Received: 25 August 2020 | Revised: 18 February 2022 | Accepted: 29 June 2022

DOI: 10.1002/acp.3984

RESEARCH ARTICLE

WILEY

Specific questions during retrieval practice are better for texts containing seductive details

Alexander Eitel[1,2] | Tino Endres[2] | Alexander Renkl[2]

[1]Department of Psychology, Justus-Liebig-University of Giessen, Giessen, Germany

[2]Department of Psychology, Albert-Ludwigs-University of Freiburg, Freiburg, Germany

Correspondence
Alexander Eitel, Department of Educational Psychology, University of Giessen, Otto-Behaghel-Str. 10D, 35394 Giessen, Germany.
Email: alexander.eitel@psychol.uni-giessen.de

Funding information
Innovation Fund of Albert-Ludwigs-University of Freiburg

Abstract

We investigated how to optimize the effectiveness of retrieval-based learning when the instructional text comprises seductive details (i.e., interesting but irrelevant text adjuncts). Specific questions during retrieval practice should help students focus their recall on main ideas—and not on seductive details, which should in turn foster delayed post-test performance. In this experiment, participants ($N = 103$) learned from an instructional text about coffee, either with or without seductive details; in subsequent retrieval practice, the participants received either unspecific or specific questions (2×2 between-subjects design). One week later, all participants received a delayed posttest assessing learning outcomes. As expected, when the instructional text comprised seductive details, participants given specific questions during retrieval practice had better learning outcomes than those given an unspecific question. We conclude that retrieval tasks should be aligned with learning materials: more specific retrieval tasks are better for materials including irrelevant information.

KEYWORDS
decorative images, learning from text, retrieval-based learning, seductive details, testing effect

1 | INTRODUCTION

If you want your students to better remember instructional contents, then have them practice retrieval! This recommendation is among the few that have found their way from cognitive psychology research to the general public and even to classroom practice, as is evident in several articles in the popular press (e.g., Murphy Paul, 2015) and in practice guides for school and university teachers (Agarwal et al., 2018). The present study investigates whether retrieval practice helps learners recall main ideas from an instructional text. The authors of instructional texts often include details such as fun-facts, anecdotes, or comics to make the text more appealing. Details that are interesting but irrelevant for the learning goals are termed seductive details (e.g., Harp & Mayer, 1998; Mayer et al., 2008), and might compromise effective retrieval-based learning from instructional texts because students might retrieve them at the expense of pertinent information. We therefore investigated how to maximize the effectiveness of retrieval practice when the instructional text includes seductive details.

2 | RETRIEVAL-BASED LEARNING

Retrieval-based learning means that learners engage in actively retrieving information from memory as a key part of their learning process (Roediger et al., 2011). More specifically, learners are provided with opportunities to retrieve what they have learned before (i.e., retrieval practice), which is usually done while preparing for a knowledge test. The act of actively retrieving information from memory during retrieval practice protects against forgetting, as proven in a

plethora of empirical studies from cognitive psychology throughout the last 100 years (e.g., Spitzer, 1939; see also Adesope et al., 2017; Rowland, 2014; Pan & Rickard, 2018; Yang et al., 2021, for recent meta-analyses). Most studies used word lists or short texts to reveal the benefits of retrieval practice on the delayed recall of knowledge (e.g., Karpicke & Roediger, 2008). These empirical studies typically compared recall (often 1-day to several weeks delay) of students who studied and practiced retrieval to students who studied the contents repeatedly. Results from most studies revealed that only after delays lasting at least 1 day, students recalled the instructional contents better when they were asked to study *and* retrieve than to just restudy the contents (e.g., Roediger & Karpicke, 2006). Since then, this benefit of retrieval practice has been identified in conjunction with word lists and more complex materials resembling study materials in school or university (see Yang et al., 2021, for a meta-analysis).

There are several explanations for retrieval-practice effects. Two particularly relevant and prominent accounts for the present study are the elaborative retrieval theory (Carpenter, 2009) and bifurcation model (Kornell et al., 2011). Carpenter (2009) argues that retrieval practice can benefit later retrieval because it triggers spreading activation and semantic elaboration. Spreading activation means that retrieving contents also activates related contents in associative memory even when the related contents are not specifically requested. This coactivation of contents and related contents provides learners with additional cues for reactivating contents from memory at a later point in time. Therefore, learners can better retrieve both the contents specifically requested (at an earlier point in time) as well as related contents. Importantly, this later effect is particularly likely when learners engage in deep semantic elaboration of the contents during the retrieval task, and thus coactivate several other parts in semantic memory. Such deeper semantic elaboration is often accompanied by high mental effort (e.g., Endres & Renkl, 2015) and may be prompted by a demanding retrieval task, for example, by an unspecific retrieval task requiring many elements to be retrieved (Carpenter, 2009). Such an unspecific retrieval task would be, for example, a question whose answer allows the learner to mention all kinds of information from a previous instructional text, such as "please write down all the contents you can remember from the previous text" (Endres et al., 2020).

Although a retrieval task should be demanding for students, it should not be too difficult. Students need to be successful in retrieving (most of) the relevant ideas from the instruction to profit from retrieval practice, at least when neither direct feedback nor a restudy opportunity is provided (Rowland, 2014). According to the bifurcation model, testing enables a more substantial increase in memory strength, but only for those items that are successfully retrieved. Unrecalled items even lose part of their memory strength (Kornell et al., 2011). Hence, the more ideas that are correctly recalled in the retrieval task, the more overall memory strength is fostered and, thus, the more ideas can be recalled later in time. Accordingly, Rowland's (2014) meta-analysis revealed that the benefits of retrieval-based learning depend on the success rate during the initial retrieval tasks. Success rates need to be at least 50% for small effects and exceed 75% for medium effects.

The previously mentioned theories suggest on the one hand that students should be provided with unspecific recall tasks so that they also retrieve ideas closely or even only remotely interrelated (elaborative retrieval theory; Carpenter, 2009). They may thus activate more parts of their semantic network, hence providing them with more retrieval cues, which supports delayed retrieval. Accordingly, results from Endres et al. (2017) suggest that more retrieval cues due to intensified elaboration (having been task-prompted) can lead to better learning outcomes in a delayed posttest. Moreover, in another recent empirical study, unspecific retrieval tasks helped students remember a broader spectrum of information and fostered motivation better than more specific retrieval tasks (Endres et al., 2020).

On the other hand, an unspecific recall task in the sense of "Please write down all the contents you can remember" may not result in the recall of a high enough ratio of central ideas when the initial instructions are longer and more complex (Carpenter et al., 2016; van Gog & Sweller, 2015). In addition, students may recall mainly peripheral or even irrelevant information (e.g., anecdotes) from previous instructions. The design and complexity of the instructional materials and tasks is, thus, a potential moderator of retrieval practice effects. This assumption concurs with the finding of Roelle and Berthold (2017), where the benefits of retrieval practice were higher when less complex rather than highly complex adjunct questions were provided to students. This idea is also consistent with Roelle and Nückles (2019), where students profited more from an unspecific retrieval task ("write down everything that you know") than from generative tasks (organization, elaboration) when they had learned from highly cohesive and elaborated texts. Hence, whether an unspecific task is an effective retrieval task seems to depend on the complexity and design of the learning text. More specific retrieval tasks (with single questions about specific main ideas from the text, e.g. "what is typical about the coffee plant?") may foster delayed retrieval more effectively than an unspecific task because they may provide relevance cues and thus help students focus on the more important idea units from text (goal-focusing model; McCrudden & Schraw, 2007).

Overall, research to date has not demonstrated a general advantage of specific questions or unspecific questions for retrieval-based learning (Rowland, 2014). One reason may be that in previous research—unlike in many realistic educational situations—the instructional materials did *not* contain irrelevant or only remotely relevant information (e.g., in the form of seductive details; Harp & Mayer, 1998). As soon as such information is added to the materials, hypotheses concerning the specificity of retrieval questions may change, as detailed below.

