducted for Experiment 1, was performed. There were no significant effects of training group, $F(1, 101) = 2.00$, $p = .16$, $\eta_p^2 = .02$, or question type, $F(1, 101) = 0.70$, $p = .41$, $\eta_p^2 = .0069$, and no significant interaction, $F(1, 101) = 0.47$, $p = .50$, $\eta_p^2 = .0046$. These results align with inspection of the criterial test results, which are detailed by segment in Table 2 (middle rows) and collapsed across segments in Figure 4, which show that mean scores were slightly

Figure 4
Experiment 2 Criterial Test Performance and Thought Probe Results



Note. See the online article for the color version of this figure.

higher for the prequestioning group in some segments, but these differences were not statistically significant, in contrast with the results from Experiment 1.

Despite the lack of significant ANOVA results, we conducted exploratory pairwise comparisons to maintain consistency with the analyses reported in the prior experiment. These comparisons revealed no significant differences between the interpolated prequestioning and interpolated statements groups on tested questions (M = 0.57, SD = 0.21 vs. M = 0.51, SD = 0.20), $t(101)$ = 1.51, p = .13, d = 0.30, 95% CI = [–0.019, 0.14], or new questions (M = 0.54, SD = 0.21 vs. M = 0.50, SD = 0.18), $t(101)$ = 0.82, p = .41, d = 0.16, 95% CI = [–0.044, 0.11]. A TOST equivalence test on the results for tested questions, however, found that the 90% CI for the mean difference [–0.006, 0.13] did not fall within the equivalence margin [–0.10, 0.10], whereas a corresponding test for new questions found that the 90% CI for the mean difference [−0.032, 0.094] fell within the equivalence margin [–0.096, 0.096]. Accordingly, although the difference between groups for tested questions was not statistically significant, it cannot be considered statistically equivalent to zero, whereas the performance difference for new questions was negligible.

Thought Probes

Thought probe results, separated by thought type and probe number, are presented in Table 3 (middle rows). As in Experiment 1, we conducted a mixed-factors ANOVA for each thought type, with training group and probe number as factors, and with results detailed in Table 5. There was no effect of probe number in all cases and no significant interaction for all thought types except for comprehension-related off-task thoughts. We next focused on the main effect of training group, reflecting patterns in collapsed data as depicted in Figure 4. Doing so involved pairwise comparisons, as specified in the preregistration, to provide further insights and maintain consistency with the analyses reported in the prior experiment.

Lecture-Related On-Task Thoughts. The interpolated prequestioning group exhibited significantly higher rates of lecture-related on-task thoughts than the interpolated statements group ($M = 0.58$, $SD = 0.37$ vs. $M = 0.43$, $SD = 0.34$). That difference reflects a 35% increase in on-task thoughts in the prequestioning group, which is statistically significant ($p = .035$) with an effect size of $d = 0.42$. Although this difference is directionally consistent with the findings in Experiment 1, and in mean percentage terms quite comparable, the effect size is notably smaller.

Lecture-Related Off-Task Thoughts. Rates of lecture-related off-task thought were generally low, as in Experiment 1 and did not significantly differ between the interpolated prequestioning and interpolated statements groups ($M = 0.076$, $SD = 0.15$ vs. $M = 0.065$, $SD = 0.14$), with an effect size of $d = 0.071$ and no significant effect in the ANOVA results (Table 5). A TOST equivalence test showed that the 90% CI for the mean difference [–0.037, 0.058] fell within the equivalence margin [–0.073, 0.073], indicating practical equivalence.

Comprehension-Related Off-Task Thoughts. The significant interaction between training group and probe for this thought type was the only such interaction observed in each of the three experiments. Examination of the means for each probe (see Table 3) suggests that this interaction was driven by higher rates of comprehension-related off-task thinking in the interpolated statement group for Probe 1 ($M_{\text{difference}} = 0.33$), which then decreased or disappeared in Probes 2–4 ($M_{\text{difference}} = 0.00$–$0.09$). Hence, the reduction in comprehension-related off-task thoughts in the interpolated prequestioning group was primarily confined to the start of the lecture in this experiment.

Task-Unrelated Thoughts. Rates of task-unrelated thoughts were noticeably lower than in Experiment 1 and did not significantly differ between the interpolated prequestioning and interpolated statements groups ($M = 0.12$, $SD = 0.24$ vs. $M = 0.18$, $SD = 0.26$), with an effect size of $d = –0.21$ and no significant effect in the ANOVA results (see Table 5). A TOST equivalence test, however, showed that the 90% CI for the mean difference [–0.14, 0.031] fell just outside the equivalence margin [–0.13, 0.13], suggesting that the observed 33% reduction in comprehension-related off-task thoughts in the interpolated prequestioning group cannot be considered statistically equivalent to zero.

Mediation Analyses

In line with our preregistered analysis plan, we did not perform mediation analyses for this experiment given that statistically significant prequestioning effects were not observed.

Supplementary Probes

The overall answer discovery rate in the interpolated prequestioning group ($M = 72\%$, $SD = 19\%$) closely matched that of Experiment 1. Difficulty ratings were similar between the interpolated

Table 5
Outcomes of Omnibus Analysis of Variance of Thought Probe Responses in Experiment 2

Analysis	df	F	p	η_p^2
Lecture-related on-task thoughts: 2 (Training Group: Interpolated Prequestioning vs. Interpolated Statements) × 4 (Probe Number: 1 vs. 2 vs. 3 vs. 4)				
Main effect of training group	1, 101	4.56	.035*	.04
Main effect of probe number	3, 303	1.55	.20	.02
Interaction	3, 303	2.42	.066	.02
Lecture-related off-task thoughts: 2 (Training Group: Interpolated Prequestioning vs. Interpolated Statements) × 4 (Probe Number: 1 vs. 2 vs. 3 vs. 4)				
Main effect of training group	1, 101	0.13	.73	.0013
Main effect of probe number	3, 303	1.68	.17	.02
Interaction	3, 303	0.37	.78	.0036
Comprehension-related off-task thoughts: 2 (Training Group: Interpolated Prequestioning vs. Interpolated Statements) × 4 (Probe Number: 1 vs. 2 vs. 3 vs. 4)				
Main effect of training group	1, 101	3.12	.080	.03
Main effect of probe number	3, 303	1.92	.13	.02
Interaction	3, 303	4.67	.0033**	.04
Task-unrelated thoughts: 2 (Training Group: Interpolated Prequestioning vs. Interpolated Statements) × 4 (Probe Number: 1 vs. 2 vs. 3 vs. 4)				
Main effect of training group	1, 101	1.10	.30	.01
Main effect of probe number	3, 303	1.99	.12	.02
Interaction	3, 303	1.12	.34	.01

* $p < .05$. ** $p < .01$.

prequestioning (M = 3.12, SD = 1.16) and interpolated statements (M = 3.34, SD = 1.17) groups (on a 1–5 scale). The interpolated prequestioning group predicted higher future test performance than the interpolated statements group (M = 60%, SD = 23% vs. M = 50%, SD = 24%), $t(101) = 2.24$, $p = .027$, $d = 0.44$. When examined separately by group, the prequestioning group's predictions were significantly correlated with actual performance, adjusted $R^2 = .275$, $F(1, 51) = 20.72$, $p < .001$, $\eta_p^2 = .29$, as were the predictions from the learning objectives group, adjusted $R^2 = .099$, $F(1, 48) = 6.38$, $p = .015$, $\eta_p^2 = .12$. Overall, these results suggest comparable answer discovery success across experiments, no greater perceived difficulty for prequestioning, and higher anticipated test outcomes for that group.

Discussion

Experiment 2 demonstrated that in an online setting, prequestions targeting content at the end of short lecture segments yielded, at best, modest learning benefits for tested content compared with reading interpolated statements of learning objectives. We observed no benefits for performance on criterial test questions targeting new content. Effects on thought patterns followed somewhat similar trends as in Experiment 1, with statistically significant increases in on-task thoughts. These findings are consistent with the observation that when a robust prequestioning effect is absent, corresponding benefits on thought patterns during lectures are also reduced. Nonetheless, a pattern resembling that of Experiment 1—where prequestioning appeared to improve focused attention—was still observed.

