

Journal of Educational Psychology

Sketching and Summarizing to Reduce Memory for Seductive Details in Science Text

Allison J. Jaeger, Mia N. Velazquez, Anastasia Dawdanow, and Thomas F. Shipley

Online First Publication, January 25, 2018. <http://dx.doi.org/10.1037/edu0000254>

CITATION

Jaeger, A. J., Velazquez, M. N., Dawdanow, A., & Shipley, T. F. (2018, January 25). Sketching and Summarizing to Reduce Memory for Seductive Details in Science Text. *Journal of Educational Psychology*. Advance online publication. <http://dx.doi.org/10.1037/edu0000254>

Sketching and Summarizing to Reduce Memory for Seductive Details in Science Text

Allison J. Jaeger, Mia N. Velazquez, Anastasia Dawdanow, and Thomas F. Shipley
Temple University

Seductive details refers to interesting pieces of information within an expository text that are only tangentially related to the target concept (Garner, Gillingham, & White, 1989). When the presence of this information results in reduced comprehension, this is called the *seductive details effect*. Previous work has found the seductive details effect to be resistant to reduction via various instructional manipulations. One avenue that has not been investigated as a tool for reducing the seductive details effect is having students generate sketches. A growing body of research suggests that sketching activities are beneficial for science learning and, moreover, that sketching can improve learning from science text (Ainsworth, Prain, & Tytler, 2011; Van Meter, 2001). The goal of the present research was to investigate the impact of sketching as compared to generating summaries or thinking silently on recall and comprehension of a text that included seductive details. Across two studies, the seductive details effect was replicated; generating sketches did not eliminate it. In Experiment 2, students compared their sketches and summaries to correct ones and were asked to identify differences between them. Results indicated that participants in the summary group recalled the most core concepts and demonstrated the highest comprehension. These results suggest that sketching may not be effective for eliminating the seductive details and that having students generate summaries with feedback may be more successful. These findings inform the design of scaffolding to support learning from naturalistic science text with its distracting details.

Educational Impact and Implications Statement

The present studies investigated the effectiveness of sketching and summarizing as tools to overcome the seductive details effect—a problematic decrease in recall and problem-solving when interesting yet irrelevant information is included in expository text, like textbooks and other educational materials. Contrary to initial hypotheses, results demonstrated that summarizing, rather than sketching, provided the greatest benefit to recall of core concepts and overall text comprehension. However, both sketching and summarizing were effective strategies to reduce the recall of seductive information. Previous research has shown that sketching is a promising educational tool, but the present studies suggest that more research is needed to determine when and for whom sketching will be most beneficial.

Keywords: seductive details, sketching, science learning, text comprehension

Using Sketching to Reduce Memory for Seductive Details

Across all levels of education, learning depends to a considerable degree on students' ability to effectively engage with and thereby comprehend informational text. In an effort to motivate

and capture the attention of readers, many texts include interesting, yet irrelevant anecdotes or flashy pictures and graphics. Though these additions may pique readers' interest in the text, they can impede successful comprehension of the important concepts—an effect referred to as the *seductive details effect* (Garner et al., 1989). One prominent hypothesis for why the presence of this irrelevant information impacts comprehension is that it distracts readers from attending to the more conceptually important information in the text (Harp & Mayer, 1998). For example, readers may selectively process, and therefore have better memory for, information about the number of deaths associated with an earthquake event at the expense of processing the conceptually relevant information about how and why that earthquake occurred. Despite the fact that irrelevant text and images can hinder comprehension and recall of core conceptual information, many researchers and educators support the idea that motivation and interest are important factors for successful learning and are therefore not keen on removing these items from expository text. Rather, there has been a push to identify interventions that can be used to direct students'

Allison J. Jaeger, Mia N. Velazquez, Anastasia Dawdanow, and Thomas F. Shipley, Department of Psychology, Temple University.

The research presented in this article and preparation of this article was supported in part by National Science Foundation Grant SBE 10-41707 to the Spatial Intelligence and Learning Center (SILC) and in part by National Science Foundation Grant 1640800. We thank Andrew Jarosz, Rachel Myer, Nora Newcombe, and other members of the SILC for their feedback on the design of materials and the development of coding schemes. We also thank Justin Peloso who was a research assistant on this project.

Correspondence concerning this article should be addressed to Allison J. Jaeger, Department of Psychology, Temple University, 1701 North 13th Street, Weiss Hall, Philadelphia, PA 19122. E-mail: allison.jaeger@temple.edu

attention to the more important core concepts and reduce the negative impact of irrelevant information. One instructional manipulation that has been shown to benefit expository science text comprehension more generally is generating concept sketches (Ainsworth et al., 2011). The goal of the present line of research was to examine if generating concept sketches can increase recall and comprehension of core concepts in a science text that contains seductive details, and if generating sketches can be an effective tool for decreasing recall of seductive details.

Learning From Expository Science Text

Learning in science, technology, engineering, and mathematics (STEM) is challenging for many students and therefore it is no surprise that students find science texts difficult to comprehend. In STEM domains, commonly used texts and textbooks tend to be loaded with technical terms that must be deciphered and remembered later on (Sadovski, 2001). In addition, many phenomena in STEM involve highly complex mechanisms with multiple components, attributes of components, relations between components, and dynamic processes that occur at spatial scales that are too large or small to perceive or temporal scales that are too fast or slow to perceive (Gentner & Stevens, 1983; Hegarty & Just, 1993; Mayer, 1989). Furthermore, many students have low knowledge about complex STEM concepts and therefore need to build their understanding with no initial foundation or support, a process that is effortful and can be demotivating. Consequently, the challenge of developing a mental model of expository science text can lead students to develop negative attitudes toward science and influence their motivation to engage in effective learning strategies such as elaboration or inference generation (Graesser, Leon, & Otero, 2002). All of these challenges point to some reasons why expository science text is difficult and why it can be difficult to get and keep students in the STEM pipeline.

A primary assumption in research on learning from text is that comprehension requires the reader to construct multiple levels of representation, both memory for the text and understanding of the meaning of the text (Kintsch, 1998; Kintsch & van Dijk, 1978; van Dijk & Kintsch, 1983). The textbase level representation is a propositional network that captures the meaning relations among elements within a single sentence or across a select few sentences in the text; in general, this level of representation is thought of as representing memory for parts of the text, but not the meaning of the text as a whole. The most relevant level of representation for understanding the complex processes described by many science texts is the mental model or situation model level representation (Gentner & Stevens, 1983; Johnson-Laird, 1983; van Dijk & Kintsch, 1983). The mental model, or situation model, captures what the text is about and supports performance on inference or application questions (Donnelly & McDaniel, 1993; Kintsch, 1994; Mayer, 1989). This level of representation includes a causal model of the phenomena being described and generally requires the reader to generate inferences or make connections across multiple pieces of information to explain how or why these phenomena occur (Chi, 2000; Graesser & Bertus, 1998; Millis & Graesser, 1994; Singer & Gagnon, 1999; Wiley & Myers, 2003). Theories of text comprehension also generally assume that both characteristics of the reader (prior knowledge and abilities) and affordances or features of a text or its topic can affect successful

comprehension (Kintsch, 1988; McNamara, Kintsch, Songer, & Kintsch, 1996; van den Broek, 2010; Wiley, Sanchez, & Jaeger, 2014).

The Seductive Details Effect

One factor that has been demonstrated to impact text comprehension is interest. For students to develop a mental model and accurately comprehend STEM concepts in expository text, they need to be willing to engage in the cognitive activity required (Schraw, 1998). Students who are interested in a task or topic may be more motivated to put in the effort to learn it. Many researchers and educators hold the view that increasing interest and motivation in a topic is important for learning (Hidi, 1990). For example, interest has been shown to correlate well with deep learning from text (Schiefele, 1999). The idea is that when a reader is more interested or engaged with a text, he or she will in turn use more effective strategies such as elaboration (Tarchi, 2017). However, previous research on learning from text has demonstrated that the information or images included in texts to make them more interesting or appealing to readers can sometimes lead to poor learning outcomes—a phenomenon known as the *seductive details effect* (Garner, Brown, Sanders, & Menke, 1992; Garner et al., 1989).

In early work investigating the impact of seductive details in text (Harp & Mayer, 1997, 1998), it was shown that both in terms of recall and problem-solving performance, individuals who were in conditions that contained any type of seductive information (textual, visual, or both) performed worse than individuals in a base text condition. Along with decreased recall and comprehension, participants in the seductive conditions also reported higher levels of emotional interest. This reduction in recall and comprehension as well as the enhancement of emotional interest has been replicated in several studies (Garner et al., 1989; Sanchez & Wiley, 2006; Wade & Adams, 1990). Further, a recent meta-analysis and review revealed a significant seductive details effect with small to medium effect sizes for retention performance and medium effect sizes for problem-solving or transfer performance (Rey, 2012). Despite the evidence that irrelevant, redundant, or distracting information can harm learning, it continues to be prevalent in learning materials across all ages and topic domains (Pozzer & Roth, 2003; Sung & Mayer, 2012) and has more recently been examined in the context of cognitive load (Park, Flowerday, & Brunken, 2015). Further, recent research has demonstrated that both students and instructors regard the inclusion of interesting yet irrelevant information in expository learning materials as beneficial to learning (Morehead, Rhodes, & DeLozier, 2016).

Explanations for the Seductive Details Effect

Different theoretical explanations for the seductive details effect exist. One prominent explanation comes from the Cognitive Theory of Multimedia Learning (CTML) and suggests that because the working memory system is inherently limited, the presence of extraneous or irrelevant information can overload that system (Sweller, 2005). If the learner must process additional information (i.e., seductive details), they may not have the required resources to engage in the cognitive processes needed to develop a situation model for the core conceptual information (Mayer, 2005). Mayer, Griffith, Jurkowitz, and Rothman (2008) found that high-interest

seductive details reduced transfer performance compared to low interest details and argued that this result supported the hypothesis that high-interest details drew more cognitive resources away from processing the core conceptual information than low-interest details. Though this result seems to support the working memory overload explanation, the two conditions were not matched on length—the high-interest condition contained more information and required more reading time—making it difficult to determine the actual cause of the working memory overload.

However, results from several studies have provided evidence against the working memory overload theory. Sanchez and Wiley (2006) demonstrated that low working memory participants had lower performance on a comprehension test and included fewer causal concepts in an essay task than high working memory capacity participants when seductive illustrations were presented alongside an expository text about ice ages. This result seems to support the working memory overload explanation, however Sanchez and Wiley also collected eyetracking data that indicated that the low working memory individuals spent more time looking at the irrelevant images than the high working memory individuals. Based on these results, they argued for an alternative explanation for the seductive details effect—namely the distraction hypothesis originally proposed by Harp and Mayer (1998).

According to the distraction hypothesis, seductive details reduce learning because they are more interesting and require little effort to understand, making it relatively easy to draw the learner's selective attention away from the important conceptual information (Harp & Mayer, 1998). Lehman, Schraw, McCrudden, and Hartley (2007) also found support for the distraction hypothesis and demonstrated that important information in an instructional text received less attention when seductive details were present. More specifically, they showed that participants spent less time with the base text when seductive information was included as compared to participants who only received the base text. Chang and Choi (2014) replicated this pattern of results using an eyetracking methodology and found that increased attention to seductive sentences, as indexed by gaze durations, was a major predictor of participants' poor performance in recall and comprehension.