3 | SEDUCTIVE DETAILS

Seductive details are often added to an instruction in an effort to make the learning situation more interesting or entertaining. Such details can be teacher jokes, a professor's anecdotes, or nice-looking but irrelevant illustrations in textbooks. Seductive details are by definition "*interesting*, *but unimportant*, information" (Garner et al., 1989,

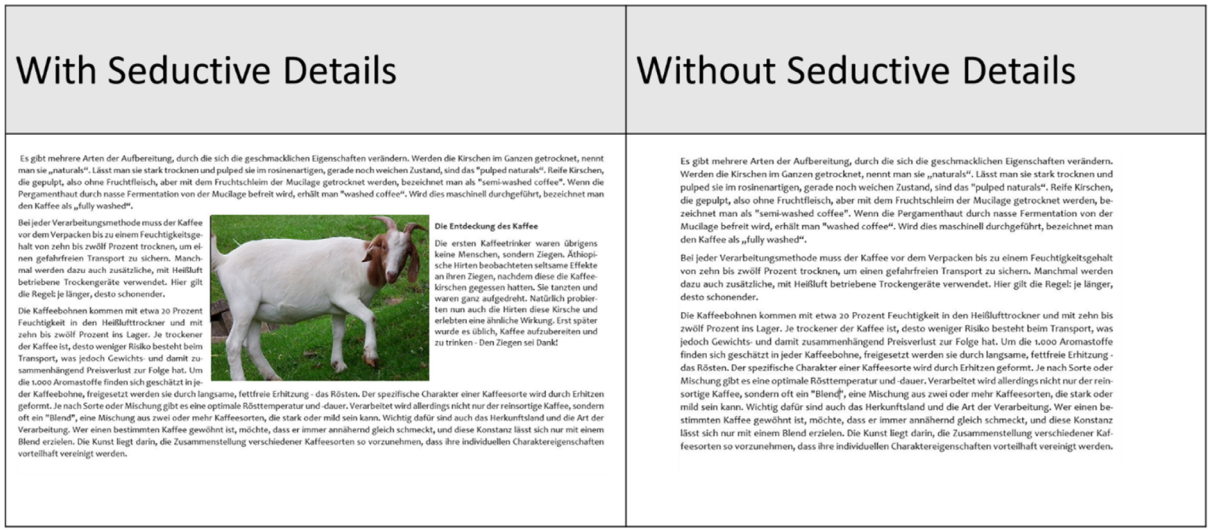

FIGURE 1 One of the eight pages of the instructional text with seductive details (left panel) and without seductive details (right panel). A slightly different picture appeared in the study materials. The current picture was originally published under https://commons.m.wikimedia.org/wiki/File:Wei%C3%9F-braune_Ziege.JPG with CC BY-SA 3.0 license

p. 41; Mayer et al., 2008). In most previous studies, seductive details were interesting or entertaining in association with the contents to-be-learned (context-dependent seductive details; Schraw, 1998). They often referred to everyday life, and were easy to process and to understand (e.g., Park et al., 2015). Moreover, seductive details were identified as *task-irrelevant*, meaning that learning and understanding the seductive details did not contribute to performing better in a subsequent test on the relevant learning contents (Alexander, 2019). For example, when the an instruction's goal is to convey information about coffee processing, the information that "goats actually discovered coffee because they were the first to eat coffee cherries, and then jump around" qualifies as a seductive detail (Figure 1, left panel).

In several studies, seductive details reduced the retrieval of structurally relevant ideas (i.e., main ideas). This finding is known as the seductive details effect (Eitel & Kühl, 2019; Rey, 2012, for overviews). It is typically explained by detrimental cognitive effects *during* learning (Harp & Mayer, 1998): Seductive details distract from processing relevant information, they disrupt from organizing the relevant information into a coherent mental representation, and they divert learners in that they build a mental model around the irrelevant seductive details information by activating inappropriate prior knowledge. To what degree such detrimental cognitive effects of seductive details are bound to the learning phase or (partially) evolve as an interaction between the lesson's design and later retrieval practice is a question addressed in this study.

As seductive details are (by definition) interesting and typically related to everyday life, they are highly likely to be retrieved after learning (e.g., a few minutes later). For example, an unspecific retrieval task such as "write everything down you can remember from the learning phase" does not prevent seductive details information from being retrieved, and they will likely be written down as part of the retrieval task. Such a task, thus, bears the risk that many seductive details will be retrieved at the expense of main ideas, thus leading to a strong seductive-details effect. Although previous research revealed mixed effects about this point (Lehman et al., 2007; McCrudden & Corkill, 2010; Schraw, 1998), the successful retrieval of seductive details bears the risk of displacing the retrieval of main ideas (Harp & Mayer, 1998).

The negative effect of seductive details in retrieval practice (i.e., displacing main ideas) might, however, only be a problem with unspecific retrieval tasks. More specific retrieval tasks, which usually tap on main ideas, may sharply reduce the retrieval of seductive details, thus avoiding the displacement of main ideas. Concerning the goal-focusing model (McCrudden & Schraw, 2007), one could argue that a retrieval task specifically addressing main ideas functions as a (retrospective) relevance cue. With this relevance cue, students may be less distracted and diverted (see Harp & Mayer, 1998) when consolidating the just-learned contents during the retrieval phase. Students focus their retrieval attempts more on main ideas, and less on seductive details, and thus activate more prior knowledge about main ideas than about seductive details during the retrieval phase. Such processing might weaken the seductive details effect. The most recent meta-analysis (Sundararajan & Adesope, 2020) tentatively supports this line of reasoning, as seductive details effect were stronger with an open-ended format ($g = -0.54$) than when specific questions were provided in a multiple-choice test format ($g = -0.20$). Nevertheless, the seductive details did not disappear entirely even with specific questions (e.g., Eitel et al., 2019; Lehman et al., 2007). The explanation for these findings is that seductive details already hampered the process of adequately encoding main ideas during the *learning* phase (e.g., because of their distracting function; Lehman et al., 2007) so that a relevance cue during the *retrieval* phase comes too late to fully counteract the negative seductive-details effect.

It is still an open question whether more specific retrieval tasks—as compared to an unspecific task—actually reduce the seductive details effect. This question is of utmost interest in light of theories about retrieval-based learning (e.g., bifurcation model; Kornell et al., 2011), because the degree to which seductive details (irrelevant information) or main ideas (relevant information) are retrieved successfully after learning may determine how much irrelevant information and how much relevant information are consolidated, and thus later accessible to learners.

4 | RETRIEVAL-BASED LEARNING AND SEDUCTIVE DETAILS: PRESENT RESEARCH AND HYPOTHESES

In the present research, we were primarily interested in how to best implement retrieval practice when the instructional text contains seductive details: Are specific or unspecific questions better after students have learned with instructional texts comprising seductive details (or not)? A delayed post-test after 1 week assessed learning outcomes. We derived our hypotheses by integrating theory and research on retrieval-based learning and on seductive details. Specifically, according to the bifurcation model, only the memory strength of successfully retrieved information is supported (Kornell et al., 2011). Therefore, retrieval as a learning task works best if many main ideas from text are recalled in the immediate test (e.g., Rowland, 2014).

4.1 | The Specific-Questions-Are-Better-with-Seductive-Details hypothesis

With seductive details in the instructional texts, specific questions during retrieval practice should lead to more main ideas being recalled than unspecific questions do, which in turn fosters learning outcomes assessed by delayed posttest performance. Specific questions during retrieval practice act as retrospective relevance cues (McCrudden & Corkill, 2010) to focus recall on relevant information from text. Unspecific questions, however, do not focus students' retrieval practice on the main ideas. In this condition, seductive details may be recalled at the expense of main ideas during retrieval practice, because the former are by definition interesting and easy-to-remember (e.g., Alexander, 2019). Recalling seductive details instead of main ideas during retrieval practice should hamper learning outcomes as assessed in the delayed posttest.