Several factors may explain the weak prequestioning effects. Positioning prequestioned content at the end of video segments might have impeded answer discovery due to proactive interference or increased difficulty. The need to remember prequestions to find answers may have been less effective in the online format, and participants who believed they had found the correct answers may not have learned them as effectively as in Experiment 1. The unsupervised environment may have introduced distractions, and characteristics unique to the online sample may have influenced responses. Ultimately, more effective prequestioning implementations may be necessary to produce larger learning gains and clearer impacts on thought patterns.

Experiment 3

The third experiment again investigated whether the patterns observed in the first experiment would conceptually replicate in an online setting. Unlike its immediate predecessor, however, it featured prequestions resembling those used in Experiment 1—that is, addressing content present at the beginning *and* later portions of each lecture segment, as opposed to only the end of each segment as in Experiment 2. The prequestion format also changed to short answer with embedded answer options instead of multiple-choice—a modification intended to improve the efficacy of prequestioning in an online environment—and study instructions and procedures were further updated. All of these changes aimed to address potential shortcomings of Experiment 2.

Method

The experimental design, target sample size, and analysis plan were preregistered at https://aspredicted.org/r72z-6ky8.pdf.

Participants

The target sample size was increased based on an a priori power analysis using the Superpower package in R. That analysis indicated that 70 participants per group were needed to achieve 80% power to detect an effect size of Cohen's $d = 0.56$ (reflecting the increase in on-task lecture-related thoughts in the interpolated prequestioning group versus the interpolated statements group in Experiment 1) via an independent-samples t test with 80% power and $\alpha = .05$. Accordingly, we recruited 173 adult participants from Prolific Academic in December 2024, following the same procedures and eligibility criteria as Experiment 2. Before formal analysis, we removed 22 participants for incomplete responding and one participant due to technical problems on the criterial test. The final sample included 151 participants (*interpolated prequestioning* group, $n = 73$; *interpolated statements* group, $n = 78$), with a mean age of 31.0 years (SD = 5.0) and 62% male. Approximately 9% identified as Asian, 33% Black, 1% Hispanic or Latino, 50% White, and 6% from other ethnic groups. Forty-eight percent were from the United Kingdom, 44% from the United States, and 8% from other eligible countries.

Design and Materials

Experiment 3 retained most of the design from Experiment 2 but introduced modifications to the prequestions and the initial procedures. Most notably, the prequestioned content reverted to the approach used in Experiment 1, with each lecture segment including one question from the first half and another from the latter half. In addition, the prequestions shifted from a multiple-choice format to a form of short answer with embedded answer options, requiring participants to compose and type their responses rather than merely select from predetermined options. This change was expected to better facilitate a prequestioning effect in an online environment: Requiring typed responses eliminated the possibility of haphazardly clicking through multiple-choice prequestions, which inattentive or distracted online participants might engage in. Each prequestion included a textbox for typed responses, followed by a submit button to proceed. Due to the logistics of this format, prequestions involving interpretation of figures or tables from Experiment 1 were replaced.

Although the prequestions required typed responses, we embedded at least two answer options within each question stem (e.g., "According to Signal Detection Theory, can past experiences or sensitivity influence decisions?"), with only one correct answer. Two questions featured three answer options. This design aimed to keep questions manageable in difficulty (cf. Hausman & Rhodes, 2018) and reduce response time (although we still expected a greater time-on-task disparity than in the prior experiments), whereas typing complete sentences was intended to encourage engagement and allow identification of disengaged participants. The criterial test retained the self-paced multiple-choice format used previously. On the criterial test, tested questions featured the same question stems as their prequestion equivalents, except that the answer options were not embedded within them. Rather, the options were listed separately, for example, "According to signal detection theory, which of the following factors can influence decisions? (a) cognitive biases, (b) past experiences, (c) sensitivity, (d) intelligence."

Both groups also received enhanced instructions at the start of the experiment to better clarify procedures. We modified the instructions for prequestioning to emphasize the requirement for sentence-length typed responses. We also expanded comprehension checks in accordance with Prolific Academic's updated policies, including increased multiple-choice options per question to minimize lucky guesses (an issue that may have occurred with the simpler checks used in the prior experiment).

Procedure, Scoring, and Analysis Plan

The procedure, scoring, and analysis plan were identical to Experiment 2 with one exception: Prequestion performance (i.e., guessing) was hand scored by a research assistant using an answer key, wherein answers that very closely matched the correct answer were counted as correct. To establish the reliability of this scoring procedure, a second research assistant scored 13.33% of the responses, yielding a high rate of interrater reliability (Cohen's $\kappa = .987$, 99.4% overall agreement). The analysis procedures were also documented in even greater detail in the preregistration for this experiment.

Results

Learning Phase

Performance in the interpolated prequestioning group on short-answer prequestions was $M = 0.62$ ($SD = 0.18$). Although very high by prequestioning literature standards, this rate likely overrepresents participants' true knowledge levels, as chance accuracy was approximately 0.50 (also uniquely high) due to the presence of answer options in each question stem. The interpolated statements group spent an average of 19.8 s ($SD = 9.9$) viewing each statement, whereas the interpolated prequestioning group spent significantly longer responding to each prequestion ($M = 123.1$ s, $SD = 74.0$). It should be emphasized, however, that a time-on-task disparity of such magnitude was unique to this experiment.

Criterial Test

A mixed-factors ANOVA on criterial test scores with factors of training group and question type, analogous to analyses conducted for the prior experiments, revealed a significant main effect of training group, $F(1, 149) = 7.68$, $p = .0063$, $\eta_p^2 = .05$, and a significant main effect of question type, $F(1, 149) = 41.43$, $p < .001$, $\eta_p^2 = .22$. These results correspond to inspection of the criterial test data, which are detailed by segment in Table 2 (bottom rows) and collapsed across segments in Figure 5, in which there are indications of a prequestioning effect and better performance on tested questions. There was also a significant interaction between training group and question type, $F(1, 149) = 30.43$, $p < .001$, $\eta_p^2 = .17$.

Pairwise comparisons revealed significant differences between the interpolated prequestioning and interpolated statements groups on tested questions ($M = 0.72$, $SD = 0.19$ vs. $M = 0.55$, $SD = 0.19$), $t(149) = 5.47$, $p < .001$, $d = 0.89$, 95% CI [0.11, 0.23], but not for new questions ($M = 0.54$, $SD = 0.21$ vs. $M = 0.54$, $SD = 0.23$), $t(149) = -0.020$, $p = .98$, $d = 0.003$, 95% CI [−0.072, 0.071], which accounts for the significant interaction. A TOST equivalence test conducted on the new question results revealed that the 90% CI for the mean difference [0.060, 0.059] fell within the equivalence margin [−0.11, 0.11], indicating practical equivalence.

Overall, the criterial test results of Experiment 3 more closely resemble those of Experiment 1 rather than Experiment 2. There was a robust prequestioning effect for tested questions and no significant effect for new questions, just as in Experiment 1. Similar to that experiment, better tested question performance was observed for all lecture segments in the interpolated prequestioning group. In this case, however, performance on new questions was practically equivalent between groups, which aligns even more closely with typical results in the prequestioning literature (St. Hilaire et al., 2024). Crucially, the robust prequestioning effect for tested questions met the precondition outlined at the outset of this article.

Thought Probes

Thought probe results, separated by thought type and probe number, are presented in Table 3 (bottom rows). As in the prior experiments, we conducted a mixed-factors ANOVA for each thought type, with training group and probe number as factors, with results detailed in Table 6. As in Experiment 1, a main effect of probe number was observed in some cases, indicating differences across probes (see Table 3), and no significant interactions were found. We next focused on the main effect of training group, reflecting patterns in data collapsed across probes as depicted in Figure 5.