Attempts to Reduce the Impact of Seductive Details

Because seductive information does appear to increase interest, educators are keen to keep it included in text as a means to motivate their students and get them to engage with the material. However, as previously described, this information can harm comprehension. Therefore, there have been several efforts aimed at reducing the impact of seductive details through various instructional manipulations meant to guide or direct student's attention away from seductive information, and rather focus them on the conceptually relevant content.

In a study by Peshkam, Mensink, Putnam, and Rapp (2011), participants read an expository text that either contained seductive text sentences or read the same text with those sentences removed. Participants were randomly assigned to one of four instructional conditions. In the relevance condition, participants were instructed to use prereading guiding question to help them understand the material. In the specific-irrelevance instruction condition, participants were told that textbooks often contain irrelevant information

and that even the prereading guiding questions can focus them on irrelevant information and may not be helpful for understanding the passage. In both of these groups (relevance and specific-irrelevance instruction), the participants received the same six prereading questions, some of which targeted seductive details and some of which targeted conceptual information. Participants in the general-irrelevance instruction group did not receive any prereading questions and were simply told that sometimes textbooks contain irrelevant information that may not be useful for understanding and to keep in mind that the passage they are going to read may contain irrelevant information that can hinder their learning. The fourth group was a control condition that received no prereading instructions and no relevance instructions. Results indicated that readers attended to seductive details across all conditions, but that participants in the general irrelevance group spent less time reading them. Participants in the general irrelevance group also recalled fewer seductive details than participants in the other groups. The authors suggested that these results demonstrated that a general instruction warning readers about the presence of irrelevant information and the fact that it may harm learning can focus reader attention on the more relevant information and reduce the effects of seductive details on attention and memory for text. This study did not include a direct measure of comprehension or mental model construction, and therefore how a general irrelevance instruction impacts comprehension remains an open question.

Wiley (2003) was able to reduce the negative impact of seductive material on comprehension of a historical topic. Specifically, she found that when seductive images were presented before a text, the negative impact on comprehension was reduced, even though students still reported a higher level of interest compared to a text with no seductive material. Similarly, Wright, Milroy, and Lickorish (1999) showed that presenting an animation before the text as opposed to presenting the animation embedded within the text reduced any detriment to learning. Taken together, these results suggest that moving seductive information to the beginning of a text rather than having it embedded with the text can reduce the seductive details effect. These results may be due to attention being solely devoted to the important conceptual information when it is presented. These results are also consistent with the idea that the effect of seductive details on comprehension occurs during the process of reading or encoding and not after reading or during a recall phase.

Sketching to Learn in Science

Learner-generated sketching or drawing is a strategy in which students construct external representations of to-be-learned content to improve learning of that material (Van Meter, Aleksic, Schwartz, & Garner, 2006). Previous work has indicated that sketching the subject matter described in an expository text can be an effective strategy for learning (for an overview see Leutner & Schmeck, 2014; Van Meter & Firetto, 2013). More recent work has also begun to investigate the effectiveness of learner-generated sketching for improving learning from expository animations (Kombartzky, Ploetzner, Schlag, & Metz, 2010; Lowe & Mason, 2017; Mason, Lowe, & Tornatora, 2013). According to the Generative Theory of Drawing Construction (GTDC), creating a sketch while reading text causes generative processing that takes

students beyond a textbase representation and helps them to develop a situation level representation (Van Meter, 2001). The act of translating the text into an external pictorial representation engages the learner in processing that is aimed at making sense of the material, selecting the most relevant information, mentally organizing it into a coherent structure, and integrating it with prior knowledge (Van Meter & Garner, 2005). Although there is a long-standing tradition of research on other generative learning activities such as self-explanation—the act of generating explanations to oneself during learning from expository text (Chi, 2000)—the use of sketching as a learning strategy is relatively new. Despite this, evidence that sketching can be an effective strategy for learning from text is accumulating (e.g., Gobert & Clement, 1999; Leopold & Leutner, 2012; Scheiter, Schleinschok, & Ainsworth, 2017; Schmeck, Mayer, Opfermann, Pfeiffer, & Leutner, 2014; Schwamborn, Mayer, Thillmann, Leopold, & Leutner, 2010; Van Meter, 2001; Van Meter et al., 2006).

Interestingly, although using sketching manipulations to improve text comprehension in research settings is relatively new, scientists have been using it as a tool for understanding and representing complex concepts and phenomena for a long time (Ainsworth et al., 2011). In fact, because STEM concepts tend to be intensely spatial (e.g., geoscience, engineering) much of the content in these areas is expressed through diagrams, maps, and other visual representations (Jee et al., 2010). As compared to verbal representations such as speech or writing, sketches more readily capture this spatial information and align with the visual-spatial demands of STEM learning (Goel, 1995; Heiser & Tversky, 2002; Suwa & Tversky, 1997; Vosniadou & Brewer, 1992).

Van Meter and Garner's (2005) GTDC presents a processing model of drawing construction that was developed as an extension of Mayer's Generative Theory of Textbook Design, a theory proposed to explain learning from illustrated text (Mayer, 1993; Mayer, Bove, Bryman, Mars, & Tapangco, 1996). GTDC is consistent with Mayer's model, as well as with research on learning more generally, in that it describes selection, organization, and integration of information as three cognitive processes required for meaningful learning. The task of sketching begins with the selection of key elements. When learners draw with no provided illustration they may only select relevant textual information. The selected elements must then be organized into a single coherent representation. During the organization process, elements are linked and new associative connections can be made. In the third step, the learner needs to construct an internal nonverbal representation of the text and connect or integrate it with the verbal representation they constructed during reading. If the learner experiences difficulty when building their mental image or external drawing they may refer back to their verbal representation or the text to detect comprehension errors. According to Van Meter and Garner, this process of aligning or integrating one's drawing with their verbal representation encourages active cognitive and metacognitive processing and thus fosters deeper learning. Further, the Prognostic Drawing Principle (Schwamborn et al., 2010) indicates that the quality of a sketch during learning can serve as a formative assessment that predicts learning and can serve as an indicator that the learner has engaged in appropriate cognitive processing.

Using Sketching to Reduce the Seductive Details Effect

When considering the presence of seductive details in an expository text, it is important that learners' attention is directed toward the important conceptual content rather than the irrelevant details. As described by the GTDC, the selection phase is a crucial step for generating a sketch because it requires the learner to decide which elements to attend to and include in their nonverbal representation. Furthermore, sketches are limited in what they can include. Because they are spatial in nature, they can constrain selection to information that can be represented in a spatial manner. For example, when considering a text about the causes and effects of earthquakes, generating a sketch may facilitate the representation of the movement of two tectonic plates in opposite directions, though it may impede developing a representation of less spatial information such as the specific number of deaths associated with that earthquake event. As such, sketching may be a useful tool for reducing the negative impact of seductive details in expository science text. More specifically, learner-generated sketches should be useful for improving recall and comprehension of an expository text about a scientific phenomenon more generally, but they should also be helpful for directing the learner's attention to the important, spatial concepts within the text and away from the interesting, yet conceptually irrelevant information in the text. To address these hypotheses, the two experiments reported here examined the effectiveness of a sketching task for improving learning and recall from an expository science text and for reducing the impact of seductive details.

Experiment 1

Students were instructed to read an expository science text describing the types of tectonic plate movements and the geologic formations that occur as a result of these movements. Half of the students received the base text only and the other half received the same text, but with seductive details about the destruction associated with volcanoes and earthquakes included (see Appendix A). In addition, students across both text conditions were instructed to complete a series of postreading learning activities: students generated sketches, generated summaries, or thought silently about what they had read. It was predicted that participants in the seductive text group would demonstrate lower recall of core concepts and lower comprehension compared to participants in the base-text only group. Because previous research has demonstrated that sketching can facilitate memory for and the comprehension of expository science text, it was hypothesized that students who were instructed to generate sketches would have better recall and comprehension than students who were instructed to write summaries or think silently.

Further, because sketching requires building a nonverbal representation and spatial information can be more readily represented in this format, it was predicted that students in the sketching group would be less likely to experience the seductive details effect. Specifically, the act of sketching should constrain students' attention to information that can be represented in a sketch (i.e., information related to plate movement and geologic processes) and reduce their attention to verbal information that is more difficult to represent in a sketch (i.e., number of deaths). On the other hand, because generating a summary should be less likely to constrain attention to the more conceptually relevant spatial information,

participants may be more likely to attend to and integrate seductive details into their mental models. Similarly, students left to their own devices in the silent think group should also be likely to attend to the seductive details, as has been demonstrated in prior research.

Method

Participants and design. One-hundred fifty-six students (121 female; $M = 20.35$ years, $SD = 2.91$) from a university in the northeastern part of the United States participated voluntarily in the experiment in exchange for course credit. Data from 38 participants had to be excluded from data analysis; five students were not native English speakers and 12 did not comply with task instructions. In addition, anyone who reported having taken introductory geology ($n = 14$) or had taken seven or more natural science courses ($n = 7$) were excluded from analysis.¹ This resulted in a final sample of 118 participants. The design was a 2 (text condition: base-only, base-plus-seductive) $\times$ 3 (activity condition: sketch, summary, silent think) between-subjects design in which participants were randomly assigned to one text condition and one activity condition. There were 58 participants in the base-only text group and 60 in the base-plus-seductive text group. In addition, 42 participants were in the sketching activity group, 40 in the summary group, and 36 in the silent think group. See Table 1 for descriptive statistics. The study was approved by the ethics committee of the university where the study was conducted, and the study followed standards for ethical treatment of human subjects.

Materials. The materials consisted of six instructional booklets (base with sketching, base with summary, base with think, seductive with sketching, seductive with summary, and seductive with think), a recall sheet, a multiple-choice test, a participant questionnaire, and a psychometric measure of spatial ability. Each instructional booklet was printed on 8.5×11 in. paper and had a cover sheet that included the general task instructions and passage title (*Plate Tectonics*). The instructions read as follows,

In this task, you will be reading a text that is titled *Plate Tectonics*. At different points in the text you will be prompted to complete an activity related to what you have just read. After reading the text and completing all of the activities, you will be asked to recall as much as you can about what you read. Please take your time and read carefully. Be sure to complete every section of this packet. After you have completed a page you will not be allowed to go back to previous pages. If you have any questions, raise your hand and the experimenter will talk to you.

After the cover sheet, the second page of all the booklets included a single Likert-scale rating item; this rating asked students to indicate how much they knew about plate tectonics ranging from 1 (*not much*) to 10 (*very much*).

Each booklet contained an 843 word (41 sentence) instructional base text describing the geologic phenomenon of plate tectonics that was developed for the present set of studies and pilot tested in a prior study. The factual information in the text was taken from high school science textbooks and information found on the U.S. Geological Survey website and was reviewed by expert geologists to insure its scientific accuracy. The base text described the three major types of plate interactions (convergent, divergent, and transform) and the kinds of geologic formations that are created at these various plate boundaries (e.g., volcanoes, island arcs, faults). Fur-

ther, the text explained that the types of geologic formations created by convergence vary depending on whether the plates are made of oceanic lithosphere or continental lithosphere. The text was generally quite difficult to read, as indicated by a relatively high Flesch-Kincaid reading score of 51.6, which is a value in the range of difficulty for nonfiction (McNamara, Graesser, & Louwerse, 2012). The text was divided into five sections that were presented on different pages. Except for the first section which presented a general introduction to plate tectonics and the layers of the Earth, each section explained a different plate interaction and resulting geologic formation: (a) introduction to plate tectonics, (b) convergence that produces stratovolcanoes, (c) convergence that produces island arcs and convergence that produces mountain ranges, (d) transform boundaries that produce recurring earthquakes, and (e) divergence that produces ridges and new crust.