4.2 | The No-Difference-without-Seductive-Details hypothesis

Without seductive details in the instructional texts, both specific and unspecific questions during retrieval practice may lead to a similar number of main ideas being recalled in the immediate test, leading to similar learning outcomes assessed in the delayed posttest. The benefit of specific questions to focus recall on relevant text information may be still present, but it is reduced when the text contains no seductive details (and therefore much less irrelevant information). On the other hand, the unspecific question may be beneficial to delayed posttest performance, because it allows many different elements to being activated and retrieved (spreading activation; Carpenter, 2009), which is accompanied by higher mental effort (Endres & Renkl, 2015).

5 | METHOD

5.1 | Participants and design

Data from 103 participants (81 female, 22 male; $M_{age} = 24.19$, $SD_{age} = 4.66$, range: 19–58 years) were used for our statistical analyses. Most of the participants (85%) were undergraduate students from a German University. The remaining participants were PhD students, unemployed, or worked in diverse professions not directly related to the topic of the instructional texts. All participants were randomly assigned to one of four conditions resulting from a 2×2 between-subjects design with the factors seductive details (with vs. without) and type of retrieval task (unspecific vs. specific). There were 24 participants in the condition with seductive details and unspecific retrieval, 19 participants with seductive details and specific retrieval, 29 participants without seductive details and unspecific retrieval, and 31 participants without seductive details and specific retrieval.

We calculated the required sample size to detect a medium-sized effect ($f^2 = .15$) with an α-level of .05, and a statistical power of $1 - \beta = .80$ using G*Power (Faul et al., 2007). Power calculations refer to a multiple regression analysis (R^2 increase) with three predictors (two main effects; one interaction), so the minimum required sample size was 77. We recruited as many participants as possible. A total of 112 participants took part in this experiment, of whom 108 participants showed up at both sessions. Recruiting was carried out via social media postings, bulletin boards, announcements in university courses, and mailing lists for which participants had volunteered. Data from three participants had to be excluded from the analyses because they had already taken part in a study using the present instructional materials. Two further participants were excluded because of dyslexia and/or zero points in the post-test (indicating total overtaxing), resulting in a total of 103 participants.

5.2 | Materials and manipulations

The instructional text was about coffee and comprised 2454 words in total (see example page in Figure 1). It was pretested in previous research to ensure its subjective readability and intelligibility (Endres et al., 2020). This text's objective readability index amounted to 48.4, meaning a medium level of difficulty resembling typical fictional literature (Lenhard & Lenhard, 2011). The text consisted of four sections: (1) the coffee plant, (2) coffee harvesting, (3) coffee preparation, and (4) coffee processing. Each of the four sections was of similar length and comprised three sub-topics. For instance, the

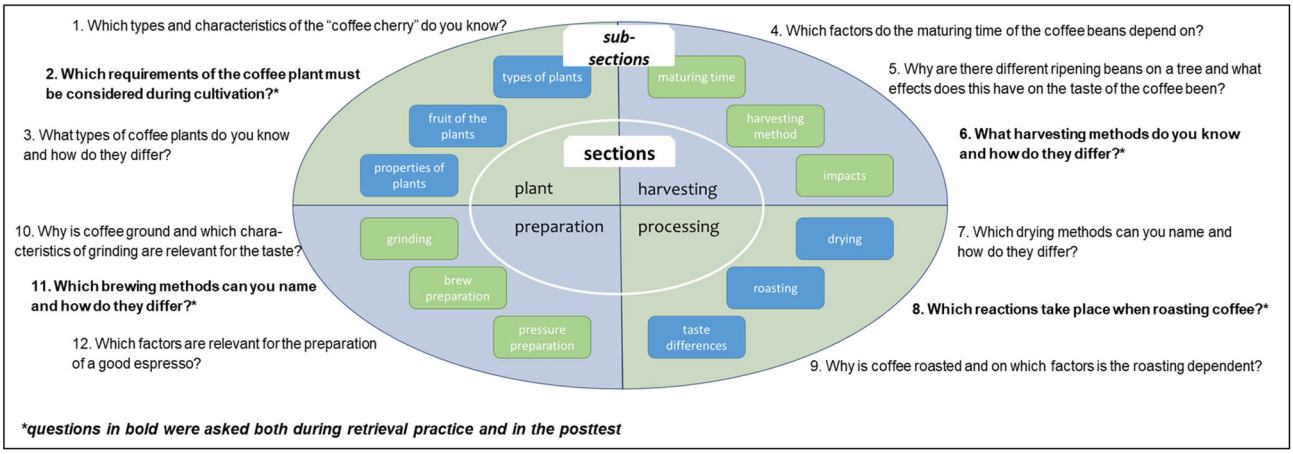

FIGURE 2 Structure and contents of the instructional text (within circle) as well as the specific questions that were asked in the post-test (those in bold were also asked at t1 in the conditions with specific retrieval practice task)

section about the coffee plant dealt with the sub-topic of (1.1) properties of plants, (1.2) fruits of plants, and (1.3) types of plants (Figure 2). To understand a specific paragraph, learners did not need to understand the previous paragraphs. There were also no references to previous sections. For instance, "The coffee plant is a tree and belongs to the group of the overgrowing, dicotyledonous, fused-crowned rubiaceae (reddish plants)" was one of the sentences in the first section. One of the first sentences from the second part was "the harvest time is determined by the degree of latitude and cultivation height." The whole instructional text was spread across eight pages.

5.2.1 | Seductive details manipulation

In two of four conditions, seductive details were added to the instructional text. Seductive details were chosen according to their typical parameters of (1) being easy-to-read, (2) having the potential to increase positive affect during learning by being interesting, entertaining, funny, and/or captivating, and (3) being remotely connected to the topic (here: coffee) but at the same time irrelevant to the task of encoding and consolidating the main ideas from text (cf. *task irrelevance definition* by Alexander, 2019; Garner et al., 1989; Park et al., 2015). Therefore, we searched textbooks and internet pages for illustrations and information about coffee that we considered to be irrelevant with respect to the learning task but of potential interest and (remotely) connected to the topic of coffee. Our search yielded 16 seductive-detail candidates. We had a sample of $N = 40$ participants who were similar in age, gender, and academic track to the participants in the main study (83% were students; 23 female, 11 male; $M_{age} = 24.19$, $SD_{age} = 6.49$, range: 19–47 years) rate the degree to which these candidate seductive details were easy-to-read, interesting, entertaining, funny, captivating, and also whether the details' contents were already known to them (scale from 1 = "not at all", to 5 = "absolutely").

We selected only those eight details for the main study that both achieved the highest aggregate scores for being interesting, entertaining, funny, and captivating (all M_s >3.5, on the 1–5 scale), were rated easy-to-read (all M_s >3.8), and were hardly known beforehand (all M_s <2.0). All these eight seductive details consisted of both a picture (qualifying as decorative; cf. Carney & Levin, 2002) and of a text with an average 63 words. For instance, one seductive detail comprised a picture of a goat together with text about the *discovery of coffee—that it was actually the goats! Goats were the first to eat coffee cherries from the straws, and danced around in consequence. The goatherds saw this, tried the berries themselves, and experienced a similar effect. That is how coffee was discovered.* These seductive details were interspersed with the instructional text (Figure 1, left panel) so that there was always one seductive detail on each of the instruction's eight pages. The seductive detail always appeared as a separate idea unit on the page, similar to a textbox in a textbook. Overall, there were 502 additional words in the instruction due to the seductive details. In the other two conditions, the learners saw no seductive details (Figure 1, right panel).