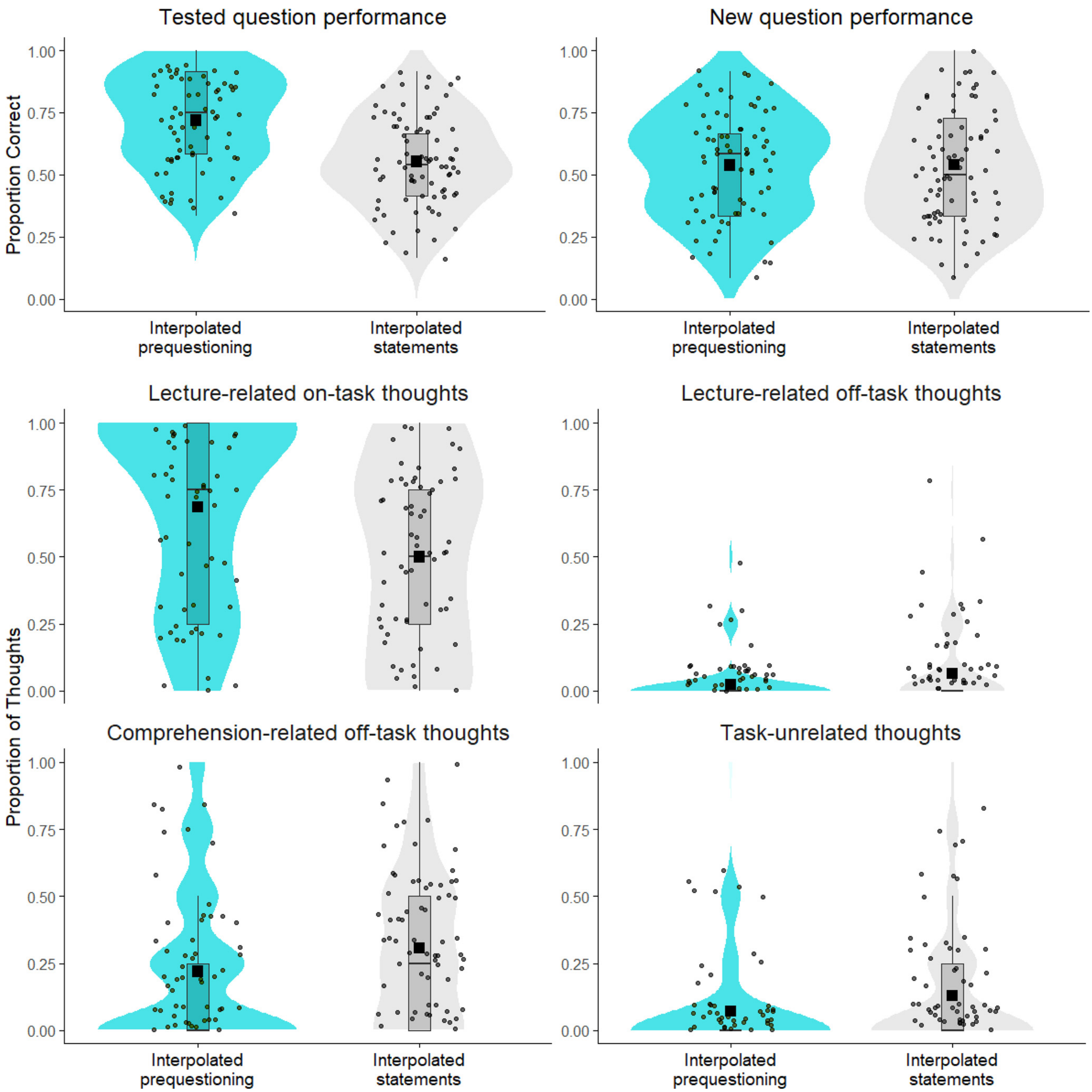
Lecture-Related On-Task Thoughts. The interpolated prequestioning group demonstrated higher rates of lecture-related on-task thoughts compared with those in the interpolated statements group ($M = 0.69$, $SD = 0.35$ vs. $M = 0.50$, $SD = 0.36$). This 36% increase was statistically significant ($p = .0017$) and reflected an effect size of $d = 0.52$, which is virtually identical to Experiment 1.

Lecture-Related Off-Task Thoughts. Although overall rates of lecture-related off-task thoughts were quite low, the interpolated prequestioning group demonstrated significantly lower rates of such thoughts compared with the interpolated statements group ($M = 0.024$, $SD = 0.085$ vs. $M = 0.064$, $SD = 0.14$). That 63% reduction was statistically significant ($p = .038$) and reflected an effect size of $d = –0.34$.

Comprehension-Related Off-Task Thoughts. Rates of comprehension-related off-task thoughts were numerically lower in the interpolated prequestioning compared with the interpolated statements group ($M = 0.22$, $SD = 0.29$ vs. $M = 0.31$, $SD = 0.28$), with an effect size of $d = –0.31$ and no significant effect in the ANOVA results (Table 6). A TOST equivalence test, however, showed that the 90% CI for the mean difference [–0.17, –0.012] fell outside the equivalence margin [–0.14, 0.14], suggesting that the observed 29% reduction in comprehension-related off-task thoughts in the interpolated prequestioning group cannot be considered statistically equivalent to zero.

Task-Unrelated Thoughts. Rates of task-unrelated thoughts were numerically lower in the interpolated prequestioning compared with the interpolated statements group ($M = 0.072$, $SD = 0.18$

Figure 5
Experiment 3 Criterial Test Performance and Thought Probe Results



Note. See the online article for the color version of this figure.

vs. M = 0.13, SD = 0.23), with an effect size of d = –0.27 and no significant effect in the ANOVA results (Table 6). A TOST equivalence test, however, showed that the 90% CI for the mean difference [–0.11, –0.001] fell outside the equivalence margin [–0.10, 0.10], suggesting that the observed 46% reduction in comprehension-related off-task thoughts in the interpolated prequestioning group cannot be considered statistically equivalent to zero.

Mediation Analyses

Given the occurrence of significant prequestioning effects, we conducted mediation analyses analogous to those performed for Experiment 1 and using the same procedures. The results are depicted in Figure 6.

On-Task Thoughts and Overall Criterial Test Performance. In the analysis examining effects on overall criterial test performance,

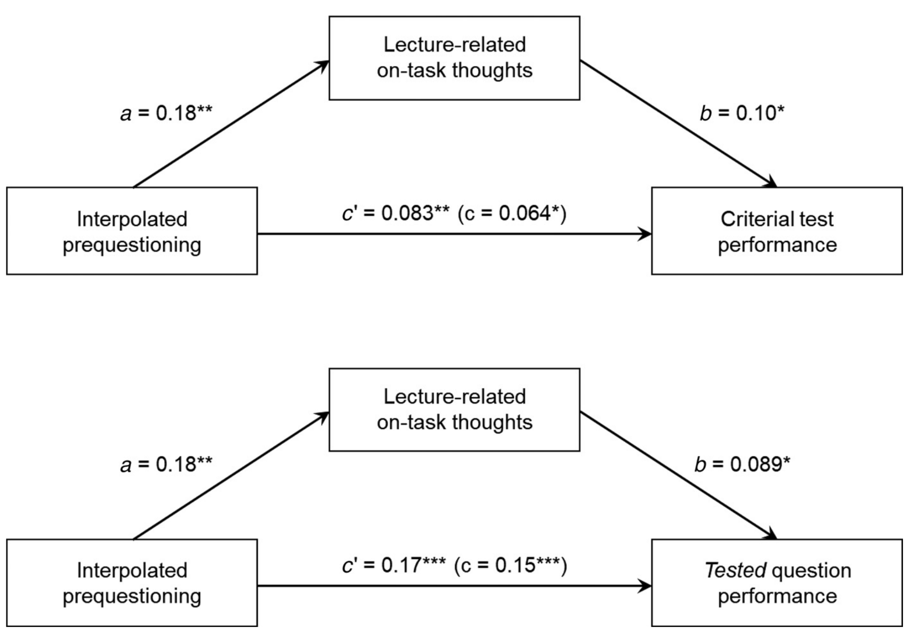
Table 6
Outcomes of Omnibus Analysis of Variance of Thought Probe Responses in Experiment 3

Analysis	*df*	*F*	*p*	η_p^2
Lecture-related on-task thoughts: 2 (Training Group: Interpolated Prequestioning vs. Interpolated Statements) × 4 (Probe Number: 1 vs. 2 vs. 3 vs. 4)				
Main effect of training group	1, 149	10.25	.00017**	.06
Main effect of probe number	3, 447	8.37	.001***	.05
Interaction	3, 447	0.51	.68	.0034
Lecture-related off-task thoughts: 2 (Training Group: Interpolated Prequestioning vs. Interpolated Statements) × 4 (Probe Number: 1 vs. 2 vs. 3 vs. 4)				
Main effect of training group	1, 149	4.36	.038*	.03
Main effect of probe number	3, 447	1.56	.20	.01
Interaction	3, 447	0.17	.92	.0012
Comprehension-related off-task thoughts: 2 (Training Group: Interpolated Prequestioning vs. Interpolated Statements) × 4 (Probe Number: 1 vs. 2 vs. 3 vs. 4)				
Main effect of training group	1, 149	3.64	.058	.02
Main effect of probe number	3, 447	1.57	.20	.01
Interaction	3, 447	0.39	.76	.0026
Task-unrelated thoughts: 2 (Training Group: Interpolated Prequestioning vs. Interpolated Statements) × 4 (Probe Number: 1 vs. 2 vs. 3 vs. 4)				
Main effect of training group	1, 149	2.78	.098	.02
Main effect of probe number	3, 447	7.463	<.001***	.05
Interaction	3, 447	0.24	.87	.0016