In the seductive text condition booklets, 13 additional sentences were included (318 words). As with prior seductive details work (Harp & Mayer, 1998; Lehman et al., 2007), the seductive details were distributed throughout the base text. The sentences containing seductive details were carefully chosen for their ability to flow within the base passage and were adapted from actual information found in science textbooks and webpages containing information about plate tectonics (see Appendix A for a sample of the base-plus-seductive details text). Other than the addition of the seductive detail sentences, the seductive text condition booklets were presented in the same manner as the base text booklets.

Because the text was developed specifically for this experiment, a pilot test with a separate sample of 28 students from the participant pool was initially conducted. In this pilot, participants first read the entire base-plus-seductive details text. After the initial reading, participants then read each sentence again, one at a time, and rated them for how interesting they found each sentence to be, ranging from 1 (*very uninteresting*) to 6 (*very interesting*), and how important the sentence was for understanding the overall meaning of the text, ranging from 1 (*very unimportant*) to 6 (*very important*). Rating order was counterbalanced. Paired-samples t tests revealed that students rated the base-text sentences ($M = 3.78$, $SD = .66$) as less interesting than the seductive sentences ($M = 4.80$, $SD = .64$), $t(27) = 7.42$, $p < .001$, $d = 1.41$, but rated the base sentences ($M = 4.36$, $SD = .63$) as more important for the overall meaning of the text than the seductive sentences ($M = 3.41$, $SD = 1.23$), $t(27) = 4.19$, $p < .001$, $d = .86$.

Across both the base and base-plus-seductive booklets, four pages were included that guided the participants activity, one activity for each major plate interaction described. In the sketching activity condition, participants were prompted to create sketches relating to the information they had just read. Because previous sketching research has indicated that some external support is

¹ To ensure that there were no differences in attrition associated with condition, a chi-square test was conducted examining the number of participants excluded as a function of text condition and activity condition. There was no difference in attrition across conditions, $\chi^2(2, n = 38) = 1.14$, $p = .56$. Because high knowledge is a learner characteristic that participants presumably come into the study with, we also conducted a chi-square test on participants who were excluded for not following task instructions. This chi-square test also revealed no difference in attrition across condition, $\chi^2(2, n = 17) = .59$, $p = .75$, suggesting that the activity or text manipulations did not lead to increased attrition in any one condition.

Table 1
Descriptive Statistics for Experiment 1

Variable	Base-only			Base-plus-seductive		
	Sketch	Summary	Think	Sketch	Summary	Think
<i>n</i>	21	18	19	19	22	19
Age	20.33 (2.65)	19.89 (2.73)	19.47 (1.07)	20.83 (6.14)	19.73 (.99)	21.00 (2.73)
Gender (% female)	.76	.82	.95	.84	.77	.79
Year in school (1–4)	2.19 (.87)	2.00 (1.14)	2.05 (1.08)	2.00 (.94)	2.05 (.84)	2.63 (1.30)
Science courses	2.50 (2.07)	2.78 (3.10)	3.94 (3.57)	3.84 (4.50)	3.21 (2.59)	2.89 (4.43)
Plate know (1–6)	2.62 (1.80)	2.22 (1.26)	2.32 (1.34)	2.00 (1.41)	1.73 (.94)	2.05 (1.27)
Interest	5.67 (2.22)	5.06 (2.90)	4.68 (2.45)	4.95 (2.48)	4.41 (1.71)	5.79 (2.57)
Try	6.24 (1.84)	6.72 (2.24)	5.68 (1.86)	5.84 (1.98)	5.82 (2.46)	5.74 (2.40)
Recall word count	110.52 (33.20)	110.83 (32.66)	115.47 (41.17)	112.42 (45.19)	110.00 (29.50)	119.26 (37.37)

Note. Year in school: 1 = freshman, 2 = sophomore, 3 = junior, 4 = senior; Science courses = average number of college science courses taken; Plate know = average self-reported knowledge of plate tectonics on a 1–6 Likert scale; Interest = how interesting they found the text on a 1–10 Likert scale; Try = how hard they tried to learn the material on a 1–10 Likert scale.

necessary for improved learning (Van Meter & Garner, 2005), students were not simply instructed to draw a picture of what they had just read but rather were instructed to draw a picture of a specific process described in the text. Specifically, the instructions asked participants to draw a picture of the plate interaction that caused a specific geologic formation to form (stratovolcanoes, island arcs, mountains ranges, recurring earthquakes, and the creation and destruction of oceanic crust). Identical instructions were given to students in the summary condition, but the phrase “draw a picture of . . .” was replaced with the phrase “write a summary about . . .” In the silent think condition, participants were instructed to spend 2 min reviewing in their mind what they had just read. The goal of all of these specific instructions were to help guide the learners about what to focus on during their learning activity.

The recall sheet consisted of a single 8.5 × 11 in. page and had the following instruction typed at the top:

We would now like you to recall everything you can about the passage you just read entitled *Plate Tectonics*. Don't worry about spelling and punctuation. Try to remember as much as you can. If you can only remember some of the meaning from a sentence, include that too. You will get 7 minutes to recall as much as you can.

The multiple-choice test consisted of 10 items presented on an 8.5 × 11 in. paper. Five of the items were classified as memory items because the answer to these questions could be found directly in the text. The other five items were classified as inference items because they required the reader to make connections between various parts of the text to generate a correct answer. Test type was counterbalanced so half of the participants received the five inference items first, whereas others received the five memory items first.

Each student completed a paper-and-pencil version of the Paper-Folding Test from the Kit of Factor Referenced Cognitive Tests (Ekstrom, French, Harman, & Dermen, 1976). In this test, the participant must determine which one of five possible patterns of holes will result after a square piece of paper goes through a sequence of folds and then is punched. Since the 1940s, this test has commonly been used as a measure of spatial visualization skill, which, broadly defined, represents one's ability to mentally transform or manipulate objects (Carroll, 1993). The test consisted of

20 items, divided into two parts, presented one at a time. Participants had 3 min to complete each part. Participants' scores were the total number of correct responses across both parts. This task was selected as the measure of spatial visualization skill because it has been demonstrated to be a strong predictor of performance or aptitude in STEM areas (Höffler & Leutner, 2011; Hsi, Linn, & Bell, 1997; Lord, 1987; Mayer & Sims, 1994; Siemankowski & MacKnight, 1971), and, more specifically, when learning from science text (Jaeger, Taylor, & Wiley, 2016; Narayanan & Hegarty, 2002; Sanchez, 2012; Sanchez & Wiley, 2010). Split-half reliability (Spearman-Brown coefficient) on this measure was .87 and Cronbach's alpha was .85 in this sample.

All students completed a paper-and-pencil final survey. This survey asked them to provide ratings on scales of 1 to 10, reporting how interesting they found the text to read, ranging from 1 (*not at all interesting*) to 10 (*very interesting*), as well as how hard they tried to learn about plate tectonics, ranging from 1 (*not at all*) to 10 (*very much*). The final survey also asked participants to report basic demographic information, including gender, age, whether they are bilingual, and the number of science courses taken. There were no differences in these variables across conditions.

Procedure. The experimental sessions were run in groups of one to three participants with all tasks being completed on paper, and each participant seated in a separate cubicle. The participants were randomly assigned to one of six treatment groups, with all participants within a session receiving the same activity treatment. However, text condition was assigned randomly within sessions such that within a given session, some participants may have read the base text and some may have read the base-plus-seductive text, however all participants in that session completed the same learning activity (sketching, summarizing, or silent thinking).

Participants first completed the paper folding task; instructions were read aloud and participants were given 3 min for each set of items. Next participants were given the instructional booklets and were told to follow along as the experimenter read the instructions aloud. Before beginning the reading task, participants answered the prior knowledge self-report item. Once participants were ready to begin the reading and learning activities, the experimenter instructed them that they would be given 2 min to work on each page. If they finished a page before the 2 min were up they were not to move on to the next page until instructed to do so. As such,

everyone was given 2 min to read each section and 2 min to complete each learning activity. Participants were given 2 min to read the first section, then 2 min to complete the learning activity (sketch, summary, think), then 2 min to read the next section, then 2 min to complete the next learning activity, and so on until all five sections of text were read and all five learning activities were completed. None of the participants were unable to fully read the text passages or complete any of the learning activities within the allotted 2 min. When the instructional booklets were completed, the experimenter collected them and handed each participant a recall sheet. Instructions were read aloud and participants were given 7 min to recall as much as they could. Next, participants were given the multiple-choice test booklet and were instructed that they would be given 5 min and they should select an answer for each item. Finally, participants completed the demographic questionnaire, were debriefed, and then thanked for their participation.

Results

Coding. The recall coding scheme gave credit for the inclusion of 15 primary causal concepts that were present in the base text. These concepts were scored as either present or absent, with the conceptual recall score representing the total number of these concepts mentioned in each free-recall protocol. For secondary analysis of the recall task, the protocols were scored for the inclusion of the 12 seductive details and again, seductive recall score represented the total number of these concepts mentioned in each free-recall protocol. For the recall task, two independent raters scored all of the protocols with interrater reliability (Krippendorff's alpha) of .94 (Krippendorff, 2013) for core concepts and .89 for seductive details; any disagreements were resolved through discussion.

For the multiple-choice test, participants were given a total correct score out of 10 possible points. As with most tests used in a classroom context, the multiple-choice test was explicitly designed for coverage of many different parts of the to-be-learned information, rather than to test for understanding of a single idea multiple times. In general, when Cronbach's alpha has been reported for inference tests, tests based on three or fewer passages (with approximately 16 sentences per passage) often have reliabilities in the .5 to .6 range. The Cronbach's alpha for the 10-item test used in this study was .51. Instead of using internal reliability as a basis for evaluation, reliability has been demonstrated through the relation of inference test performance to other measures of comprehension (Griffin, Wiley, & Thiede, 2008; Sanchez & Wiley, 2006, 2010). Thus, importantly, the multiple-choice test showed a

significant correlation with conceptual recall, $r = .46$, $p < .001$, suggesting they were both capturing aspects of student understanding about plate tectonics.

Previous research has indicated that the quality of student-generated sketches is correlated with learning and that sketch quality can be more predictive of learning than summary quality (Scheiter et al., 2017). Therefore, participants' sketches and summaries were also coded for the inclusion of the primary causal concepts that were present in the base text. For the summaries and sketches depicting/describing the plate interactions that cause stratovolcanoes, 5 points were possible. Five points were also possible for the summaries and sketches depicting/describing island arcs, recurring earthquakes, and the creation of ridges and new crust. The summaries and sketches for the formation of mountain ranges were only worth a total of 4 points because there was less content in the text on this topic. In sum, the highest possible score for the summaries or sketches was 24 (see Appendix B for examples of high- and low-scoring sketches). Interrater reliability on the five sketches and five summaries was adequate (Krippendorff's $\alpha = .83-.92$) and disagreements were resolved through discussion.