5.3 | Type of retrieval practice manipulation

In two of four conditions, students received one unspecific question "What do you know about the topic of coffee?" to prompt retrieval practice a few minutes after reading the respective text about coffee (immediate test at time point 1; *t1*). Students were prompted by the program to answer this question in 16 to 20 sentences (800–1600 characters). Specifically, their word count appeared in red if it fell outside the boundaries, and in green when inside it (even though students could nevertheless write more than 1600 characters). In the other two conditions, students received four specific questions about pertinent information from the previously read text. There was one question on each of the four sections in the instructional text: (1) For the coffee plant it was "which requirements of the coffee plant must

be considered during cultivation?,", (2) for coffee harvesting it was "what harvesting methods do you know and how do they differ?," (3) for coffee processing it was "what reactions occur when coffee is being roasted?," (4) for coffee preparation it was "which brewing methods can you name and how do they differ?" Students were prompted to answer each of the four questions in four to five sentences (200–400 characters). Note that overall, learners in both retrieval-practice conditions were prompted to deliver the same amount of text (i.e., 800–1600 words). It is also important to note that not all the information from each subsection needed to be retrieved to answer the specific question.

5.4 | Measures

5.4.1 | Control variables

We asked for the participant's grade point average (GPA) across all school subjects during their last two school years in Germany, where the lower values reflect better overall school performance (from 6 to 1). Prior knowledge was assessed by having participants rate on a scale "How much did you know about the content in the coffee text beforehand?" (from 0% to 100%). To capture participants' working memory capacity, they had to work on the dual *n*-back task (Jaeggi et al., 2010). Specifically, we administered a dual 1-back task, followed by a dual 2-back task, followed by a dual 3-back task. In this dual 1-back task, participants had to memorize whether a letter appeared in the same position within a 3 × 3 grid as the letter from the *previous trial* ("1-back") as well as whether a sound delivered via headphone was the same as the sound heard in the previous trial ("*dual* 1-back"). After 10 training and 50 experimental trials with dual 1-back, participants had to work on the dual 2-back task, meaning that they needed the match letter position and sounds with the letter and sound that had appeared *two trials* before. After 10 training and 50 experimental trials with dual 2-back, participants had to work on the dual 3-back, meaning that they had to match the letter position and sound with those appearing *three trials* earlier. After 10 training and 50 experimental trials, students were finished with the *n*-back task. Because there were ceiling effects for the 1-back task, we aggregated the %correct scores for the 2-back and 3-back tasks into one aggregate *n*-back score for further analyses. Finally, we obtained the control variable of individual study times by summarizing logged study times for the eight pages of instructional text. Note that we assessed other variables, such as having students evaluate the learning and retrieval phase, which are not the focus here. They are reported in the Appendix A.

5.4.2 | Recall of main ideas during retrieval practice

All participants' answers were quantified by first counting the number of main ideas retrieved from the text—with partial credit steps of 0, 0.5, or 1 point depending on how accurately a main idea had been retrieved. In conditions with *unspecific* retrieval task (one question: "What do you know about the topic of coffee?"), students could achieve a maximum score of 43 points for recalling main ideas during retrieval practice, as there were three to five main ideas in each of the text's 12 sub-topics (Figure 2, within circle). Cronbach's alpha was .88. In conditions with *specific* retrieval task, the students' maximum score was 15 points for recalling main ideas during retrieval practice, as there were four questions (one per subtopic; Figure 2, printed in bold outside the circle) that were each assigned up to four points. Cronbach's alpha was .91. Because the maximum scores for successful retrieval practice were different between conditions, we *z*-standardized results separately in these conditions, hence eliminating scale differences as a biasing factor.

5.4.3 | Recall of seductive details during retrieval practice

The number of seductive details recalled from text during retrieval practice were counted; partial credits of 0, 1, and 2 points were provided depending on how accurate those contents were recalled (Cronbach's alpha = .86). This resulted in a maximum score of 16 points for this measure. From this score, we calculated the score for the relative recall of seductive details during retrieval practice, being the number of seductive details recalled during retrieval practice divided by the sum of seductive details and main ideas recalled during retrieval practice. This score should indicate how well students focused their retrieval practice on seductive details.

5.4.4 | Delayed post-test for learning outcomes

The delayed post-test assessing learning outcomes (at time point 2; *t2*) comprised both the unspecific question "What do you know about the topic of coffee?" (maximum score 43 points) and 12 specific questions about the instructional text's contents (see questions outside the circle in Figure 2). The rationale for these 12 questions were: There was always one question on each of the three subtopics (e.g., properties of plants) within each of the four sections (e.g., the coffee plant), leading to 3 × 4 = 12 questions. For instance, one question was "Which types and characteristics of the 'coffee cherry' do you know?". Four of these 12 questions (33%) were identical to the retrieval-practice questions in conditions involving specific retrieval (see highlighted questions in Figure 2).[i] Answers to these questions were also coded by counting the number of main ideas from the text, resulting in maximum score of 43 points. Because both the unspecific and specific questions required retrieving the same information (resulting in a high correlation between these two measures [*r* = .76]), we aggregated scores from those two types of questions by adding them up into one posttest score representing the retrieval of main ideas from text (i.e., learning outcomes). The posttest score's Cronbach's alpha was .83. The theoretical maximum on the post-test scale for main ideas was 86 points. A second rater scored 28% of the data. Inter-rater agreement was analyzed via intraclass correlation (ICC) for consistency with two-way random effects. Inter-rater agreement was good for the post-test scale for main ideas (ICC = .85).

5.4.5 | Delayed post-test for seductive details

The number of seductive details retrieved from text were assigned 0, 1, or 2 points each, resulting in a maximum score of 16 points for this measure (Cronbach's alpha = .82). Finally, the posttest also comprised eight short questions about seductive details, for instance "who discovered coffee?" Again, participants' answers were awarded with 0, 1, or 2 points per seductive detail depending on how accurately they had been recalled (max. 16 points; Cronbach's alpha = .91). Inter-rater agreement was excellent for the seductive-details questions (ICC = .99).

5.5 | Procedure

Participants were invited to the lab for a two-session experimental study. They were informed that the study consisted of two sessions, and that they should return for the second session 1 week later (see Figure 3). Participants provided informed consent during the first session. Specifically, they were informed that their data would be stored and processed anonymously, that participation is voluntary, and that they could end the study without having to fear repercussions. After providing informed consent, students answered questions about demographics (age, gender, dyslexia, German language skills), and worked on the *n*-back task. Afterward, the learning phase began. Students were instructed to read the text attentively and that they would be asked questions about its contents later on. We did not further specify what the to-be-learned contents from the text about coffee were. There was no study time limit. Participants could click through the eight pages of instructional text at their own pace, but they could not go back to previous pages. They were informed about that. Right after learning, participants were asked about what they believed they knew about coffee before they read the text (i.e., retrospective prior knowledge assessment), and about whether they had read this text or a very similar one before. They then reported their mental effort, extraneous cognitive load, situational interest, and affect before completing the filler task (play hang-man game). After 10 min, participants were asked to answer the retrieval practice questions recalling main ideas from text (and potentially also seductive details), with no time limit. Finally, they described the mental effort they had invested in the test, were thanked, and asked to return to the lab 1 week later.