* $p < .05$. ** $p < .01$. *** $p < .001$.

the total effect of interpolated prequestioning was significant (β = 0.083, p = .0053; 95% CI [0.025, 0.14]). Upon inclusion of the mediator, lecture-related on-task thoughts, the direct effect was reduced (β = 0.064, p = .033; 95% CI [0.0053, 0.12]), although it remained significant. The indirect effect through lecture-related on-task thoughts was significant (β = 0.019; 95% CI [0.0033, 0.040]), indicating that lecture-related on-task thoughts mediated the relationship between prequestioning and criterial test performance. Hence, interpolated prequestioning increased lecture-related on-task thoughts, which, in turn, enhanced overall criterial test performance.

Figure 6
Experiment 3 Mediation Analyses



Note. Relationships between interpolated prequestioning, lecture-related on-task thoughts, and criterial test performance (upper panel) and tested question performance (lower panel).
* $p < .05$. ** $p < .01$. *** $p < .001$.

On-Task Thoughts and Performance on Tested Questions. In the analysis of performance on tested questions, the total effect of interpolated prequestioning was significant ($\beta = 0.17$, $p < .001$; 95% CI [0.11, 0.23]). When including lecture-related on-task thoughts as a mediator, the direct effect remained significant ($\beta = 0.15$, $p < .001$; 95% CI [0.088, 0.21]) and the indirect effect was also significant ($\beta = 0.016$; 95% CI [0.0013, 0.038]). Hence, lecture-related on-task thoughts mediated the effect of prequestioning on performance for tested questions.

Supplementary Probes

The overall rate of answer discovery in the interpolated prequestioning group was somewhat higher than in prior experiments ($M = 84\%$, $SD = 17\%$), possibly due to the reduced difficulty of the questions. Overall difficulty ratings were roughly comparable between the interpolated prequestioning ($M = 2.74$, $SD = 0.89$) and interpolated statements ($M = 2.83$, $SD = 0.93$) groups, however, and somewhat lower than the ratings for Experiment 2, whereas the interpolated prequestioning group had higher average predictions of future test performance than the interpolated statements group ($M = 71\%$, $SD = 18\%$ vs. $M = 55\%$, $SD = 20\%$), $t(149) = 5.03$, $p < .001$, $d = 0.82$. When examined separately by group, the prequestioning group's predictions were significantly correlated with actual performance, adjusted $R^2 = .047$, $F(1, 73) = 4.686$, $p = .034$, $\eta_p^2 = .06$, as were the predictions from the learning objectives group, adjusted $R^2 = .192$, $F(1, 77) = 19.59$, $p < .001$, $\eta_p^2 = .20$. These patterns are very similar to those observed in Experiment 2, and moreover, match overall test performance for tested questions and all questions in the interpolated prequestioning and interpolated statements groups, respectively (see Table 2).

Discussion

Experiment 3 demonstrated that benefits of prequestioning for learning and on-task are replicable in an online setting. As in Experiment 1, interpolated prequestioning yielded better criterial test performance than reading interpolated statements, but in this case via short answer with embedded answer options as opposed to multiple-choice prequestions. Correspondingly, interpolated prequestioning significantly increased lecture-related on-task thoughts, significantly reduced lecture-related off-task thoughts, and appeared to yield meaningful reductions in comprehension-related and task-unrelated thoughts. Furthermore, lecture-related on-task thoughts mediated criterial test performance in the prequestioning group, just as in Experiment 1. Together, these results suggest that many of the speculated causes of the reduced prequestioning effects in Experiment 2 are not determinative or not insurmountable. These factors include sample characteristics, unsupervised environments, and inability to remember the prequestions or learn the correct answers. Overall, they reinforce the conclusion that interpolated prequestioning can improve focused attention and, in turn, enhance learning from video lectures.

General Discussion

Across three experiments—using validated thought probes, a competitive control condition, and a design that avoided contamination from other learning activities—interpolated prequestioning influenced learners' thought patterns during prerecorded video lectures and, in two of three instances, significantly improved learning outcomes. In all experiments, interpolated prequestioning increased lecture-related on-task thoughts relative to reading interpolated learning objectives ($ds = 0.42$–0.56). Comprehension-related and task-unrelated off-task thoughts also showed consistent numerical decreases across experiments, whereas lecture-related off-task thoughts significantly decreased in the third experiment ($d = 0.34$). Furthermore, in a supplementary analysis of combined data from all three experiments (see online supplemental materials for full details), performed in alignment with preregistered analysis plans, there were significant benefits of interpolated prequestioning for lecture-related on-task thoughts ($d = 0.46$), comprehension-related off-task thoughts ($d = –0.26$), and task-unrelated thoughts ($d = –0.24$), largely reflecting patterns observed in Experiments 1 and 3.

As for learning outcomes, interpolated prequestioning significantly enhanced criterial test performance in Experiments 1 and 3, but not Experiment 2 (which, uniquely in this study, had prequestions which only targeted content toward the end of each video segment). Mediation analyses indicated that the learning benefits observed in Experiments 1 and 3—which were most apparent for tested information (i.e., a specific benefit; $ds = 1.10$ and 0.89, respectively) but possibly extended to untested information—were driven by increases in lecture-related on-task thoughts. Together, these results highlight attentional engagement as a potential mechanism through which prequestioning enhances learning. In contrast, Experiment 2 showed no significant prequestioning effect and the weakest statistical evidence for increases in on-task thoughts, suggesting that a failure to improve a prequestioning effect for learning outcomes occurs in conjunction with reduced attentional benefits.

Overall, the present study advances our understanding of how interpolated prequestioning influences attentional engagement and learning from video lectures—a mode of learning that is increasingly the default option in many educational contexts and for which the efficacy of question-based instructional approaches has yet to be fully established. The results conceptually replicate and extend those of Pan, Sana, Schmitt, et al. (2020), who reported initial evidence with numerical probes that prequestioning can reduce mind wandering during video lectures relative to an irrelevant control activity, plus imply that the null findings of Welhaf et al. (2022) for two different approaches to prequestioning and mind wandering during video lectures, albeit a well-powered study that made use of well-validated thought probes, may nevertheless not generalize across all prequestioning contexts. Notably, attentional benefits emerged even against a strong comparison condition—studying interpolated learning objectives—that also previewed upcoming content and has shown prior efficacy (Osueke et al., 2018; Yu et al., 2025). Hence, the attentional benefits of interpolated prequestioning are not just apparent versus control conditions with minimal relevance to the task at hand. Experience sampling further showed that the effects of interpolated prequestioning extended beyond dampening task-unrelated thoughts, primarily boosting lecture-related on-task thoughts. As discussed next, these results have theoretical and practical implications, including for research on prequestioning and real-world educational practices.

Theoretical Implications for Prequestioning and Relevant Research

The distinct benefits of interpolated prequestioning versus studying interpolated learning objectives suggest that although both can act as advance organizers—alerting learners to upcoming instructional content and potentially activating relevant prior knowledge—prequestioning uniquely engages productive learning processes. Specifically, because it requires participants to generate guesses and, in most cases, search for the correct answers, prequestioning appears to invoke generative and active learning behaviors (Brod, 2021; Chi & Wylie, 2014; Fiorella & Mayer, 2016; Freeman et al., 2014) to an extent that studying learning objectives does not. Instead, the latter may encourage a more passive form of learning (i.e., although the learning objectives informed participants of important concepts and ideas they would encounter, this information may have simply led them to focus more closely, at best, only when those specific elements appeared). Hence, the present research identifies a unique property of prequestioning that differentiates it from at least one type of advanced organizer or relevance instructions. It also raises the prospect that prequestioning may confer similar advantages over yet other approaches which aim to direct learners' attention to different aspects of to-be-learned materials. In our estimation, it is likely that interpolated prequestioning highlighted knowledge gaps and prompted a search for answers that, at least in Experiments 1 and 3, persisted throughout much of each video segment. This ongoing search probably sustained attention throughout a greater portion of the lecture, with participants keeping their initial guesses in mind until they could confirm or correct them.