Conceptual recall. Given that one main goal of this experiment was to reduce the seductive details effect, a preliminary question was whether the seductive details effect on recall was evident in this experiment. A 2 (text condition: base-only, base-plus-seductive) $\times$ 3 (activity condition: sketch, summary, think) between-subjects analysis of covariance (ANCOVA) controlling for self-reported plate tectonics knowledge indicated that there was a significant main effect for text condition such that students in the base-only group recalled significantly more core concepts than students in the base-plus-seductive text group, $F(1, 111) = 6.09$, $p < .02$, $\eta_p^2 = .05$ (see Table 2). These results are consistent with the hypothesis that seductive details interfere with students' recall of important information and replicate findings on the seductive details effect (e.g., Garner et al., 1989; Harp & Mayer, 1997, 1998; Lehman et al., 2007). A further goal was to investigate whether sketching would serve as an effective tool for reducing the impact of seductive details on recall. Results indicated no main effect for activity condition and no interaction, $F_s < 1.35$, suggesting that there was no difference in recall as a function of whether participants sketched, wrote summaries, or thought silently.

Seductive recall. An additional question for this experiment was how the learning activities impacted recall of seductive information. Because only half of the participants received the base-plus-seductive text, a one-way ANCOVA controlling for self-reported plate tectonics knowledge was conducted on this subsample to investigate the impact of learning activity on recall of

Table 2
Mean Conceptual and Seductive Recall as a Function of Text and Activity Condition for Experiment 1

Recall	Base-only			Base-plus-seductive		
	Sketch	Summary	Think	Sketch	Summary	Think
Conceptual recall	4.28 (2.97)	3.83 (1.86)	4.68 (2.38)	3.21 (1.96)	3.59 (1.44)	3.21 (2.51)
Seductive recall	—	—	—	.47 (.84)	1.00 (1.23)	1.95 (1.43)

Note. Standard deviations are presented in parentheses. There are no values for seductive recall in the base-text condition because these participants were not exposed to seductive text.

seductive concepts. Results of this test revealed a significant effect, $F(2, 57) = 7.56, p < .001, \eta_p^2 = .21$ (see Table 2). Follow-up Bonferroni adjusted pairwise comparisons indicated that participants in the sketching group ($p < .001$) and the summary group ($p = .04$) recalled significantly fewer seductive details than participants in the silent think group. Although the pattern of recall shows the lowest seductive recall in the sketching group, there was no significant difference in seductive recall between the sketch and summary groups ($p = .47$).

Multiple choice. Beyond the question of how the presence of seductive information and the various learning activities would impact recall of conceptual information, we considered how these factors would impact comprehension as measured by the multiple-choice test. To address this question, another 2 (text condition) $\times$ 3 (activity condition) between-subjects ANCOVA was conducted with performance on the multiple-choice test as the dependent variable.² This analysis revealed no main effect for activity condition, $F < 1$, no effect for text condition, $F(1, 111) = 2.27, p = .13$, and no interaction, $F < 1$. Although this test did not reveal any significant effects, the pattern of means is similar to those for concept recall such that participants who received the base-plus-seductive text ($M = 3.87, SD = 1.74$) performed less well on the multiple-choice test than participants in the base-only text group ($M = 4.47, SD = 2.20$).

Spatial skills. A regression model including text condition, activity condition, and paper folding scores significantly predicted conceptual recall, $F(3, 114) = 8.81, p < .001$. Consistent with the analysis presented above, there was no effect of activity condition on conceptual recall ($\beta = .02, t < 1, ns$), but there was an effect of text condition such that students in the base-text group recalled more core concepts than students in the base-plus-seductive text group ($\beta = .23, t = 2.71, p < .01$). In addition, there was a significant independent effect for paper folding scores ($\beta = .36, t = 4.32, p < .001$), such that students with higher paper folding scores recalled more core concepts. Another regression model including text condition, activity condition, and paper folding score significantly predicted performance on the multiple-choice test, $F(3, 114) = 3.87, p = .01$. This analysis revealed no main effect for activity condition ($\beta = .02, t < 1, ns$) and a marginal effect for text condition ($\beta = .15, t = 1.63, p = .11$). There was however, a significant effect for paper folding score ($\beta = .26, t = 2.96, p < .01$) again, indicating that students with higher paper folding scores performed better on the multiple-choice test.

Activity quality. Mean activity quality score for the sketching group was 10.75 ($SD = 5.86$) and for the summary group was 11.55 ($SD = 4.05$), which did not significantly differ, $t < 1$. A correlation analysis revealed that sketching quality score was significantly correlated with conceptual recall, $r = .78, p < .001$, and with multiple choice performance, $r = .73, p < .001$. For the summary condition, summary quality was also significantly correlated with conceptual recall, $r = .60, p < .01$, and with multiple-choice performance, $r = .53, p < .001$. Finally, a correlation analysis also revealed that paper folding score significantly correlated with activity quality for the sketching group, $r = .49, p < .001$, but not for the summary group, $r = .10, ns$. A Fisher's exact difference test (one-tailed) indicated that these two correlations did significantly differ, $z = 1.88, p = .03$.

Discussion

Experiment 1 investigated the effectiveness of sketching and summarizing for overcoming the seductive details effect. Consistent with existing literature, participants in the seductive text group performed worse than participants in the base-only text group, on the recall task. There were, however, no significant differences in performance on the comprehension test as a function of text condition. Contrary to prior literature demonstrating a learning benefit for generating sketches from science text, the current study found no differences in recall or comprehension as a function of activity condition. Also, contrary to the primary hypotheses for this experiment, there was no interaction between text condition and activity condition on recall or comprehension. However, when looking at recall of seductive information, results demonstrated that students in the sketching group recalled the fewest seductive details. This aligns with the hypothesis that sketching can benefit expository science text comprehension by constraining attention and forcing the reader to select and organize relevant information. Further, the impact of activity quality on recall and comprehension was investigated. Results indicated that higher quality sketches and higher quality summaries were related to better recall and comprehension and aligns with the Prognostic Drawing Principle (Schwamborn et al., 2010). Interestingly, results also indicated that performance on the paper folding task was correlated with sketching quality, but not with summary quality. This suggests that although generating either a sketch or a summary of high quality are related to better recall and comprehension, generating a high-quality sketch may rely at least in part on spatial thinking skills whereas generating a high-quality summary may not. This result is consistent with the idea that sketching focuses readers on spatial information, but further suggests that individuals with low spatial skills may need additional support to effectively create a mental model of that spatial information.

Experiment 2

Contrary to the hypothesis that sketching would improve recall and comprehension and thus reduce the impact of seductive details, Experiment 1 found no overall effect of activity condition on recall or comprehension. Therefore, the primary goal of Experiment 2 was to improve the impact of the sketching condition. Prior research on sketching to learn from science text has indicated that feedback may be necessary (Fan, 2015; Van Meter & Garner, 2005). Van Meter (2001) suggested that sketching exercises need to be supported to be effective. Specifically, in her study students were given a text passage about the human nervous system. In the most supported condition, students read the passage and inspected two illustrations before creating their own sketch. After completing their sketch, they answered questions that required them to compare their sketch to the provided illustrations. In the second condition students generated sketches and were allowed to edit their sketches based on the provided illustrations but were not

² The effects of text and activity condition were also examined for each subtype of multiple-choice item. Two 2×3 ANCOVAs were conducted and revealed no significant effects for performance on the memory or inference items. Additionally, a one sample t test testing overall multiple choice score against chance (25%) demonstrated that performance was significantly above chance, $t(117) = 9.04, p < .001$.

explicitly instructed to compare them. In the third condition, students read the text and generated sketches but were not shown the illustrations, and in the final condition students read the text and were shown the illustrations but did not sketch. Results from this study indicated that students who were prompted to compare their sketch to the provided illustrations scored higher on the free recall. One hypothesis for this result is that the additional feedback provided in the most supported condition may have supported the correction of spatial errors (Gagnier, Atit, Ormand, & Shipley, 2017).

Thus, in Experiment 2 a feedback manipulation was added with the idea that having students compare their sketches and summaries to ideal sketches and summaries would result in better encoding of the material and better comprehension. An additional prediction was that the providing feedback would reduce the correlation between paper folding and sketch quality because it would help to direct low spatial individuals to the important spatial information they should include in their mental models. Further, prior research has indicated that sketching is especially beneficial for the development of the mental model representations as compared to rote memory or textbase representations (Alesandrini, 1981; Cromley et al., 2013; Gobert, 2005). Based on this idea, a short-answer comprehension measure was added in an attempt to further identify comprehension differences across activity conditions.

Method

Participants and design. One-hundred thirty-two students (100 female; $M = 20.07$ years, $SD = 2.28$) from a university in the northeastern part of the United States participated voluntarily in the experiment in exchange for course credit. Data from nine participants had to be excluded from data analysis; three participants were missing data due to an experimenter error, one participant was a nonnative English speaker, one did not comply with task instructions, and four reported having previously taken introductory geology. This resulted in a final sample of 123 participants.³ The design was the same as that in Experiment 1. There were 59 participants in the base-only text group and 64 in the base-plus-seductive text group. In addition, 41 participants were in the sketching activity group, 43 in the summary group, and 39 in the think group. See Table 3 for descriptive statistics. The study was approved by the ethics committee of the university where the study was conducted, and the study followed standards for ethical treatment of human subjects.

Materials. The text, activity pages, demographics survey, paper folding test, recall task, and multiple-choice test were all identical to those used in Experiment 1. The primary difference for Experiment 2 was the addition of a feedback portion following each activity. After completing each learning activity (sketch or summary), participants were given feedback in the form of a correct sketch or summary and asked to provide an explanation of how their sketch or summary differed from the correct sketch or summary (see Appendix C for examples of a correct sketch and summary). More specifically, for each of the five activities, participants were shown an ideal sketch or summary and were asked to explain how their artifact differed from the ideal one provided. For example, after participants wrote a summary describing, or created a sketch depicting, the plate interaction that causes strato-

volcanoes to form, they were then given an ideal summary or sketch about that topic and were told:

Here is a sketch/summary of the plate interaction that causes strato-volcanoes to form. This sketch/summary is from a previous student in the study who did a good job of including all the important information. Please take a minute to compare your sketch/summary to this ideal sketch/summary. In the space below the line please indicate how your sketch/summary differs from this ideal sketch/summary.

Participants in the silent think group were not provided with an ideal sketch or summary, but rather were asked to report what they thought about during the 2-min silent think activity. Specifically, they were told:

You just spent two minutes reviewing in your mind what you just read. On this sheet please take a minute to describe what you were thinking about during the silent think activity. If you were thinking about things related to what you just read, please describe. If you were not thinking about things related to what you just read, please indicate "personal matters" but you do not need to provide further description.

The other primary difference in Experiment 2 was the addition of five short-answer application questions. These questions were intended to serve as an additional measure of comprehension as well as an opportunity for participants to demonstrate their ability to apply their understanding of plate tectonics to more specific scenarios. Three of the items required students to use a diagram depicting some aspect of plate tectonics to answer the question. The five short-answer application questions were as follows: (a) "At which point in the diagram would the oldest oceanic crust be found, explain why." (b) "Why is it possible for earthquakes to occur along all types of plate boundaries?" (c) "Suppose you see a string of mountain peaks, but none of them are volcanic. How could this be?" (d) "Volcanoes and earthquakes are more likely to occur near Seattle, Washington than Atlanta, Georgia. Using the map below, explain why this is the case?" (e) "Based off the movement (indicated by the arrows) of the Pacific Plate and North American Plate in the image below, where will Los Angeles be located in the future? Explain why."