Students came back to the lab exactly 1 week later. They were first asked about whether they had sought additional information about the coffee topic between the two sessions. Afterwards, they were told to take their time while answering the questions in the delayed posttest assessing learning outcomes. The first question in the posttest for all participants was the unspecific one "What do you know about the topic of coffee?" about which they were prompted to answer in 16–20 sentences (800–1600 characters). Participants then worked on 12 specific questions. Participants were prompted to answer each of those 12 questions in four to five sentences (200–400 characters). Participants were always asked to indicate after three specific questions how much effort they had invested and how difficult they had found it answering the questions (via one question each). Last of all, they were given the delayed seductive-details posttest comprising eight short questions about those details. At the end of the second session,

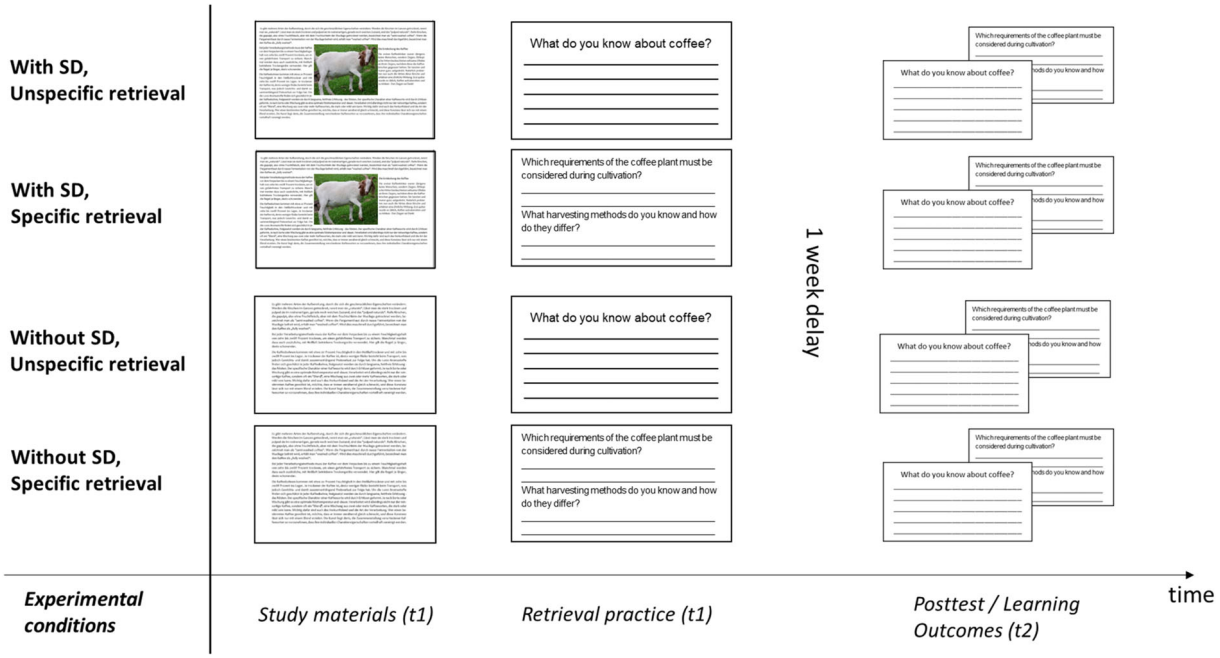

FIGURE 3 Experimental procedure in the four conditions. Note that there were four questions in the specific retrieval conditions. The posttest comprised both the unspecific question and the four specific questions from the retrieval practice phase

participants were debriefed, thanked for their participation, and rewarded with 20 Euros.

6 | RESULTS

We first analyzed the control variables, followed by the learning outcomes, main ideas recalled during retrieval practice, and the recall of seductive details. We applied a .05 alpha level for all statistical tests. We relied on two-sided tests in all analyses. We used partial eta squared as effect size measure for ANOVA results, with values of .01, .06, and .14 indicating a small, medium, and large effect size. We further used Cohens d as an effect size index for planned comparisons, with values of .20, .50, and .80 indicating a small, medium, and large effect size, respectively.

6.1 | Control variables

We first checked to see whether students differed in the control variables across the four conditions. Descriptive values are in Table 1. As none of the control variables was normally distributed, we conducted Kruskal-Wallis tests for independent samples. These tests revealed that students did *not* differ in their school grade point averages (GPA), working memory capacity, age, and gender, *all* $p_s > .26$. However, the number of native German speakers, $\chi^2(3) = 9.01$, $p = .03$, their estimated prior knowledge levels, $H(3) = 10.63$, $p = .01$, and study times, $H(3) = 10.30$, $p = .02$, differed between conditions. Study times were longer among students studying texts with seductive details than among those studying texts without seductive details, $p = .001$.

A linear regression analysis revealed that school GPA, $\beta = -.28$, $p = .003$, working memory capacity, $\beta = .25$, $p = .007$, being a native German speaker (vs. not), $\beta = -.33$, $p = .001$, and study times, $\beta = .32$, $p = .001$, correlated significantly with learning outcomes. In addition, we noted a slight tendency toward significance for the predictor of estimated prior knowledge in the regression analysis, $\beta = .14$, $p = .12$. We thus controlled for the influence of individual scores in school GPAs, working memory capacity, native German speaker, estimated prior knowledge, and study times in further analyses.

6.2 | Delayed post-test for learning outcomes

We analyzed learning outcomes (delayed post-test scores) via a 2 × 2 between-subjects ANCOVA with the factors seductive details (with vs. without) and type of retrieval task (unspecific vs. specific). We detected main effects of seductive details, $F(1, 94) = 3.93$, $p = .050$, $\eta_p^2 = .04$, and of the type of retrieval task, $F(1, 94) = 5.54$, $p = .02$, $\eta_p^2 = .06$. We also found a significant interaction between those two factors, $F(1, 94) = 4.07$, $p = .047$, $\eta_p^2 = .04$. In particular, results are best explained by an ordinal interaction term (contrast: $-3 + 1 + 1 + 1$), $F(1, 94) = 12.67$, $p = .001$, $\eta_p^2 = .12$[1]. Students given a specific retrieval task outperformed those given an unspecific retrieval task when the texts comprised seductive details, $p = .005$, $d = .67$, but not when seductive details were left out, $p = .83$, $d = .19$. These findings support both the Specific-Questions-Are-Better-with-Seductive-Details Hypothesis and the No-Difference-without-Seductive-Details Hypothesis.

6.3 | Recall of main ideas during retrieval practice

We analyzed the recall of main ideas during retrieval practice via a 2 × 2 between-subjects ANCOVA with the factors seductive details (with vs. without) and type of retrieval task (unspecific vs. specific), and identified neither a main effect of seductive details, $F(1, 94) = 2.94$, $p = .09$, $\eta_p^2 = .03$, nor a main effect of the retrieval-task type, $F(1, 94) = 1.43$, $p = .23$, $\eta_p^2 = .02$. However, we did note a significant interaction, $F(1, 94) = 4.25$, $p = .04$, $\eta_p^2 = .04$. Again, these results are best explained by the ordinal interaction term (contrast: $-3 + 1 + 1 + 1$), $F(1, 94) = 7.83$, $p = .006$, $\eta_p^2 = .08$. Again, students given a specific retrieval task outperformed those given an unspecific one when the texts comprised seductive details, $p = .04$, $d = .67$, but not when seductive details were left out, $p = .48$, $d = .13$.

Furthermore, we tested via mediation analysis with 10,000 bootstrap samples (Hayes, 2018) whether the recall of main ideas during retrieval practice would explain the effects of the experimental condition (contrast code: $-3 + 1 + 1 + 1$) on learning outcomes. The indirect (mediation) effect was significant, $b = 0.71$, $SE = 0.23$, Bca CI95 [0.29, 1.19], and reduced the total effect by 56%. Specifically, the

TABLE 1 Means (and standard deviations) of the control variables as a function of experimental condition

	With seductive details		Without seductive details		All subjects
Type of retrieval task	Unspecific ($n = 24$)	Specific ($n = 19$)	Unspecific ($n = 29$)	Specific ($n = 31$)	($N = 103$)
Age	24.25 (2.63)	23.68 (3.35)	23.45 (3.07)	25.16 (7.16)	24.19 (4.66)
Gender (female/male)	18/6	16/3	23/6	24/7	81/22
School GPA (min. = 6.0, max. = 1.0)	1.93 (0.60)	1.87 (0.70)	1.73 (0.63)	2.06 (0.68)	1.90 (0.65)
Prior knowledge (min. = 0, max. = 100)	27.29 (18.80)	19.47 (12.39)	16.45 (18.17)	15.77 (12.93)	19.33 (16.36)
Working memory (min. = 0, max. = 100)	75.02 (10.20)	74.37 (11.01)	75.98 (6.82)	74.71 (10.77)	75.08 (9.60)
German as mother language (frequency in %)	100	73.7	86.2	91.7	89.3
Study times (in seconds)	937.75 (253.18)	925.32 (229.73)	811.65 (243.81)	783.45 (211.02)	853.52 (240.42)