When considering prominent theories of the prequestioning effect, the present results are consistent with, but not meant to adjudicate between, the attention-direction (e.g., Little & Bjork, 2016), error correction (e.g., Kang et al., 2011), and curiosity accounts (e.g., Berlyne, 1966). To reiterate, these accounts can be interpreted as implying that prequestioning fosters improved attentional engagement during subsequent learning activities. The present results also challenge certain interpretations of other accounts—such as the possibility that prequestioning activates schemata or background knowledge (e.g., Mayer et al., 2002)—which posit that attentional processes are unaffected. Instead, they support alternate interpretations, such as that alongside activated schemata or knowledge, attentional focus improves as well. Our findings also converge with empirical evidence showing changes in eye gaze and reading time behavior following prequestioning (Lewis & Mensink, 2012; Yang et al., 2021). Ultimately, it appears that when prequestioning prompts learners to actively search for answers, improved attentional focus tends to occur. Relatedly, it is notable that the extent to which participants successfully discovered the answers to the prequestions was predictive of criterial test performance (for further details, see the online supplemental materials).

It should also be noted that we repeatedly observed prequestioning effects for tested but not new questions. That pattern is common in prequestioning studies (St. Hilaire et al., 2024). As previously mentioned, a recent explanation for that pattern has been that learners maintain heightened attention until the answers to prequestions are discovered (Sana & Carpenter, 2023); if so, having prequestions target the materials toward the end of a text or video should yield greater learning benefits. The results of Experiment 2, however, imply that such heightened attention may not always be sustainable over time. An alternate possibility is that interpolated prequestioning has a focusing effect on learning rather than enhancing learning more broadly. It may also be the case that directly prequestioned content occupies substantial working memory and attentional resources, leading to less successful learning of nonprequestioned information. Nevertheless, the results of Pan, Sana, Schmitt, et al. (2020; Experiment 2) demonstrate that there are cases wherein prequestioning can enhance learning of both directly tested and untested information.

The present findings also highlight the importance of distinguishing between different types of on-task and off-task thoughts, a distinction increasingly emphasized in contemporary mind wandering research (Kawagoe & Kase, 2021; Pavlova, 2024; Seli et al., 2016). From the standpoint of resource-allocation theories of attention—which posit that executive control processes dynamically distribute cognitive resources between task-focused processing and task-unrelated thoughts (Thomson et al., 2015; see also Lavie, 2010)—it appears that interpolated prequestioning effectively shifts this balance toward enhanced on-task engagement. Further investigation, however, is needed to determine the extent to which increases in on-task thoughts invariably entail concomitant reductions in comprehension-related and task-unrelated off-task thoughts.

In addition, across experiments, it is notable that interpolated prequestioning did not increase negative emotions or perceived difficulty compared with studying interpolated learning objectives (cf. Pan, Sana, Samani, et al., 2020). A decline in attention over the course of a lecture, as observed in some prior research (e.g., Bunce et al., 2010; Pan, Sana, Schmitt et al., 2020) but not in other cases (e.g., Wammes et al., 2016), also did not manifest. Accordingly, it is possible that both types of interpolated learning activities helped maintain more stable levels of attention throughout the lecture, and because emotional valence and perceived difficulty were similar across conditions, these factors evidently did not contribute to differences in attentional engagement (cf. Mills et al., 2013; Smallwood et al., 2009). Finally, it is notable that prequestioning led to higher judgments of learning in the latter two experiments, but only in the final experiment were those judgments truly justified (although it should be noted that the question soliciting those judgments only referred to an upcoming test in general terms, without specifying whether those questions would reflect previously viewed prequestions or learning objectives directly). These results add to the evidence that learners are not always accurate in detecting the benefits of prequestioning for learning (Pan & Rivers, 2023).

Implementation Factors Involving Prequestioning

The results of Experiments 1 and 3 suggest that interpolated prequestioning can enhance attentional engagement and learning from video lectures across varied conditions, including in-person and online formats, with undergraduate or more diverse adult learners, and using multiple-choice or cued recall tests. Hence, interpolated prequestioning can, with some flexibility, support improvements in attention and learning across a variety of educationally relevant circumstances. An important implementation factor may have been that in both experiments, the two prequestions preceding each segment targeted different portions of the segment, allowing participants to encounter answers at two points rather than waiting until the end. The lack of such distributed prequestioned content may have contributed to the weak prequestioning effects in Experiment 2 (which were also inconsistent with the results of Sana & Carpenter, 2023, although differences including the use of interpolated rather than noninterpolated prequestioning, the number of prequestions, and sample characteristics may have also been responsible).

It is important to emphasize that we explicitly instructed participants to discover the correct answers to the prequestions during the video segments, which likely helped them maintain focus and may have contributed to the observed prequestioning effects. Answer search instructions were exclusive to the prequestioning condition (although in the case of learning objectives, we essentially gave participants directions to learn specific concepts and processes). In the prequestioning literature, similar instructions have been used in studies reporting strong learning benefits (e.g., Sana et al., 2020; St. Hilaire & Carpenter, 2020), although in other cases, the search for answers is only implied (e.g., St. Hilaire et al., 2019) or omitted entirely (e.g., Richland et al., 2009). Hence, explicitly requiring a search for answers might increase the likelihood of significant prequestioning effects, but may not always be essential. It should also be noted that the present experiments featured even greater interpolation than Pan, Sana, Schmitt, et al. (2020)—across eight lecture segments rather than four—thus providing additional opportunities for "attention resets" that may have sustained engagement over time. Further research could determine the extent to which the amount of interpolation is influential, particularly for lectures of varying lengths, and in comparison to massing all the prequestions before the lecture itself.

Study Limitations and Future Directions

Despite addressing limitations of prior research, the present experiments remain limited in other respects that future studies could explore. For instance, the total video lecture duration—including interpolated learning activities—was still relatively short at approximately half an hour; future work could apply the present design to longer lectures and different subject materials (e.g., history vs. mathematics). Such research could also further explore issues raised in the present research, including the location of prequestioned content in lecture segments and how participants are instructed (or not) to discover the correct answers.

Another potential limitation is that participants from online crowdsourcing platforms can misinterpret categorical thought probes of the type popularized by Kane et al. (2017, 2021), yielding potentially misleading estimates. Even after we clarified that participants should respond honestly and without fear of violating instructions or suffering reprisals—which, alongside instructions to be attentive, are standard in online crowdsourcing research (Hauser & Schwarz, 2016; Muszynski, 2023)—rates of on-task thought in Experiments 2–3 were higher and task-unrelated thought were lower than in the in-person sample in Experiment 1 (see Table 4). That result, which may stem from the taboo against straying off-task in online studies, raises the possibility that other probe types may be needed to capture thought patterns with greater precision. Such research might also involve lecture materials that tend to yield different distributions of thought patterns (e.g., higher rates of task-unrelated thoughts).

It is also acknowledged that although all three experiments shared many design features, more than one aspect—including prequestion format and the type of content targeted by the prequestions—was altered across experiments. That approach reflected our interest in identifying conditions under which prequestioning most effectively enhances learning, rather than isolating design elements individually. In some cases, however, interpretation is consequently limited; in particular, Experiment 3 ruled out online sampling as the driving factor behind the weak prequestioning effects in Experiment 2 but leaves prequestion format and the content targeted by the prequestions as potentially crucial factors. Additional experimentation could provide further insights into their relative contributions.

Future research could also compare prequestioning against yet other alternative learning activities. For instance, most prequestioning research to date has pitted prequestioning against a reading control condition (Pan & Carpenter, 2023; St. Hilaire et al., 2024), in some cases with precise controls for time-on-task (in the present experiments, participants in the prequestioning groups may have spent longer due to time needed for answer deliberation or type out short answer responses (Experiment 3); it should be noted, however, that exploratory mediation analyses reported in the online supplemental materials show that time spent on prevideo activities did not significantly mediate the observed prequestioning effect for tested questions). In the context of learning from lectures, doing so might involve reading factual statements drawn from the lecture, which are similar, but not identical, to statements outlining learning objectives. Both reading factual statements and reading learning objectives can be considered passive and nonactive activities, and it is possible that broadly similar results might have been obtained had the former been employed.