Procedure. The procedure matched that of Experiment 1, except that after each learning activity was completed the feedback/explain task was completed. For participants in the sketching group, they were shown a correct sketch and then were given 1 min to explain how their sketch differed from the provided correct sketch. Similarly, participants in the summary group were given a correct summary to look over and were given 1 min to explain how theirs differed. In the silent think group participants were instructed to spend 1 min reporting what they thought about during the silent think activity. Participants completed this feedback and explanation task for all 5 activities. After completing the reading and learning activities all participants were given 7 min to recall

³ Again, to ensure that there were no differences in attrition associated with condition, a chi-square test examining the number of participants excluded as a function of text condition and activity condition was conducted. There was no difference in attrition across conditions, $\chi^2(2, n = 9) = .23, p = .89$, when looking at all excluded participants, and no difference when looking only at participants with missing data or who did not follow task instructions, $\chi^2(2, n = 5) = .83, p = .66$. This suggests that the activity or text manipulations did not lead to increased attrition in any one condition.

Table 3
Descriptive Statistics for Experiment 2

Variable	Base-only			Base-plus-seductive		
	Sketch	Summary	Think	Sketch	Summary	Think
<i>n</i>	20	19	19	20	24	20
Age	19.52 (1.40)	19.33 (1.94)	20.56 (1.25)	20.26 (2.18)	20.33 (2.39)	20.53 (3.88)
Gender (% female)	.76	.74	.74	.70	.75	.80
Year in school (1–4)	2.14 (.96)	2.00 (1.09)	2.72 (.96)	2.53 (1.07)	2.32 (1.25)	2.26 (.87)
Science courses	2.75 (2.17)	2.44 (2.04)	3.56 (4.94)	2.05 (2.72)	3.23 (2.87)	2.32 (2.87)
Plate know (1–6)	2.48 (1.50)	2.44 (1.50)	2.22 (1.70)	3.11 (1.66)	1.77 (1.27)	2.42 (1.31)
Interest	4.67 (2.52)	5.28 (1.81)	3.94 (2.55)	5.05 (1.84)	4.27 (1.93)	3.89 (2.40)
Try	6.52 (2.16)	6.22 (2.02)	5.89 (2.25)	5.11 (1.97)	6.23 (2.43)	5.05 (2.24)
Recall word count	115.43 (43.87)	110.11 (41.22)	87.58 (26.86)	101.20 (37.99)	98.71 (40.62)	113.55 (34.71)

Note. Year in school: 1 = freshman, 2 = sophomore, 3 = junior, 4 = senior; Science courses = average number of college science courses taken; Plate know = average self-reported knowledge of plate tectonics on a 1–6 Likert scale; Interest = how interesting they found the text on a 1–10 Likert scale; Try = how hard they tried to learn the material on a 1–10 Likert scale.

what they could from the text. Then, they were given the test packet, which included the 10 multiple-choice items and the 5 short-answer items, and were given 10 min to complete the items.

Results

Coding. The scoring for the multiple-choice test was the same as in Experiment 1; Cronbach's alpha was .58 and importantly, performance on the multiple-choice test showed high correlations with conceptual recall, $r = .67$, $p < .001$, and short-answer performance, $r = .56$, $p < .001$, suggesting they were all capturing aspects of student understanding about plate tectonics. Scoring for the paper folding test was also the same as in Experiment 1; split-half reliability (Spearman-Brown coefficient) on this measure was .75 and Cronbach's alpha was .72 in this sample. The recall coding scheme was also the same as in Experiment 1. For the recall task, two independent raters scored all of the protocols with interrater reliability (Krippendorff's alpha) of .94 for core concepts and .93 for seductive details. For the short-answer test, each of the five items were scored by assigning 1 point for each acceptable answer. Acceptable answers were based on predetermined correct responses to each question. For Question 1, 2 points were possible; 1 point for selecting the correct location and 1 point for indicating that crust gets progressively older as it moves away from the ridge. For Question 2, 2 points were possible; 1 point for mentioning that earthquakes are caused by plate movement and 1 point for mentioning that all types of plate interactions involve movement. For Question 3, 3 points were possible; 1 point for mentioning the convergence of two continental plates, 1 point for mentioning that

it is a plate boundary where no subduction is occurring, and 1 point for mentioning that there is no magma formation. For Question 4, 2 points were possible; 1 point for mentioning that volcanoes and earthquakes are more common near Seattle because it is near a plate boundary, and 1 point for mentioning that they are less likely near Atlanta because it is on a single plate and not near a plate boundary. Lastly, for Question 5, 2 points were possible; 1 point for mentioning that the plates are moving in opposite directions (or a transform boundary), and 1 point for saying that Los Angeles is going to end up further north (or near San Francisco). This coding scheme resulted in a total possible score of 12 points for the short-answer test. Two independent raters scored all of the short-answer responses with interrater reliability (Krippendorff's alpha) of .90. Participants' postreading activities were also coded for the inclusion of core concepts. Coding was conducted in the same manner as in Experiment 1 and reliability was high (Krippendorff's alphas = .83–.93).

Conceptual recall. A 2 (text condition: base-only, base-plus-seductive) $\times$ 3 (activity condition: sketch, summary, think) between-subjects analysis ANCOVA controlling for self-reported plate tectonics knowledge indicated that there was a significant main effect for text condition such that students in the base-only group recalled significantly more core concepts than students in the base-plus-seductive text group, $F(1, 116) = 8.59$, $p < .01$, $\eta_p^2 = .07$ (see Table 4), suggesting that seductive details interfere with students' recall of important information. There was also a main effect for activity condition, $F(2, 116) = 6.90$, $p < .001$, $\eta_p^2 = .11$. Follow-up Bonferroni adjusted pairwise comparisons

Table 4
Mean Conceptual and Seductive Recall as a Function of Text and Activity Condition for Experiment 2

Recall	Base-only			Base-plus-seductive		
	Sketch	Summary	Think	Sketch	Summary	Think
Conceptual recall	4.43 (2.93)	5.63 (2.94)	4.26 (2.38)	2.80 (1.96)	4.75 (2.99)	2.45 (2.11)
Seductive recall	—	—	—	.80 (1.24)	.75 (1.33)	1.95 (1.64)

Note. Standard deviations are presented in parentheses. There are no values for seductive recall in the base-text condition because these participants were not exposed to seductive text.

indicated that participants in the summary group recalled significantly more core concepts than participants in the silent think group, $p < .01$, and significantly more than participants in the sketching group, $p = .03$. There was no difference in conceptual recall between the sketching and silent think groups. Results of the overall ANCOVA also revealed no significant interaction between text condition and activity condition, $F < 1$.

Seductive recall. A one-way ANCOVA controlling for self-reported plate tectonics knowledge was conducted on participants in the seductive text group to investigate the impact of learning activity on recall of seductive concepts. Results of this test revealed a significant effect, $F(2, 60) = 4.77, p < .02, \eta_p^2 = .14$ (see Table 4). Follow-up Bonferroni adjusted pairwise comparisons indicated that participants in the sketching group recalled significantly fewer seductive details than participants in the silent think group, $p < .04$. Similarly, participants in the summary group recalled significantly fewer seductive details than participants in the silent think group, $p = .02$. There was no difference in seductive recall between the sketch and summary groups, $p = 1.00$.

Multiple choice. Another 2 (text condition) $\times$ 3 (activity condition) between-subjects ANCOVA was conducted with performance on the multiple-choice test as the dependent variable.⁴ This analysis revealed a significant main effect for text condition, $F(1, 116) = 7.76, p < .01, \eta_p^2 = .06$, such that participants in the base-text group ($M = 4.93, SD = 2.15$) performed better than participants in the base-plus-seductive-details group ($M = 3.91, SD = 2.06$). There was a marginal effect for activity condition, $F(2, 116) = 2.34, p = .10, \eta_p^2 = .04$, and no interaction, $F < 1$. Although there was no a significant main effect for activity condition, participants in the summary group ($M = 4.91, SD = 2.34$) performed the best on the multiple-choice test (Sketch: $M = 4.10, SD = 1.93$; Think: $M = 4.15, SD = 2.12$).

Short-Answer. A 2×3 ANCOVA was conducted with performance on the short-answer test as the dependent variable. Results demonstrated a main effect for text condition, $F(1, 116) = 4.95, p < .03, \eta_p^2 = .04$, such that participants in the base-text group scored higher than participants in the base-plus-seductive-text group. This analysis also revealed a marginal effect for activity condition, $F(2, 116) = 2.92, p = .06, \eta_p^2 = .05$, but no interaction, $F < 1$. Although the effect of activity condition did not reach significance, the pattern of means was that participants in the summary group ($M = 4.88, SD = 2.48$) scored higher than participants in either the sketching ($M = 3.78, SD = 2.52$) or silent think groups ($M = 4.05, SD = 2.37$).

Spatial skills. To investigate the role that spatial skills might play in Experiment 2, a regression model including text condition, activity condition, and paper folding scores was run and significantly predicted conceptual recall, $F(3, 119) = 14.88, p < .001$. Consistent with the analysis presented above, there was no effect of activity condition on conceptual recall ($\beta = -.07, t < 1, ns$), but there was an effect of text condition such that students in the base-text group recalled more core concepts than students in the base-plus-seductive text group ($\beta = .17, t = 2.14, p < .05$). In addition, there was a significant independent effect for paper folding scores ($\beta = .47, t = 5.90, p < .001$), such that students with higher paper folding scores recalled more core concepts. Another regression model including text condition, activity condition and paper folding score significantly predicted performance

on the multiple-choice test, $F(3, 119) = 7.82, p < .001$. This analysis revealed no main effect for activity condition ($\beta = .02, t < 1, ns$) and a significant effect for text condition, $\beta = .19, t = 2.27, p < .03$. There was also a significant effect for paper folding score ($\beta = .33, t = 3.91, p < .001$) again indicating that students with higher paper folding scores performed better on the multiple-choice test. The same regression analysis was also conducted with short-answer performance as the dependent measure and revealed a significant overall model, $F(3, 119) = 12.62, p < .001$. Activity condition was not a significant predictor ($\beta = .05, t < 1$) in the model, text condition was marginal ($\beta = .13, t = 1.59, p = .11$), and there was a significant effect for paper folding score ($\beta = .45, t = 5.63, p < .001$).

Activity quality. Mean activity quality score for the sketching group was 10.12 ($SD = 4.88$) and for the summary group was 12.65 ($SD = 5.11$), which did significantly differ, $t(82) = 2.32, p < .03, d = .51$. A correlation analysis revealed that sketching quality score was significantly correlated with conceptual recall, $r = .70, p < .001$, with multiple choice performance, $r = .48, p < .001$, and with short answer performance, $r = .70, p < .001$. For the summary group, summary quality was also significantly correlated with conceptual recall, $r = .75, p < .001$, with multiple-choice performance, $r = .67, p < .001$, and with short answer performance, $r = .59, p < .001$. Finally, a correlation analysis also revealed that paper folding score significantly correlated with activity quality in the sketching group, $r = .60, p < .001$, and in the summary group, $r = .33, p = .03$. A Fisher's exact difference test (one-tailed) indicated that these two correlations marginally differed, $z = 1.88, p = .06$. For the silent-think group, participants were asked to report and write down what they thought about during each 2-min silent think period. These reports were transcribed and coded using the same coding scheme as was used for scoring summary and sketch activity quality. Although prompted to think about the information they had just read, participants in the silent think group only reported thinking about an average of 4.97 core concepts ($SD = 4.49$). This was significantly lower than the activity quality scores for both sketching and summary groups, $ps < .001$. In addition, the activity score in the silent think group was only correlated with recall, $r = .50, p < .001$, but not with multiple-choice, $r = .17, ns$, or short answer, $r = .19, ns$.