TABLE 2 Adjusted means (and standard errors) of the main dependent variables as a function of experimental condition

	With seductive details		Without seductive details	
Type of retrieval task	Unspecific ($n = 24$)	Specific ($n = 19$)	Unspecific ($n = 29$)	Specific ($n = 31$)
Delayed posttest for learning outcomes (min. = 0, max. = 86)	13.35 (1.23)	18.56 (1.35)	18.21 (1.09)	18.54 (1.07)
Recall of main ideas during retrieval practice (z-scores)	−0.48 (0.19)	0.13 (0.21)	0.24 (0.17)	0.07 (0.17)
Recall of main ideas during retrieval practice (min. = 0, max. = 15/43)[a]	6.39 (0.55)	7.22 (0.61)	8.68 (0.49)	7.13 (0.48)
Delayed posttest for seductive details (min. = 0, max. = 16)	9.00 (0.47)	7.62 (0.52)	0.29 (0.42)	0.61 (0.41)
Relative recall of seductive details during retrieval practice (min. = 0, max. = 1)[b]	0.22 (0.02)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)
Recall of seductive details during retrieval practice (min. = 0, max. = 16)	2.23 (0.23)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)

[a]The theoretical maximum was 43 in the conditions unspecific retrieval, and 15 in the conditions with specific retrieval; therefore, these values were z-standardized in the retrieval conditions separately.
[b]This score is calculated as the number of seductive details retrieved at t1 divided by the number of seductive details + main ideas retrieved at t1.

experimental condition affected the recall of main ideas, $b = 0.16$, $p = .006$, which then affected learning outcomes, $b = 1.27$, $p < .001$.

6.4 | Recalling seductive details

Interestingly, only participants given unspecific questions during retrieval practice recalled seductive details at all (see Table 2). The better these students' relative recall of seductive details was during retrieval practice, the worse their performance was in the delayed posttest, $r = -.25$, $p = .01$, and the better their performance was in the delayed posttest for seductive details, $r = .63$, $p < .001$. Overall, this suggests that specific questions do indeed foster retrieval-based learning by focusing retrieval practice on main ideas from the text - and away from the seductive details.

7 | DISCUSSION

The goal of the present research was to understand how retrieval practice can be made effective when the instructional text contains seductive details. We thus investigated which types of questions during retrieval practice (specific vs. unspecific) work better for instructional texts that either comprise seductive details or not. We derived two hypotheses that we discuss separately.

7.1 | Specific questions are better for texts with seductive details

We expected that specific questions during retrieval practice would improve learning outcomes more than unspecific questions when the instructional texts comprised seductive details. Specific questions should help students focus more closely on the main ideas during retrieval practice so that more of them are recalled and consolidated during retrieval practice to improve learning outcomes in a delayed post-test 1 week later. We found support for this hypothesis. Of the students who had learned from texts comprising seductive details, those who were later asked specific rather than unspecific questions recalled more main ideas from the text both during retrieval practice and in the delayed post-test for learning outcomes. Recalling more main ideas during retrieval practice further explained the beneficial effects of specific versus unspecific questions on learning outcomes.

These findings are clarified via the goal-focusing model (McCrudden & Schraw, 2007) and bifurcation model (Kornell et al., 2011). Specific questions helped students focus their recall activities during retrieval practice on the main ideas from text, and not on the seductive details. The specific questions acted as retrospective relevance cues (McCrudden & Corkill, 2010), signaling which information from the previously read text was important and should be recalled from memory, leading to the successful retrieval of more main ideas from text. Because (according to the bifurcation model) only the memory strength of successfully retrieved information is supported (Kornell et al., 2011), more successful retrieval practice (i.e., more main ideas recalled) explained better learning outcomes as assessed in the delayed posttest.

With *unspecific* questions during retrieval practice, we expected seductive details to be recalled at the expense of main ideas. In line with this expectation, we first found that only students given unspecific questions recalled seductive details at all. As in previous research (e.g., Lehman et al., 2007), the unspecific question prompted seductive details retrieval to a certain degree, whereas the specific questions did not. These findings also support the idea that seductive details are generally well recalled—because they are by definition interesting and easy-to-remember (e.g., Alexander, 2019)—even when one is not specifically asked to recall them. We also found that the relatively more seductive details being recalled during retrieval practice led to their better retrieval 1 week later. This finding is also in line with the

bifurcation model (Kornell et al., 2011) stating that successfully retrieved information (here: seductive details) increases in its relative activation in memory, and thus in its likelihood of being successfully retrieved in a delayed test.

More generally, these findings suggest that relevance cues help keep students from becoming too distracted or diverted by seductive details, not just when such cues are presented in the learning phase (Eitel et al., 2019; Bender et al., 2021) but also when they are presented in the retrieval phase. Retrieval practice that focuses on main ideas can buffer the distracting effects of seductive details while encoding (Harp & Mayer, 1998), and thus protect from the details' detrimental effects on consolidating knowledge.

7.2 | The No-Difference-without-Seductive-Details hypothesis

Without seductive details in the instructional texts, even students given unspecific questions are *not* at risk of recalling seductive details at the expense of main ideas. This greatly weakens the advantage of specific questions to focus recall on main ideas during retrieval practice. We thus hypothesized no significant learning-outcome differences in conjunction with specific and unspecific questions. Supporting this hypothesis was the fact that we found that students in both conditions achieved very similar scores in the delayed posttest for learning outcomes. This finding is line with research evidence of no overall benefit of specific vs. unspecific questions regarding the testing effect when instructional materials contained no seductive details (Endres et al., 2020). Our data also suggest that it was similarly difficult for students with either unspecific or specific questions to retrieve text information, because students in both conditions demonstrated similar recall of main ideas during retrieval practice (see Table 2), and also rated their effort invested in retrieval similarly (see Appendix A). In addition, the recall of main ideas during retrieval practice was similarly associated with the delayed posttest performance with both types of questions.

7.3 | Limitations and further research

Students in our study differed in some of their learning prerequisites (e.g., estimated prior knowledge) across experimental conditions, even though we had assigned them randomly to the conditions. As learning prerequisites also correlated with the dependent variables, we included them as covariates in our analyses following the recommendation by Field (2013). To check for the robustness of these findings, however, we invite further studies with bigger sample sizes, and thus a more likely successful randomization, to replicate the present findings.

Another limitation of our study is that we did not assess the transfer of knowledge based on the text information. While further research should assess such a variable, note that our assessment of immediate and delayed recall of text information reflects previous research in the context of retrieval-based learning (e.g., Endres & Renkl, 2015; Rummer et al., 2017). Furthermore, our questions asked students not just for isolated facts - they sometimes had to integrate ideas from one text to another (e.g., "which brewing methods can you name and how do they differ"). Answers to such questions reflect understanding on a text-base level (van Dijk & Kintsch, 1983) besides reflecting the ability to recall facts. Finally, note that the unspecific and specific questions differed somewhat in the present experiment. The unspecific question asked very generally about what students knew about coffee, whereas each of the four specific questions asked for specific information within one of four text sections, such as what reactions occur when roasting coffee. In doing so, we strongly manipulated the question's specificity to detect any effects in combination with the presence of seductive details in the texts. As we did indeed observe such effects, it is certainly interesting for future research to formulate questions more specifically, such as asking for very specific information within a given text section versus requesting general information from that text section, to find out whether similar effects still appear with such parallel questions.