Educational Implications

The present results add to growing evidence that interpolating prequestions throughout a lecture can scaffold attention and enhance learning. Given the longstanding challenge of maintaining engagement during lectures, whether prerecorded or live, this finding has considerable practical importance, particularly in the increasingly common contexts of online and remote learning (Brammer & Punyanunt-Carter, 2022; see also Schacter & Szpunar, 2015). In addition, our results suggest that interpolated prequestions may serve as more effective advance organizers than studying learning objectives or other relatively passive approaches, yielding greater attentional and learning benefits.

In online, asynchronous, and other minimally supervised learning environments where student attention can be difficult to sustain, interpolated prequestioning stands out as a potential solution. Moreover, although further research is needed, such prequestioning may also be beneficial for live lectures and other instructional formats, such as assigned readings. Accordingly, integrating interpolated prequestions into prerecorded lectures via learning management systems or into live sessions using interactive platforms could provide a direct way to apply the present findings into educational practice. Ultimately, adding interpolated prequestions into instructional design may offer a scalable and practical means of enhancing attentional engagement and learning outcomes across various educational contexts.

References

Anderson, R. C., & Biddle, W. B. (1975). On asking people questions about what they are reading. In G. H. Bower (Ed.), *Psychology of learning and motivation* (Vol. 9, pp. 89–132). Academic Press.

Appelbaum, M., Cooper, H., Kline, R. B., Mayo-Wilson, E., Nezu, A. M., & Rao, S. M. (2018). Journal article reporting standards for quantitative research in psychology: The APA Publications and Communications Board task force report. *American Psychologist*, *73*(1), 3–25. https://doi.org/10.1037/amp0000191

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Ben-Shachar, J. R., Lüdecke, L., & Makowski, P. (2024). effectsize: Estimating the effect size of statistical models. (R Package Version 0.8.3). https://cran.r-project.org/package=effectsize

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, E. L., Soderstrom, N. C., & Little, J. L. (2015). Can multiple-choice testing induce desirable difficulties? Evidence from the laboratory and the classroom. *American Journal of Psychology*, *128*(2), 229–239. https://doi.org/10.5406/amerjpsyc.128.2.0229

Brammer, S. E., & Punyanunt-Carter, N. M. (2022). Getting the attention of online learners. *Communication Education*, *71*(2), 155–157. https://doi.org/10.1080/03634523.2021.2022732

Brod, G. (2021). Generative learning: Which strategies for what age? *Educational Psychology Review*, *33*(4), 1295–1318. https://doi.org/10.1007/s10648-020-09571-9

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

Bunce, D. M., Flens, E. A., & Neiles, K. Y. (2010). How long can students pay attention in class? A study of student attention decline using clickers. *Journal of Chemical Education*, *87*(12), 1438–1443. https://doi.org/10.1021/ed100409p

Butterfield, B., & Mangels, J. (2003). Neural correlates of error detection and correction in a semantic retrieval task. *Brain Research*, *17*(3), 793–817. https://doi.org/10.1016/S0926-6410(03)00203-9

Caldwell, K. (Ed.). (2023). Learning out loud: A framework for learning in an era of information abundance. In *Reimagining education—The role of e-learning, creativity, and technology in the post-pandemic era*. IntechOpen. https://doi.org/10.5772/intechopen.1002923

Carpenter, S. K., & Toftness, A. R. (2017). The effect of prequestions on learning from video presentations. *Journal of Applied Research in Memory & Cognition*, *6*(1), 104–109. https://doi.org/10.1016/j.jarmac.2016.07.014

Carrier, M., & Pashler, H. (1992). The influence of retrieval on retention. *Memory & Cognition*, *20*(6), 633–642. https://doi.org/10.3758/BF03202713

Chan, J. C. K., Ahn, D., Szpunar, K. K., Assadipour, Z., & Gill, H. (2025). In-lecture quizzes improve online learning for university and community college students. *Communications Psychology*, *3*(1), Article 54. https://doi.org/10.1038/s44271-025-00234-5

Chan, J. C. K., Meissner, C. A., & Davis, S. D. (2018). Retrieval potentiates new learning: A theoretical and meta-analytic review. *Psychological Bulletin*, *144*(11), 1111–1146. https://doi.org/10.1037/bul0000166

Chi, M. T. H., & Wylie, R. (2014). The ICAP framework: Linking cognitive engagement to active learning outcomes. *Educational Psychologist*, *49*(4), 219–243. https://doi.org/10.1080/00461520.2014.965823

Cumming, G. (2014). The new statistics: Why and how. *Psychological Science*, *25*(1), 7–29. https://doi.org/10.1177/0956797613504966

Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J., & Willingham, D. T. (2013). Improving students' learning with effective learning techniques: Promising directions from cognitive and educational psychology. *Psychological Science in the Public Interest*, *14*(1), 4–58. https://doi.org/10.1177/1529100612453266

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

Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., & Wenderoth, M. P. (2014). Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(23), 8410–8415. https://doi.org/10.1073/pnas.1319030111

Garcia, M. (2015). *Collector* [Software]. https://github.com/gikeymarcia/Collector

Geller, J., Carpenter, S. K., Lamm, M. H., Rahman, S., Armstrong, P. I., & Coffman, C. R. (2017). Prequestions do not enhance the benefits of retrieval in a STEM classroom. *Cognitive Research: Principles and Implications*, *2*(1), Article 42. https://doi.org/10.1186/s41235-017-0078-z

Gupta, M. W., Pan, S. C., & Rickard, T. C. (2024). Interaction between the testing and forward testing effects in the case of cued-recall: Implications for theory, individual difference studies, and application. *Journal of Memory and Language*, *134*, Article 104476. https://doi.org/10.1016/j.jml.2023.104476

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

Hauser, D. J., & Schwarz, N. (2016). Attentive Turkers: MTurk participants perform better on online attention checks than do subject pool participants. *Behavior Research Methods*, *48*(1), 400–407. https://doi.org/10.3758/s13428-015-0578-z

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

Hollins, T. J., Seabrooke, T., Inkster, A., Wills, A., & Mitchell, C. J. (2023). Pre-testing effects are target-specific and are not driven by a generalised state of curiosity. *Memory*, *31*(2), 282–296. https://doi.org/10.1080/09658211.2022.2153141

Hollis, R. B., & Was, C. A. (2016). Mind wandering, control failures, and social media distractions in online learning. *Learning and Instruction*, *42*, 104–112. https://doi.org/10.1016/j.learninstruc.2016.01.007

Igartua, J.-J., & Hayes, A. F. (2021). Mediation, moderation, and conditional process analysis: Concepts, computations, and some common confusions. *The Spanish Journal of Psychology*, *24*, Article e49. https://doi.org/10.1017/SJP.2021.46

Imundo, M. N., Zung, I., Whatley, M. C., & Pan, S. C. (2025). When two learners are better than one: Using flashcards with a partner improves metacognitive accuracy. *Metacognition and Learning*, *20*(1), Article 3. https://doi.org/10.1007/s11409-024-09406-w

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*, Article 104076. https://doi.org/10.1016/j.compedu.2020.104076

Jing, H. G., Szpunar, K. K., & Schacter, D. L. (2016). Interpolated testing influences focused attention and improves integration of information during a video-recorded lecture. *Journal of Experimental Psychology: Applied*, *22*(3), 305–318. https://doi.org/10.1037/xap0000087

Kane, M. J., Smeekens, B. A., Meier, M. E., Welhaf, M. S., & Phillips, N. E. (2021). Testing the construct validity of competing measurement approaches to probed mind-wandering reports. *Behavior Research Methods*, *53*(6), 2372–2411. https://doi.org/10.3758/s13428-021-01557-x

Kane, M. J., Smeekens, B. A., von Bastian, C. C., Lurquin, J. H., Carruth, N. P., & Miyake, A. (2017). A combined experimental and individual-differences investigation into mind wandering during a video lecture. *Journal of Experimental Psychology: General*, *146*(11), 1649–1674. https://doi.org/10.1037/xge0000362