For all activity conditions, participants were also given a score for whether or not they included seductive information in their activities. For each activity, participants received a 1 if they included any seductive information or a 0 if they included none. This resulted in a total possible score of 5 (1 point for each of the five activities). Reliability on this coding was high (two raters, Krippendorff's $\alpha = .90$). A one-way analysis of variance looking at the inclusion of seductive details in the postreading activities as a function of activity condition was conducted and revealed a significant effect, $F(2, 61) = 13.49, p < .001, d = .89$. A follow-up Bonferroni test revealed that students in the silent think

⁴ The effects of text and activity condition were again also examined for each subtype of multiple-choice item. Two 2×3 ANCOVAs were conducted and only the main effect for text type was significant for performance on the memory and inference items. In addition, a one sample t test testing overall multiple choice score against chance (25%) demonstrated that performance was significantly above chance, $t(122) = 9.76, p < .001$.

group reported significantly more seductive details in their postreading activities than participants in the sketching or summary groups ($ps < .001$), which did not differ from each other.

Explain task quality. In the sketching and summary groups, participants were prompted after completing each activity to compare their sketch or summary to a correct sketch or summary and explain how theirs differed from the correct one. These explanations were coded using the same activity quality coding scheme. For example, if a participant included three core concepts in their initial summary or sketch and then identified the remaining two core concepts as missing from their summary or sketch in the explanation task, they received 2 points. Again, two coders coded all explanation tasks and reliability was high (Krippendorff's $\alpha = .95$). An independent samples t test indicated that the number of concepts participants included in their postactivity explanations significantly differed as a function of activity condition, $t(82) = 2.37$, $p = .02$, $d = .52$. Participants in the summary group ($M = 2.09$, $SD = 2.39$) identified more core concepts missing from their postreading activities in their explanations than participants in the sketching group ($M = 1.12$, $SD = 1.12$). In other words, despite having higher activity quality scores and thus, less opportunity to earn additional points during the explanation task, students in the summary group still benefitted more from the feedback and were able to identify any concepts they missed in their initial activity.

Discussion

The goal of Experiment 2 was to further investigate the effectiveness of sketching and summarizing for learning from expository science text and for overcoming the seductive details effect. Because prior work on sketching has indicated that external support is critical, Experiment 2 included a feedback phase where students were instructed to compare their sketches and summaries with correct versions and describe how theirs differed. The results of Experiment 2 once again demonstrated a seductive details effect such that students who read the base-only text showed better recall for core concepts, better performance on the multiple-choice test and better performance on the short answer test than students who read the base-plus-seductive text. Contrary to initial hypotheses, sketching did not reduce the impact of seductive details and in fact students in the summary group tended to show better performance than students in the sketching group. More specifically, students in the summary group had significantly better recall of core concepts and marginally better short answer performance compared to students in the sketching or silent think groups, but there was again no impact of activity condition on multiple-choice performance. However, when looking at recall for seductive details, results revealed that both sketching and summarizing resulted in reduced irrelevant detail recall compared to students in the silent think group. These results indicate that students in the summary group were more able to take advantage of the feedback that was provided by presenting them with correct sketches or summaries. In line with this interpretation, analyses of the explanations students provided regarding their sketches and summaries demonstrated that students in the summary group identified more core concepts missing from their activities than students in the sketching group. The impact of activity quality on recall and comprehension was also investigated. Results again indicated that higher quality sketches and higher quality summaries were related to better recall

and comprehension and that performance on the paper folding task was more correlated with sketching quality than with summary quality.

General Discussion

Overall, findings across both experiments showed that the presence of seductive details in an expository science text does lead to reduced recall of core conceptual information and that the presence of seductive details can lead to reduced comprehension, as measured by the multiple choice and short answer tests in Experiment 2. Together, these results add to the robust literature suggesting that, although interesting and irrelevant information in a text is meant to increase reader engagement and enjoyment, it can harm learning (Rey, 2012).

Contrary to hypotheses, there was no overall benefit for sketching; across both studies, prompting students to generate sketches did not improve recall or comprehension compared to generating summaries or thinking silently. This result is at odds with the growing literature that has demonstrated learning gains from student-generated sketching activities (with medium effect sizes) when reading expository science text (e.g., Gobert & Clement, 1999; Leopold & Leutner, 2012; Schmeck et al., 2014; Schwamborn et al., 2010; Van Meter, 2001; Van Meter et al., 2006). Although no effect of sketching was seen for conceptual recall or comprehension, there was an effect for recall of seductive details; across both studies students in the sketching group recalled the fewest seductive details. This result is in line with the initial hypothesis that sketching would be useful for directing attention away from the seductive details. However, the hypothesis that sketching would reduce or eliminate the seductive details effect is not fully supported by the data. It was expected that directing attention away from the seductive details would in turn focus attention more on the spatial and conceptual information and foster the development of a richer and more accurate mental model, but the lack of a gain in conceptual recall or comprehension suggest that this was not the case. This result is also at odds with the hypothesis that attention to seductive details causes the decrement in performance typically reported in seductive details research. It is possible that the sketching interfered with consolidation of the seductive details, but not attention to them in the moment of reading.

Referring back to the GTDC model, it seems that the sketching activity was useful for the selection phase (as indexed by lack of seductive details in the sketches) but did not offer enough support for the integration phase. In particular, sketching was successful for directing attention to the relevant spatial information, but it was not successful for helping students to build connections between elements in the verbal and visual representations. This interpretation is well aligned with the growing body of research that has only found positive effects for sketching if additional instructional support is provided (Leutner & Schmeck, 2014). Previous research indicating positive effects for sketching have provided support in a variety of ways including having students complete learning-strategy training prior to reading the target text (Leopold & Leutner, 2012), providing guidance about the key pictorial elements to be included in the sketches (Alesandrini, 1981; Schmeck et al., 2014; Schwamborn et al., 2010), or by providing a partially complete series of sketches where students fill in the missing stages

(Britton & Wandersee, 1997). The idea behind the addition of these instructional supports is that they help to reduce the extraneous cognitive load generated by the act of sketching. In cases where students are asked to generate sketches without instructional support, the act of sketching can consume a large amount of cognitive resources, leaving an insufficient amount of cognitive resources for the integration phase (Leutner & Schmeck, 2014). In both the present experiment and the study by Scheiter et al. (2017), no instructional support or pictorial elements were provided and neither study finds a significant benefit for sketching when compared to generating summaries.

The goal of Experiment 2 was to provide students with additional support in the form of a feedback phase in which they were able to compare their sketch or summary with the correct version. Despite the addition of this opportunity for alignment, model integration and updating, the student-generated sketching activity still did not lead to improved conceptual recall or comprehension as compared to generating summaries or thinking silently. However, students in the summary group were able to take advantage of the feedback activity and improve conceptual recall and comprehension while simultaneously reducing the recall of seductive details. Van Meter (2001) found that simply giving participants a correct illustration and telling them to compare it to their sketch was not as effective as giving participants the correct illustrations and providing them with specific questions that required them to compare their sketch to the provided illustrations. These results indicate that simply showing participants the correct sketch and asking them to explain how theirs differed still did not offer enough support. Follow-up studies should investigate whether providing participants with specific questions that guide them to attend to each of the major points in each sketch could improve comprehension and reduce the seductive details effect.

A possible explanation for the lack of a sketching effect, even in Experiment 2, is that creating a high-quality sketch may rely more heavily on spatial thinking skills than creating a high-quality summary. Activity quality was positively correlated with conceptual recall and comprehension for both sketching and summarizing, demonstrating support for the Prognostic Drawing Principle (Schwamborn et al., 2010). In addition, in both groups activity quality was significantly correlated with paper folding score, however the correlation between sketch quality and paper folding was higher than the correlation between summary quality and paper folding. This suggests that generating a sketch relies heavily on spatial thinking skills and that simply providing students with an opportunity to compare their sketch to an ideal sketch is not supportive enough, perhaps especially for low spatial students. If low spatial individuals are already struggling to develop an accurate internal mental model and translate that into an external representation, they may not have enough cognitive resources to fully take advantage of the opportunity for comparison and mental model updating. On the other hand, generating a summary may rely less on spatial thinking skills, therefore allowing low spatial individuals to benefit more from the corrective feedback. Again, a follow-up study in which more specific questions prompting participants to consider each important component of a sketch may be necessary, especially for low spatial students, for making the correct mappings and integration.

Results from Jaeger, Taylor, and Wiley (2016) offer support for this prediction and found that presenting analogies in an inter-

leaved manner throughout a science text (as opposed to presenting the analogy at the beginning of the text) was particularly beneficial for low spatial students. They suggested that the interleaving helped low spatial individuals by circumventing the need for them to generate the mappings between the analogy and the science phenomenon on their own. Future research that takes a more process-focused approach, such as by collecting measures of reading time or eyetracking traces, would be useful for providing more direct evidence for determining how different instructional manipulations impact attention to seductive details and how sketching tasks alter attention patterns during reading. Specifically, process measures could help clarify why there was a decrease in recall of seductive details, but not an increase in conceptual recall or comprehension.

Another explanation for the lack of an overall effect of sketching, even when given the opportunity to identify and explain what the initial sketch was missing, could stem from the fact that some of the important concepts in the text were not readily "sketchable." For example, when considering the plate interactions that cause mountain ranges to form or recurring earthquakes to form, a key concept is that subduction (the process of one plate being pushed beneath the other and eventually melted in the mantle) does not occur. Naturally, the lack of something may be difficult to represent spatially, and may in fact be better represented verbally. If a sketch did represent the lack of subduction in these cases, it was through a verbal label or caption, rather than through the sketch itself. Therefore, future research should be careful to only assess the effectiveness of sketching when all relevant concepts in a text can be represented by nonverbal representations. Further, though it is likely that all students have experience with writing summaries, many may not have experience with using sketching as a tool for learning from science text. Thus, students may not understand what it means to generate a sketch, what a good sketch should include, and how to represent certain types of information in a visual manner. In this case, a prereading sketching tutorial could alleviate this potential source of overload.

Although the present set of studies did not find an effect for sketching beyond that of generating summaries or thinking silently, it is important to note that many studies that have found significant learning benefits for sketching have used read-only control groups (e.g., Schmeck et al., 2014; Schwamborn et al., 2010; Van Meter, 2001). That is, in these studies the generative activity of sketching is only significant when compared to the conditions where students generate nothing; the primary learning difference is between sketching and the control group and no difference is found between sketching and the other generative conditions. When sketching is compared to other generative learning activities such as writing summaries or self-explaining, the impact of sketching on comprehension becomes less clear. For instance, Cromley et al. (2013) compared a sketching activity to a verbal elaboration activity and a self-explanation activity. Results showed that verbal elaboration and self-explanation improved inference test performance, but sketching only improved memory performance. Further, they found that only students who completed the self-explanation activity were able to transfer their learning to comprehension tasks from another science discipline. Ultimately, they concluded that the sketching activities, at least as they had designed them, were not as effective as self-explanation

tasks in isolation. Scheiter et al. (2017) compared sketching and summarization and found no difference in recall between the two groups. Similarly, Gobert and Clement (1999) suggested that they found a benefit for sketching as compared to summarization, but none of their results were statistically significant (all were marginal).