7.4 | Implications and conclusions

With the present research, we wanted to broaden the empirical and theoretical perspective on research about retrieval-based learning. More specifically, we studied retrieval-based learning without merely focusing on the design and successful implementation of retrieval (for other exceptions, see e.g., Roelle & Berthold, 2017; Roelle & Nückles, 2019), but rather by focusing on the combined effects of the type of study materials (i.e. including seductive details or not) and the type of retrieval practice (unspecific vs. specific). Our results suggest that the effectiveness of unspecific questions during retrieval practice depends on the study materials' coherence. Unspecific questions during retrieval practice are effective when study materials contain little irrelevant information. Specific questions during retrieval practice are also effective when study materials contain more irrelevant information.

As a practical implication, we can conclude that specifically asking for main ideas during retrieval practice is often the better teaching alternative than asking in an unspecific manner. As seductive details are often present in instructional materials in schools or universities, unspecific questions during retrieval practice carry the risk of students retrieving the "wrong" or undesired (seductive details) information at the expense of retrieving the important information. Specific questions lower this risk, and are thus more broadly applicable.

ACKNOWLEDGMENT

Open Access funding enabled and organized by Projekt DEAL.

FUNDING INFORMATION

This research was funded by the Innovation Fund of Albert-Ludwigs-University of Freiburg.

CONFLICT OF INTEREST

The authors have no conflict of interest to disclose.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID

Alexander Eitel https://orcid.org/0000-0001-6403-3467

ENDNOTE

[i] Note that we also ran the ANCOVA separately (1) for questions that were identical between retrieval practice and post-test, and (2) for questions that were not identical between retrieval practice and post-test as dependent variables. We detected a significant ordinal interaction term (contrast: $-3 + 1 + 1 + 1$) for both identical questions, $F(1, 94) = 15.80$, $p < .001$, $\eta_p^2 = .14$, and for questions that were not identical between retrieval practice and posttest, $F(1, 94) = 9.27$, $p = .003$, $\eta_p^2 = .09$. Covariates again were school GPA, working memory capacity, native German-speaking, estimated prior knowledge, and study times.

REFERENCES

Adesope, O. O., Trevisan, D. A., & Sundararajan, N. (2017). Rethinking the use of tests: A meta-analysis of practice testing. *Review of Educational Research*, *87*(3), 659–701. https://doi.org/10.3102/0034654316689306

Agarwal, P. K., Roediger, H. L., McDaniel, M. A., & McDermott, K. B. (2018). *How to use retrieval practice to improve learning*. Washington University in St.

Alexander, P. A. (2019). The art (and science) of seduction: Why, when, and for whom seductive details matter. *Applied Cognitive Psychology*, *33*(1), 142–148. https://doi.org/10.1002/acp.3510

Carpenter, S. K. (2009). Cue strength as a moderator of the testing effect: The benefits of elaborative retrieval. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *35*(6), 1563–1569. https://doi.org/10.1037/a0017021

Carpenter, S. K., Lund, T. J. S., Coffman, C. R., Armstrong, P. I., Lamm, M. H., & Reason, R. D. (2016). A classroom study on the relationship between student achievement and retrieval-enhanced learning. *Educational Psychology Review*, *28*(2), 353–375. https://doi.org/10.1007/s10648-015-9311-9

Eitel, A., Bender, L., & Renkl, A. (2019). Are seductive details seductive only when you think they are relevant? An experimental test of the moderating role of perceived relevance. *Applied Cognitive Psychology*, *33*(1), 20–30. https://doi.org/10.1002/acp.3479

Eitel, A., & Kühl, T. (2019). Harmful or helpful to learning? The impact of seductive details on learning and instruction. *Applied Cognitive Psychology*, *33*(1), 3–8. https://doi.org/10.1002/acp.3513

Endres, T., Carpenter, S., Martin, A., & Renkl, A. (2017). Enhancing learning by retrieval: Enriching free recall with elaborative prompting. *Learning and Instruction*, *49*, 13–20. https://doi.org/10.1016/j.learninstruc.2016.11.010

Endres, T., Kranzdorf, L., Schneider, V., & Renkl, A. (2020). It matters how to recall – Task differences in retrieval practice. *Instructional Science*, *48*(6), 699–728. https://doi.org/10.1007/s11251-020-09526-1

Endres, T., & Renkl, A. (2015). Mechanisms behind the testing effect: An empirical investigation of retrieval practice in meaningful learning. *Frontiers in Psychology*, *6*(July), 1–6. https://doi.org/10.3389/fpsyg.2015.01054

Endres, T., Kranzdorf, L., Schneider, V., & Renkl, A. (2020). It matters how to recall–task differences in retrieval practice. *Instructional Science*, *48*(6), 699–728.

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G* power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. https://doi.org/10.3758/BF03193146

Field, A. (2013). *Discovering statistics using IBM SPSS statistics* (4th ed.). Sage.

Garner, R., Gillingham, M. G., & White, C. S. (1989). Effects of 'seductive details' on macroprocessing and microprocessing in adults and children. *Cognition and Instruction*, *6*(1), 41–57. https://doi.org/10.1207/s1532690xci0601_2

Harp, S. F., & Mayer, R. E. (1998). How seductive details do their damage: A theory of cognitive interest in science learning. *Journal of Educational Psychology*, *90*(3), 414–434. https://doi.org/10.1037/0022-0663.90.3.414

Jaeggi, S. M., Buschkuehl, M., Perrig, W. J., & Meier, B. (2010). The concurrent validity of the N-back task as a working memory measure. *Memory*, *18*, 394–412. https://doi.org10.1080/09658211003702171

Karpicke, J. D., & Roediger, H. L. (2008). The critical importance of retrieval for learning. *Science*, *319*(5865), 966–968. https://doi.org/10.1126/science.1152408

Kornell, N., Bjork, R. A., & Garcia, M. A. (2011). Why tests appear to prevent forgetting: A distribution-based bifurcation model. *Journal of Memory and Language*, *65*(2), 85–97. https://doi.org/10.1016/j.jml.2011.04.002

Klepsch, M., Schmitz, F., & Seufert, T. (2017). Development and validation of two instruments measuring intrinsic, extraneous, and germane cognitive load. *Frontiers in Psychology*, *8*, 425. https://doi.org/10.3389/fpsyg.2017.01997

Knekta, E., & Eklöf, H. (2015). Modeling the test-taking motivation construct through investigation of psychometric properties of an expectancy-value-based questionnaire. *Journal of Psychoeducational Assessment*, *33*(7), 662–673. https://doi.org/10.1177/0734282914551956

Lehman, S., Schraw, G., McCrudden, M. T., & Hartley, K. (2007). Processing and recall of seductive details in scientific text. *Contemporary Educational Psychology*, *32*(4), 569–587.

Lenhard, W., & Lenhard, A. (2011). *Berechnung des Lesbarkeitsindex LIX nach Björnson [calculation of the readability index after Björnson]*. Psychometrica. https://doi.org/10.13140/RG.2.1.1512.3447, http://www.psychometrica.de/lix.html

Magner, U. I., Schwonke, R., Aleven, V., Popescu, O., & Renkl, A. (2014). Triggering situational interest by decorative illustrations both fosters and hinders learning in computer-based learning environments. *Learning and Instruction*, *29*, 141–152. https://doi.org/10.1016/j.learninstruc.2012.07.002

McCrudden, M. T., & Corkill, A. J. (2010). Verbal ability and the processing of scientific text with seductive detail sentences. *Reading Psychology*, *31*(3), 282–300. https://doi.org/10.1080/02702710903256486

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

Murphy Paul, A. (2015). A new vision for testing. *Scientific American*, *313*(2), 54–61. https://doi.org/10.1038/scientificamerican0815-54

Pan, S. C., & Rickard, T. C. (2018). Transfer of test-enhanced learning: Meta-analytic review and synthesis. *Psychological Bulletin*, *144*(7), 710–756. https://doi.org/10.1037/bul0000151

Park, B., Korbach, A., & Brünken, R. (2015). Do learner characteristics moderate the seductive-details-effect? A cognitive-load-study using eye-tracking. *Journal of Educational Technology & Society*, *18*, 24–36.