Kang, S. H. K., McDermott, K. B., & Roediger, H. L. (2007). Test format and corrective feedback modify the effect of testing on long-term retention. *European Journal of Cognitive Psychology*, *19*(4–5), 528–558. https://doi.org/10.1080/09541440601056620

Kang, S. H. K., Pashler, H., Cepeda, N. J., Rohrer, D., Carpenter, S. K., & Mozer, M. C. (2011). Does incorrect guessing impair fact learning? *Journal of Educational Psychology*, *103*(1), 48–59. https://doi.org/10.1037/a0021977

Kawagoe, T., & Kase, T. (2021). Task-related thought and metacognitive ability in mind wandering reports: An exploratory study. *Psychological Research*, *85*(4), 1626–1632. https://doi.org/10.1007/s00426-020-01346-9

Klimovich, M., & Richter, T. (2024). Metacognitive training can reduce mindless reading. *Journal of Educational Psychology*, *116*(5), 836–852. https://doi.org/10.1037/edu0000884

Knopps, A., & Wissman, K. (2023). Collaborative retrieval practice reduces mind-wandering during learning. *Experimental Psychology*, *70*(4), 241–248. https://doi.org/10.1027/1618-3169/a000589

Kornell, N., & Vaughn, K. E. (2016). How retrieval attempts affect learning: A review and synthesis. *Psychology of Learning and Motivation*, *65*, 183–215.

Kornell, N., Hays, M. J., & Bjork, R. A. (2009). Unsuccessful retrieval attempts enhance subsequent learning. *Journal of Experimental Psychology; Learning, Memory, and Cognition*, *35*(4), 989–998. https://doi.org/10.1037/a0015729

Kornell, N., Rabelo, V. C., & Klein, P. J. (2012). Tests enhance learning—Compared to what? *Journal of Applied Research in Memory & Cognition*, *1*(4), 257–259. https://doi.org/10.1016/j.jarmac.2012.10.002

Lakens, D. (2017). Equivalence tests: A practical primer for t tests, correlations, and meta-analyses. *Social Psychological and Personality Science*, *8*(4), 355–362. https://doi.org/10.1177/1948550617697177

Lakens, D., & Caldwell, A. R. (2021). Simulation-based power analysis for factorial analysis of variance designs. *Advances in Methods and Practices in Psychological Science*, *4*(1), Article 251524592095150. https://doi.org/10.1177/2515245920951503

Lavie, N. (2010). Attention, distraction, and cognitive control under load. *Current Directions in Psychological Science*, *19*(3), 143–148. https://doi.org/10.1177/0963721410370295

Lenth, R. P. (2024). *emmeans: Estimated marginal means, aka least-squares means*. R Package (Version 1.7.1). https://cran.r-project.org/package=emmeans

Lewis, M. R., & Mensink, M. C. (2012). Prereading questions and online text comprehension. *Discourse Processes*, *49*(5), 367–390. https://doi.org/10.1080/0163853X.2012.662801

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

Little, J. L., Bjork, E. L., Bjork, R. A., & Angello, G. (2012). Multiple-choice tests exonerated, at least of some charges: Fostering test-induced learning and avoiding test-induced forgetting. *Psychological Science*, *23*(11), 1337–1344. https://doi.org/10.1177/0956797612443370

Lorch, J. R. F., Pugzles Lorch, E., & Klusewitz, M. A. (1995). Effects of typographical cues on reading and recall of text. *Contemporary Educational Psychology*, *20*(1), 51–64. https://doi.org/10.1006/ceps.1995.1003

Lorch, R. F. (1989). Text-signaling devices and their effects on reading and memory processes. *Educational Psychology Review*, *1*(3), 209–234. https://doi.org/10.1007/BF01320135

Makel, M. C., & Plucker, J. A. (2014). Facts are more important than novelty: Replication in the education sciences. *Educational Researcher*, *43*(6), 304–316. https://doi.org/10.3102/0013189X14545513

Martin, L., Mills, C., D'Mello, S. K., & Risko, E. F. (2018). Re-watching lectures as a study strategy and its effect on mind wandering. *Experimental Psychology*, *65*(5), 297–305. https://doi.org/10.1027/1618-3169/a000412

Mayer, R., Mathias, A., & Wetzell, K. (2002). Fostering understanding of multimedia messages through pre-training: Evidence for a two-stage theory of mental model construction. *Journal of Experimental Psychology: Applied*, *8*(3), 147–154. https://doi.org/10.1037//1076-898x.8.3.147

McCrudden, M. T., & Rapp, D. N. (2017). How visual displays affect cognitive processing. *Educational Psychology Review*, *29*(3), 623–639. https://doi.org/10.1007/s10648-015-9342-2

McCrudden, M. T., & Schraw, G. (2007). Relevance and goal-focusing in text processing. *Educational Psychology Review*, *19*(2), 113–139. https://doi.org/10.1007/s10648-006-9010-7

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

Mera, Y., Rodríguez, G., & Marin-Garcia, E. (2022). Unraveling the benefits of experiencing errors during learning: Definition, modulating factors, and explanatory theories. *Psychonomic Bulletin & Review*, *29*(3), 753–765. https://doi.org/10.3758/s13423-021-02022-8

Metcalfe, J. (2017). Learning from errors. *Annual Review of Psychology*, *68*(1), 465–489. https://doi.org/10.1146/annurev-psych-010416-044022

Metcalfe, J., & Huelser, B. J. (2020). Learning from errors is attributable to episodic recollection rather than semantic mediation. *Neuropsychologia*, *138*, Article 107296. https://doi.org/10.1016/j.neuropsychologia.2019.107296

Mills, C., D'Mello, S., Lehman, B., Bosch, N., Strain, A., & Graesser, A. (2013). What makes learning fun? Exploring the influence of choice and difficulty on mind wandering and engagement during learning. In H. C. Lane, K. Yacef, J. Mostow, & P. Pavlik (Eds.), *Artificial intelligence in education* (pp. 71–80). Springer. https://doi.org/10.1007/978-3-642-39112-5_8

Montagni, I., Cariou, T., Tzourio, C., & González-Caballero, J. L. (2019). "I don't know", "I'm not sure", "I don't want to answer": A latent class analysis explaining the informative value of nonresponse options in an online survey on youth health. *International Journal of Social Research Methodology*, *22*(6), 651–667.

Mooneyham, B. W., & Schooler, J. W. (2013). The costs and benefits of mind-wandering: A review. *Canadian Journal of Experimental Psychology*, *67*(1), 11–18. https://doi.org/10.1037/a0031569

Motz, B., Chinni, A., Leeuw, J., Jankowski, H., Aggarwal, A., Amato, M., Berlin, K., Britten, K., Brown, A., Cerchiaro, M., Evans, N., Findley, A., Gorman, R., Gregg, K., Hansen, K., Hullinger, R., Larkin, P., Lion, M., Long, R., & Fyfe, E. (2024). *ManyClasses 2: The effects of prequestions on media interactions and learning*. OSF. https://doi.org/10.31234/osf.io/3xbma

Muszynski, M. (2023). Attention checks and how to use them: Review and practical recommendations. *Ask Research and Methods*, *32*, 3–38. https://doi.org/10.18061/ask.v32i1.0001

Navarro, D. (2015). *Learning statistics with R: A tutorial for psychology students and other beginners* (Version 0.6). https://learningstatisticswithr.com

Nosek, B. A., Hardwicke, T. E., Moshontz, H., Allard, A., Corker, K. S., Dreber, A., Fidler, F., Hilgard, J., Struhl, M. K., Nuijten, M. B., Rohrer, J. M., Romero, F., Scheel, A. M., Scherer, L. D., Schönbrodt, F. D., & Vazire, S. (2022). Replicability, robustness, and reproducibility in psychological science. *Annual Review of Psychology*, *73*(1), 719–748. https://doi.org/10.1146/annurev-psych-020821-114157

Onan, E., Wiradhany, W., Biwer, F., Janssen, E. M., & de Bruin, A. B. H. (2022). Growing out of the experience: How subjective experiences of effort and learning influence the use of interleaved practice. *Educational Psychology Review*, *34*(4), 2451–2484. https://doi.org/10.1007/s10648-022-09692-3

Osueke, B., Mekonnen, B., & Stanton, J. D. (2018). How undergraduate science students use learning objectives to study. *Journal of Microbiology & Biology Education*, *19*(2), Article e19.2.69. https://doi.org/10.1128/jmbe.v19i2.1510