Positive results obtained for the summary group in Experiment 2 do align with prior research. Wang, Sundararajan, Adesope, and Ardasheva (2016) found that note taking reduced the seductive details effect. In contrast to the present experiment, they used seductive images rather than seductive text and suggested that the act of taking notes helped to direct readers' attention away from the seductive images. The difference in stimuli is critical because it may be more challenging to represent the content of seductive images in a note taking task, hence its effectiveness in their study. On the other hand, it is less challenging to include seductive text information in notes or summaries, which could be why the summary manipulation was not effective in Experiment 1. However, in Experiment 2 when students were made aware of the core concepts they should be focusing on, the summary group not only became effective for reducing the seductive details effect, but also became more effective than sketching or thinking silently. Recently, Chang and Choi (2014) conducted a study where both seductive text and images were present and showed with eyetracking that increased attention to seductive text was the major determinant of poor recall and comprehension, not attention to seductive images.

Another open question is how robust the effects of these different generation activities are and if sketching may be more effective for reducing the seductive details effect with a delay. If sketching improves learning from expository science text by helping students to develop a more robust mental model, as opposed to merely a textbase representation, then it could be hypothesized that over time, the textbase representation will fade whereas the situation model will remain. Gobert (2000) found that sketching resulted in better performance on inference questions, whereas summarizing resulted in better performance on memory questions. This finding indicates that sketching selectively supports the development of a "higher-level" mental model, one that includes complex spatial relations, which is important for deeper inferential reasoning (Kintsch, 1994) and may be more robust to decay. This may be especially true for text with seductive details because, as the results from the present experiment demonstrated, sketching focuses attention away from the seductive details but summarizing may not.

In conclusion, this set of studies indicates that seductive details, although interesting and engaging, are harmful for comprehension and their use and placement should be carefully considered when formulating expository text. Although there is support in the literature for sketching as a method for improving learning from text, the conditions under which it is used and the instructional supports that accompany it should also be considered carefully. If sketching is to be an effective method for decreasing the negative impact of seductive details on text comprehension, then students may need more explicit instruction or support to not only direct attention to the important content, but also to foster mental model construction.

References

- Ainsworth, S., Prain, V., & Tytler, R. (2011). Science education. Drawing to learn in science. *Science*, 333, 1096–1097. <http://dx.doi.org/10.1126/science.1204153>
- Alesandrini, K. L. (1981). Pictorial-verbal and analytic-holistic learning strategies in science learning. *Journal of Educational Psychology*, 73, 358–368. <http://dx.doi.org/10.1037/0022-0663.73.3.358>
- Britton, L. A., & Wandersee, J. H. (1997). Cutting up text to make moveable, magnetic diagrams: A way of teaching and assessing biological processes. *The American Biology Teacher*, 59, 288–291. <http://dx.doi.org/10.2307/4450310>
- Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor analytical studies*. New York, NY: Cambridge University Press. <http://dx.doi.org/10.1017/CBO9780511571312>
- Chang, Y., & Choi, S. (2014). Effects of seductive details evidenced by gaze duration. *Neurobiology of Learning and Memory*, 109, 131–138. <http://dx.doi.org/10.1016/j.nlm.2014.01.005>
- Chi, M. T. H. (2000). Self-explaining expository texts: The dual processes of generating inferences and repairing mental models. In R. Glaser (Ed.), *Advances in instructional psychology* (pp. 161–238). Mahwah, NJ: Erlbaum.
- Cromley, J. G., Bergey, B. W., Fitzhugh, S., Newcombe, N., Wills, T. W., Shipley, T. F., & Tanaka, J. C. (2013). Effects of three diagrams instruction methods on transfer of diagram comprehension skills: The critical role of inference while learning. *Learning and Instruction*, 26, 45–58. <http://dx.doi.org/10.1016/j.learninstruc.2013.01.003>
- Donnelly, C. M., & McDaniel, M. A. (1993). Use of analogy in learning scientific concepts. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 19, 975–987. <http://dx.doi.org/10.1037/0278-7393.19.4.975>
- Ekstrom, R. B., French, J. W., Harman, H. H., & Dermen, D. (1976). *Manual for kit of factor-referenced cognitive tests*. Princeton, NJ: Educational testing service.
- Fan, J. E. (2015). Drawing to learn: How producing graphical representations enhances scientific thinking. *Translational Issues in Psychological Science*, 1, 170–181. <http://dx.doi.org/10.1037/tps0000037>
- Gagnier, K. M., Atit, K., Ormand, C. J., & Shipley, T. F. (2017). Comprehending 3D diagrams: Sketching to support spatial reasoning. *Topics in Cognitive Science*, 9, 883–901. <http://dx.doi.org/10.1111/tops.12233>
- Garner, R., Brown, R., Sanders, S., & Menke, D. J. (1992). "Seductive details" and learning from text. In A. Renninger, S. Hidi, & A. Krapp (Eds.), *The role of interest in learning and development* (pp. 239–254). Mahwah, NJ: Erlbaum, Inc.
- 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, 41–57. http://dx.doi.org/10.1207/s1532690xc0601_2
- Gentner, D., & Stevens, A. L. (1983). *Mental models*. Hillside, NJ: Erlbaum.
- Gobert, J. D. (2000). A typology of causal models for plate tectonics: Inferential power and barriers to understanding. *International Journal of Science Education*, 22, 937–977. <http://dx.doi.org/10.1080/095006900416857>
- Gobert, J. D. (2005). The effects of different learning tasks on model-building in plate tectonics: Diagramming versus explaining. *Journal of Geoscience Education*, 53, 444–455. <http://dx.doi.org/10.5408/1089-9995-53.4.444>
- Gobert, J. D., & Clement, J. J. (1999). Effects of student-generated diagrams versus student-generated summaries on conceptual understanding of causal and dynamic knowledge in plate tectonics. *Journal of Research in Science Teaching*, 36, 39–53. [http://dx.doi.org/10.1002/\(SICI\)1098-2736\(199901\)36:1<39::AID-TEA4>3.0.CO;2-I](http://dx.doi.org/10.1002/(SICI)1098-2736(199901)36:1<39::AID-TEA4>3.0.CO;2-I)
- Goel, V. (1995). *Sketches of thought*. Cambridge, MA: MIT Press.

- Graesser, A. C., & Bertus, E. L. (1998). The construction of causal inferences while reading expository texts on science and technology. *Scientific Studies of Reading*, 2, 247–269. http://dx.doi.org/10.1207/s1532799xssr0203_4
- Graesser, A. C., Leon, J. A., & Otero, J. (2002). Introduction to the psychology of science text comprehension. In J. Otero, J. A. Leon, & A. C. Graesser (Eds.), *The psychology of science text comprehension* (pp. 1–15). New York, NY: Routledge.
- Griffin, T. D., Wiley, J., & Thiede, K. W. (2008). Individual differences, rereading, and self-explanation: Concurrent processing and cue validity as constraints on metacomprehension accuracy. *Memory & Cognition*, 36, 93–103. <http://dx.doi.org/10.3758/MC.36.1.93>
- Harp, S. F., & Mayer, R. E. (1997). The role of interest in learning from scientific text and illustrations: On the distinction between emotional interest and cognitive interest. *Journal of Educational Psychology*, 89, 92–102. <http://dx.doi.org/10.1037/0022-0663.89.1.92>
- 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, 414–434. <http://dx.doi.org/10.1037/0022-0663.90.3.414>
- Hegarty, M., & Just, M. A. (1993). Constructing mental models of machines from text and diagrams. *Journal of Memory and Language*, 32, 717–742. <http://dx.doi.org/10.1006/jmla.1993.1036>
- Heiser, J., & Tversky, B. (2006). Arrows in comprehending and producing mechanical diagrams. *Cognitive Science*, 30, 581–592. http://dx.doi.org/10.1207/s15516709cog0000_70
- Hidi, S. (1990). Interest and its contribution as a mental resource for learning. *Review of Educational Research*, 60, 549–571. <http://dx.doi.org/10.3102/00346543060004549>
- Höfler, T. N., & Leutner, D. (2011). The role of spatial ability in learning from instructional animations—Evidence for an ability-as-compensator hypothesis. *Computers in Human Behavior*, 27, 209–216. <http://dx.doi.org/10.1016/j.chb.2010.07.042>
- Hsi, S., Linn, M. C., & Bell, J. E. (1997). The role of spatial reasoning in engineering and the design of spatial instruction. *The Research Journal for Engineering Education*, 86, 151–158. <http://dx.doi.org/10.1002/j.2168-9830.1997.tb00278.x>
- Jaeger, A. J., Taylor, A. R., & Wiley, J. (2016). When, and for whom, analogies help: The role of spatial skills and interleaved presentation. *Journal of Educational Psychology*, 108, 1121–1139. <http://dx.doi.org/10.1037/edu0000121>
- Jee, B. D., Uttal, D. H., Gentner, D., Manduca, C., Shipley, T. F., Tikoff, B., . . . Sageman, B. (2010). Commentary: Analogical thinking in geoscience education. *Journal of Geoscience Education*, 58, 2–13. <http://dx.doi.org/10.5408/1.3544291>
- Johnson-Laird, P. N. (1983). *Mental models: Towards a cognitive science of language, inference, and consciousness*. Cambridge, MA: Harvard.
- Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review*, 95, 163–182. <http://dx.doi.org/10.1037/0033-295X.95.2.163>
- Kintsch, W. (1994). Text comprehension, memory, and learning. *American Psychologist*, 49, 294–303. <http://dx.doi.org/10.1037/0003-066X.49.4.294>
- Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. Cambridge, UK: Cambridge University Press.
- Kintsch, W., & van Dijk, T. A. (1978). Toward a model of text comprehension and production. *Psychological Review*, 85, 363–394. <http://dx.doi.org/10.1037/0033-295X.85.5.363>
- Kombartzky, U., Ploetzner, R., Schlag, S., & Metz, B. (2010). Developing and evaluating a strategy for learning from animations. *Learning and Instruction*, 20, 424–433. <http://dx.doi.org/10.1016/j.learninstruc.2009.05.002>
- Krippendorff, K. (2013). *Content analysis: An introduction to its methodology* (3rd ed.). Thousand Oaks, CA: Sage.
- Lehman, S., Schraw, G., McCrudden, M. T., & Hartley, K. (2007). Processing and recall of seductive details in scientific text. *Contemporary Educational Psychology*, 32, 569–587. <http://dx.doi.org/10.1016/j.cedpsych.2006.07.002>
- Leopold, C., & Leutner, D. (2012). Science text comprehension: Drawing, main idea selection, and summarizing as learning strategies. *Learning and Instruction*, 22, 16–26. <http://dx.doi.org/10.1016/j.learninstruc.2011.05.005>
- Leutner, D., & Schmeck, A. (2014). The generative drawing principle in multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 433–448). Cambridge, UK: Cambridge University Press. <http://dx.doi.org/10.1017/CBO9781139547369.022>
- Lord, T. R. (1987). A look at spatial abilities in undergraduate women science majors. *Journal of Research in Science Teaching*, 24, 757–767. <http://dx.doi.org/10.1002/tea.3660240808>
- Lowe, R., & Mason, L. (2017). Self-generated drawing: A help or hindrance to learning from animation? In R. Lowe & R. Ploetzner (Eds.), *Learning from dynamic visualization: Innovations in research and application* (pp. 309–331). Berlin, Germany: Springer. http://dx.doi.org/10.1007/978-3-319-56204-9_13
- Mason, L., Lowe, R., & Tornatora, M. C. (2013). Self-generated drawings for supporting comprehension of a complex animation. *Contemporary Educational Psychology*, 38, 211–224. <http://dx.doi.org/10.1016/j.cedpsych.2013.04.001>
- Mayer, R. E. (1989). Models for understanding. *Review of Educational Research*, 59, 43–64. <http://dx.doi.org/10.3102/00346543059001043>
- Mayer, R. E. (1993). Illustrations that instruct. In R. Glaser (Ed.), *Advances in instructional psychology* (pp. 253–284). Hillsdale, NJ: Erlbaum.
- Mayer, R. E. (Ed.). (2005). *Cambridge handbook of multimedia learning*. New York, NY: Cambridge University Press. <http://dx.doi.org/10.1017/CBO9780511816819>
- Mayer, R. E., Bove, W., Bryman, A., Mars, R., & Tapangco, L. (1996). When less is more: Meaningful learning from visual and verbal summaries of science textbook lessons. *Journal of Educational Psychology*, 88, 64–73. <http://dx.doi.org/10.1037/0022-0663.88.1.64>
- Mayer, R. E., Griffith, E., Jurkowitz, I. T. N., & Rothman, D. (2008). Increased interestingness of extraneous details in a multimedia science presentation leads to decreased learning. *Journal of Experimental Psychology: Applied*, 14, 329–339. <http://dx.doi.org/10.1037/a0013835>
- Mayer, R. E., & Sims, V. (1994). For whom is a picture worth a thousand words? Extensions of a dual-coding theory of multimedia learning. *Journal of Educational Psychology*, 86, 389–401. <http://dx.doi.org/10.1037/0022-0663.86.3.389>
- McNamara, D. S., Graesser, A. C., & Louwerse, M. M. (2012). Sources of text difficulty: Across genres and grades. In J. P. Sabatini, E. Albro, & T. O'Reilly (Eds.), *Measuring up: Advances in how we assess reading ability* (pp. 89–116). Lanham, MD: R & L. Education.
- McNamara, D. S., Kintsch, E., Songer, N. B., & Kintsch, W. (1996). Are good texts always better? Interactions of text coherence, background knowledge, and levels of understanding in learning from text. *Cognition and Instruction*, 14, 1–43. http://dx.doi.org/10.1207/s1532690xci1401_1
- Millis, K. K., & Graesser, A. C. (1994). The time-course of constructing knowledge-based inferences for scientific texts. *Journal of Memory and Language*, 33, 583–599. <http://dx.doi.org/10.1006/jmla.1994.1028>
- Morehead, K., Rhodes, M. G., & DeLozier, S. (2016). Instructor and student knowledge of study strategies. *Memory*, 24, 257–271. <http://dx.doi.org/10.1080/09658211.2014.1001992>
- Narayanan, N. H., & Hegarty, M. (2002). Multimedia design for communication of dynamic information. *International Journal of Human-Computer Studies*, 57, 279–315. <http://dx.doi.org/10.1006/ijhc.2002.1019>