Rey, G. D. (2012). A review of research and a meta-analysis of the seductive detail effect. *Educational Research Review*, *7*(3), 216–237. https://doi.org/10.1016/j.edurev.2012.05.003

Roediger, H. L., & Karpicke, J. D. (2006). Test-enhanced learning: Taking memory tests improves long-term retention. *Psychological Science*, *17*(3), 249–255. https://doi.org/10.1111/j.1467-9280.2006.01693.x

Roediger, H. L., Putnam, A. L., & Smith, M. A. (2011). Ten benefits of testing and their applications to educational practice. In *Psychology of learning and motivation* (Vol. 55, pp. 1–36). Academic Press. https://doi.org/10.1016/B978-0-12-387691-1.00001-6

Roelle, J., & Berthold, K. (2017). Effects of incorporating retrieval into learning tasks: The complexity of the tasks matters. *Learning and Instruction*, *49*, 142–156. https://doi.org/10.1016/j.learninstruc.2017.01.008

Roelle, J., & Nückles, M. (2019). Generative learning versus retrieval practice in learning from text: The cohesion and elaboration of the text matters. *Journal of Educational Psychology*, *111*(8), 1341–1361. https://doi.org/10.1037/edu0000345

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

Rummer, R., Schweppe, J., Gerst, K., & Wagner, S. (2017). Is testing a more effective learning strategy than note-taking? *Journal of Experimental Psychology: Applied*, *23*(3), 293–300. https://doi.org/10.1037/xap0000134

Satow, L. (2012). *Big-Five-Persönlichkeitstest (B5T): Test- und Skalendokumentation*. Retrieved January, 2019, from https://www.drsatow.de/tests

Schraw, G. (1998). Processing and recall differences among selective details. *Journal of Educational Psychology*, *90*(1), 3–12. https://doi.org/10.1037/0022-0663.90.1.3

Schiefele, U. (2009). Situational and individual interest. In K. R. Wentze & A. Wigfield (Eds.), *Handbook of motivation at school* (pp. 197–222). Routledge.

Spitzer, H. F. (1939). Studies in retention. *Journal of Educational Psychology*, *30*, 641–656. https://doi.org/10.1037/h0063404

Sundararajan, N., & Adesope, O. (2020). Keep it coherent: A meta-analysis of the seductive details effect. *Educational Psychology Review*, *32*, 707–734. https://doi.org/10.1007/s10648-020-09522-4

van Dijk, T. A., & Kintsch, W. (1983). *Strategies of discourse comprehension*. Academic Press.

van Gog, T., & Sweller, J. (2015). Not new, but nearly forgotten: The testing effect decreases or even disappears as the complexity of learning materials increases. *Educational Psychology Review*, *27*(2), 247–264. https://doi.org/10.1007/s10648-015-9310-x

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: the PANAS scales. *Journal of Personality and Social Psychology*, *54*(6), 1063–1070.

Yang, C., Luo, L., Vadillo, M. A., Yu, R., & Shanks, D. R. (2021). Testing (quizzing) boosts classroom learning: A systematic and meta-analytic review. *Psychological Bulletin*, *147*, 399–435. https://doi.org/10.1037/bul0000309

How to cite this article: Eitel, A., Endres, T., & Renkl, A. (2022). Specific questions during retrieval practice are better for texts containing seductive details. *Applied Cognitive Psychology*, *36*(5), 996–1008. https://doi.org/10.1002/acp.3984

APPENDIX A

A1 | ASSESSMENT OF FURTHER VARIABLES

Participants' degree of conscientiousness was assessed by having them respond to seven items such as "I always fulfill my tasks precisely", or "I am very conscientious", on a 7 point Likert scale from „not at all "to „absolutely "(Satow, 2012; Cronbach's alpha = .64). With regard to the *learning phase*, we assessed subjective mental effort via five items (adapted from Knekta & Eklöf, 2015). Specifically, items were "I did my best during reading", or „I could have tried harder during reading "(reversely coded), Cronbach's alpha = .64. We assessed intrinsic cognitive load via two items ("for this task, many things needed to be kept in mind simultaneously"; "this task was very complex"; Cronbach's alpha = .58; Klepsch et al., 2017), and extraneous cognitive load via three items (e.g., "during this task it was exhausting to find the important information; "the design of this task was very inconvenient for learning"; Cronbach's alpha = .77; Klepsch et al., 2017). We also assessed situational interest (cf. Endres et al., 2020; Schiefele, 2009). Specifically, three items were used for the "catch" component (I found the design of the text exciting / entertaining / boring [reverse-coded]"; Cronbach's α = .82), and three items were used for the hold component (I found the contents of the text useful / unnecessary [reverse-coded] / unimportant [reverse-coded]; Cronbach's α = .79). We also assessed the positive affect via seven items from Positive-And-Negative-Affect-Scale (PANAS; Watson et al., 1988); participants had to indicate how active, interested, happy, aroused, awake, determined, and attentive they felt on a five-point scale (from "not at all" to "absolutely").

With regard to the *retrieval practice phase* (retrieval within the first session, at t1), we assessed mental effort via three items (adapted from Knekta & Eklöf, 2015). Specifically, items were "I did my best on this test", "I could have tried harder on this test", and "I felt motivated to do my best on this test" (Cronbach's alpha = .72). These items were also asked after the final posttest.

A2 | RESULTS FOR FURTHER VARIABLES

We analyzed whether there were effects of seductive details on the evaluation of the learning phase. Seductive details significantly affected the catch-component of situational interest ($p < .001$), which is in line with previous research (e.g., Magner et al., 2014). There were no further main effects. Also, subjective retrieval effort was neither affected by seductive details nor by the type of retrieval task nor by an interaction by the two factors (all $ps > .05$) (Table A1).

TABLE A1 Results for further variables

Type of retrieval task	With seductive details		Without seductive details		All subjects ($N = 103$)
	Unspecific ($n = 24$)	Specific ($n = 19$)	Unspecific ($n = 29$)	Specific ($n = 31$)	
Conscientiousness (min. = 1, max. = 7)	2.69 (0.54)	3.06 (0.91)	2.86 (0.70)	2.82 (0.91)	2.85 (0.74)
Study effort (min. = 1, max. = 7)	5.31 (1.08)	5.01 (0.72)	5.23 (0.73)	5.43 (0.77)	5.27 (0.83)
Extraneous cognitive load (min. = 1, max. = 5)	3.53 (1.31)	4.56 (1.15)	4.25 (1.38)	4.00 (1.28)	4.07 (1.32)
Intrinsic cognitive load (min. = 1, max. = 5)	4.00 (1.63)	4.32 (1.46)	4.67 (1.40)	4.34 (1.47)	4.35 (1.48)
Situational interest: catch (min. = 1, max. = 5)	4.21 (1.44)	4.53 (1.39)	3.41 (1.80)	2.84 (1.85)	3.63 (1.77)
Situational interest: hold (min. = 1, max. = 5)	5.15 (1.12)	4.95 (1.16)	5.22 (1.06)	4.80 (1.39)	5.03 (1.19)
Positive affect (min. = 1, max. = 5)	2.90 (0.74)	2.92 (0.62)	2.77 (0.80)	2.77 (0.70)	2.83 (0.72)
Retrieval effort (min. = 1, max. = 7)	4.96 (0.78)	5.21 (0.93)	5.06 (0.67)	5.02 (0.81)	5.05 (0.78)

Notes: Specific References for Further Variables: Klepsch et al. (2017); Knekta and Eklöf (2015); Magner et al. (2014); Satow (2012); Schiefele (2009); Watson et al. (1988).