Palan, S., & Schitter, C. (2017). Prolific.ac—A subject pool for online experiments. *Journal of Behavioral and Experimental Finance*, *17*, 22–27. https://doi.org/10.1016/j.jbef.2017.12.004

Pan, S. C., & Carpenter, S. K. (2023). Prequestioning and pretesting effects: A review of empirical research, theoretical perspectives, and implications for educational practice. *Educational Psychology Review*, *35*(4), Article 97. https://doi.org/10.1007/s10648-023-09814-5

Pan, S. C., & Rivers, M. L. (2023). Metacognitive awareness of the pretesting effect improves with self-regulation support. *Memory & Cognition*, *51*(6), 1461–1480. https://doi.org/10.3758/s13421-022-01392-1

Pan, S. C., Sana, F., Samani, J., Cooke, J., & Kim, J. A. (2020). Learning from errors: Students' and instructors' practices, attitudes, and beliefs. *Memory*, *28*(9), 1105–1122. https://doi.org/10.1080/09658211.2020.1815790

Pan, S. C., Sana, F., Schmitt, A. G., & Bjork, E. L. (2020). Pretesting reduces mind wandering and enhances learning during online lectures. *Journal of Applied Research in Memory & Cognition*, *9*(4), 542–554. https://doi.org/10.1016/j.jarmac.2020.07.004

Pavlova, M. (2024). A dual process model of spontaneous conscious thought. *Consciousness and Cognition*, *118*, Article 103631. https://doi.org/10.1016/j.concog.2023.103631

Peterson, D. J., & Wissman, K. (2020). Using tests to reduce mind-wandering during learning review. *Memory*, *28*(4), 582–587. https://doi.org/10.1080/09658211.2020.1748657

Poerio, G. L., Totterdell, P., & Miles, E. (2013). Mind-wandering and negative mood: Does one thing really lead to another? *Consciousness and Cognition*, *22*(4), 1412–1421. https://doi.org/10.1016/j.concog.2013.09.012

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

R Core Team. (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org

Randall, J. G., Hanson, M. D., & Nassrelgrgawi, A. S. (2022). Staying focused when nobody is watching: Self-regulatory strategies to reduce mind wandering during self-directed learning. *Applied Psychology*, *71*(4), 1428–1464. https://doi.org/10.1111/apps.12366

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

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

Rowland, C. A. (2014). The effect of testing versus restudy on retention: A meta-analytic review of the testing effect. *Psychological Bulletin*, *140*(6), 1432–1463. https://doi.org/10.1037/a0037559

Sana, F., & Carpenter, S. K. (2023). Broader benefits of the pretesting effect: Placement matters. *Psychonomic Bulletin & Review*, *30*(5), 1908–1916. https://doi.org/10.3758/s13423-023-02274-6

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

Schacter, D. L., & Szpunar, K. K. (2015). Enhancing attention and memory during video-recorded lectures. *Scholarship of Teaching and Learning in Psychology*, *1*(1), 60–71. https://doi.org/10.1037/stl0000011

Schweppe, J., Radloff, P., & Fenzl, A. (2025). Do prequestions support learning from text more than learning objectives? *Journal of Experimental Psychology: Applied*. Advance online publication. https://doi.org/10.1037/xap0000558

Seabrooke, T., Mitchell, C. J., Wills, A. J., Inkster, A. B., & Hollins, T. J. (2022). The benefits of impossible tests: Assessing the role of error-correction in the pretesting effect. *Memory & Cognition*, *50*(2), 296–311. https://doi.org/10.3758/s13421-021-01218-6

Seli, P., Risko, E. F., Smilek, D., & Schacter, D. L. (2016). Mind-wandering with and without intention. *Trends in Cognitive Sciences*, *20*(8), 605–617. https://doi.org/10.1016/j.tics.2016.05.010

Signorell, A. (2021). DescTools: Tools for Descriptive Statistics. https://cran.r-project.org/package=DescTools

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2012). *A 21 Word Solution* (SSRN Scholarly Paper No. 2160588). Social Science Research Network. https://doi.org/10.2139/ssrn.2160588

Singmann, H. (2022). *Afex: Analysis of Factorial Experiments*. https://CRAN.R-project.org/package=afex

Smallwood, J., & Schooler, J. W. (2006). The restless mind. *Psychological Bulletin*, *132*(6), 946–958. https://doi.org/10.1037/0033-2909.132.6.946

Smallwood, J., Fitzgerald, A., Miles, L. K., & Phillips, L. H. (2009). Shifting moods, wandering minds: Negative moods lead the mind to wander. *Emotion*, *9*(2), 271–276. https://doi.org/10.1037/a0014855

St. Hilaire, K. J., & Carpenter, S. K. (2020). Prequestions enhance learning, but only when they are remembered. *Journal of Experimental Psychology: Applied*, *26*(4), 705–716. https://doi.org/10.1037/xap0000296

St. Hilaire, K. J., Carpenter, S. K., & Jennings, J. M. (2019). Using prequestions to enhance learning from reading passages: The roles of question type and structure building ability. *Memory*, *27*(9), 1204–1213. https://doi.org/10.1080/09658211.2019.1641209

St. Hilaire, K. J., Chan, J. C. K., & Ahn, D. (2024). Guessing as a learning intervention: A meta-analytic review of the prequestion effect. *Psychonomic Bulletin & Review*, *31*(2), 411–441. https://doi.org/10.3758/s13423-023-02353-8

Szpunar, K. K., Khan, N. Y., & Schacter, D. L. (2013). Interpolated memory tests reduce mind wandering and improve learning of online lectures. *Proceedings of the National Academy of Sciences*, *110*(16), 6313–6317. https://doi.org/10.1073/pnas.1221764110

Thomson, D. R., Besner, D., & Smilek, D. (2015). A resource-control account of sustained attention: Evidence from mind-wandering and vigilance paradigms. *Perspectives on Psychological Science*, *10*(1), 82–96. https://doi.org/10.1177/1745691614556681

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*(3), 370–378. https://doi.org/10.1016/j.jarmac.2018.06.003

Unsworth, N. (2019). Individual differences in long-term memory. *Psychological Bulletin*, *145*(1), 79–139. https://doi.org/10.1037/bul0000176

Wammes, J. D., Boucher, P. O., Seli, P., Cheyne, J. A., & Smilek, D. (2016). Mind wandering during lectures I: Changes in rates across an entire semester. *Scholarship of Teaching and Learning in Psychology*, *2*(1), 13–32. https://doi.org/10.1037/stl0000053

Weinstein, Y. (2018). Mind-wandering, how do I measure thee with probes? Let me count the ways. *Behavior Research Methods*, *50*(2), 642–661. https://doi.org/10.3758/s13428-017-0891-9

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), Article 26. https://doi.org/10.1186/s41235-022-00372-y

Wong, A. Y., Smith, S. L., McGrath, C. A., Flynn, L. E., & Mills, C. (2022). Task-unrelated thought during educational activities: A meta-analysis of its occurrence and relationship with learning. *Contemporary Educational Psychology*, *71*, Article 102098. https://doi.org/10.1016/j.cedpsych.2022.102098

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*, Article 102055. https://doi.org/10.1016/j.lindif.2021.102055

Yu, L., Heng, J. T., Arden-Gardner, L. S., Ding, X. P., & Pan, S. C. (2025). Prequestioning with feedback fosters learning in young children. *Learning and Instruction*, *100*, Article 102187. https://doi.org/10.1016/j.learninstruc.2025.102187

Received June 29, 2025
Revision received February 8, 2026
Accepted April 1, 2026 ■

Journal of Educational Psychology

Interpolated Prequestioning Increases On-Task Thoughts, Dampens Off-Task Thoughts, and Facilitates Learning During Video Lectures

Steven C. Pan, Elisya Tung, Luke Chiang, Alyssa Indrajaya, and Michelle E. Kaku

Online First Publication, June 29, 2026. https://dx.doi.org/10.1037/edu0001055

CITATION

Pan, S. C., Tung, E., Chiang, L., Indrajaya, A., & Kaku, M. E. (2026). Interpolated prequestioning increases on-task thoughts, dampens off-task thoughts, and facilitates learning during video lectures. *Journal of Educational Psychology*. Advance online publication. https://dx.doi.org/10.1037/edu0001055