- Park, B., Flowerday, T., & Brünken, R. (2015). Cognitive and affective effects of seductive details in multimedia learning. *Computers in Human Behavior*, 44, 267–278. <http://dx.doi.org/10.1016/j.chb.2014.10.061>
- Peshkam, A., Mensink, M. C., Putnam, A. L., & Rapp, D. N. (2011). Warning readers to avoid irrelevant information: When being vague might be valuable. *Contemporary Educational Psychology*, 36, 219–231. <http://dx.doi.org/10.1016/j.cedpsych.2010.10.006>
- Pozzer, L. L., & Roth, M. (2003). Prevalence, function, and structure of photographs in high school biology textbooks. *Journal of Research in Science Teaching*, 40, 1089–1114. <http://dx.doi.org/10.1002/tea.10122>
- Rey, G. D. (2012). A review of research and a meta-analysis of the seductive detail effect. *Educational Research Review*, 7, 216–237. <http://dx.doi.org/10.1016/j.edurev.2012.05.003>
- Sadoski, M. (2001). Resolving the effects of concreteness on interest, comprehension, and learning important ideas from text. *Educational Psychology Review*, 13, 263–279. <http://dx.doi.org/10.1023/A:1016675822931>
- Sanchez, C. A. (2012). Enhancing visuospatial performance through video game training to increase learning in visuospatial science domains. *Psychonomic Bulletin & Review*, 19, 58–65. <http://dx.doi.org/10.3758/s13423-011-0177-7>
- Sanchez, C. A., & Wiley, J. (2006). An examination of the seductive details effect in terms of working memory capacity. *Memory & Cognition*, 34, 344–355. <http://dx.doi.org/10.3758/BF03193412>
- Sanchez, C. A., & Wiley, J. (2010). Sex differences in science learning: Closing the gap through animations. *Learning and Individual Differences*, 20, 271–275. <http://dx.doi.org/10.1016/j.lindif.2010.01.003>
- Scheiter, K., Schleinschok, K., & Ainsworth, S. (2017). Why sketching may aid learning from science texts: Contrasting sketching with written explanations. *Topics in Cognitive Science*, 9, 866–882. <http://dx.doi.org/10.1111/tops.12261>
- Schiefele, U. (1999). Interest and learning from text. *Scientific Studies of Reading*, 3, 257–279. http://dx.doi.org/10.1207/s1532799xssr0303_4
- Schmeck, A., Mayer, R. E., Opfermann, M., Pfeiffer, V., & Leutner, D. (2014). Drawing pictures during learning from scientific text: Testing the generative drawing effect and the prognostic drawing effect. *Contemporary Educational Psychology*, 39, 275–286. <http://dx.doi.org/10.1016/j.cedpsych.2014.07.003>
- Schraw, G. (1998). Processing and recall differences among selective details. *Journal of Educational Psychology*, 90, 3–12. <http://dx.doi.org/10.1037/0022-0663.90.1.3>
- Schwaborn, A., Mayer, R. E., Thillmann, H., Leopold, C., & Leutner, D. (2010). Drawing as a generative activity and drawing as a prognostic activity. *Journal of Educational Psychology*, 102, 872–879. <http://dx.doi.org/10.1037/a0019640>
- Siemankowski, F. T., & MacKnight, F. C. (Eds.). (1971). In *Proceedings from 44th annual meeting of the national association for research in science teaching*. Silver Spring, MD: National Association for Research in Science Teaching.
- Singer, M., & Gagnon, N. (1999). Detecting causal inconsistencies in scientific text. In S. R. Goldman, A. C. Graesser, & P. van den Broek (Eds.), *Narrative comprehension, causality, and coherence: Essays in honor of Tom Trabasso* (pp. 179–194). Mahwah, NJ: Erlbaum.
- Sung, E., & Mayer, R. E. (2012). When graphics improve liking but not learning from online lessons. *Computers in Human Behavior*, 28, 1618–1625. <http://dx.doi.org/10.1016/j.chb.2012.03.026>
- Suwa, M., & Tversky, B. (1997). What do architects and students perceive in their design sketches? A protocol analysis. *Design Studies*, 18, 385–403. [http://dx.doi.org/10.1016/S0142-694X\(97\)00008-2](http://dx.doi.org/10.1016/S0142-694X(97)00008-2)
- Sweller, J. (2005). Implications of cognitive load theory for multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 19–30). New York, NY: Cambridge University Press. <http://dx.doi.org/10.1017/CBO9780511816819.003>
- Tarchi, C. (2017). Comprehending expository texts: The role of cognitive and motivational factors. *Reading Psychology*, 38, 154–181. <http://dx.doi.org/10.1080/02702711.2016.1245229>
- van den Broek, P. (2010). Using texts in science education: Cognitive processes and knowledge representation. *Science*, 328, 453–456. <http://dx.doi.org/10.1126/science.1182594>
- van Dijk, T. A., & Kintsch, W. (1983). *Strategies of discourse comprehension*. New York, NY: Academic Press.
- Van Meter, P. (2001). Drawing construction as a strategy for learning from text. *Journal of Educational Psychology*, 93, 129–140. <http://dx.doi.org/10.1037/0022-0663.93.1.129>
- Van Meter, P., Aleksic, M., Schwartz, A., & Garner, J. (2006). Learner-generated drawing as a strategy for learning from content area text. *Contemporary Educational Psychology*, 31, 142–166. <http://dx.doi.org/10.1016/j.cedpsych.2005.04.001>
- Van Meter, P., & Firetto, C. M. (2013). Cognitive model of drawing construction. In G. Schraw, M. T. McCrudden, & D. Robinson (Eds.), *Learning through visual displays* (pp. 247–280). Charlotte, NC: Information Age Publishing, Inc.
- Van Meter, P., & Garner, J. (2005). The promise and practice of learner-generated drawing: Literature review and synthesis. *Educational Psychology Review*, 17, 285–325. <http://dx.doi.org/10.1007/s10648-005-8136-3>
- Vosniadou, S., & Brewer, W. F. (1992). Mental models of the earth: A study of conceptual change in childhood. *Cognitive Psychology*, 24, 535–585. [http://dx.doi.org/10.1016/0010-0285\(92\)90018-W](http://dx.doi.org/10.1016/0010-0285(92)90018-W)
- Wade, S. E., & Adams, R. B. (1990). Effects of importance and interest on recall of biographical text. *Journal of Reading Behavior*, 22, 331–353. <http://dx.doi.org/10.1080/10862969009547717>
- Wang, Z., Sundararajan, N., Adesope, O. O., & Ardasheva, Y. (2016). Moderating the seductive details effect in multimedia learning with note-taking. *British Journal of Educational Technology*, 48, 1380–1389. <http://dx.doi.org/10.1111/bjet.12476>
- Wiley, J. (2003). Cognitive and educational implications of visually rich media: Images and imagination. In M. E. Hocks & M. R. Kendrick (Eds.), *Eloquent images: Word an image in the age of new media* (pp. 201–215). Cambridge, MA: MIT Press.
- Wiley, J., & Myers, J. L. (2003). Availability and accessibility of information and causal inferences from scientific text. *Discourse Processes*, 36, 109–129. http://dx.doi.org/10.1207/S15326950DP3602_2
- Wiley, J., Sanchez, C. A., & Jaeger, A. J. (2014). The individual differences in working memory capacity principle for multimedia learning. In R. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2nd ed., pp. 598–619). New York, NY: Cambridge University Press. <http://dx.doi.org/10.1017/CBO9781139547369.029>
- Wright, P., Milroy, R., & Lickorish, A. (1999). Static and animated graphics in learning from interactive texts. *European Journal of Psychology of Education*, 14, 203–224. <http://dx.doi.org/10.1007/BF03172966>

(Appendices follow)

Appendix A

Example Text Passage With Italicized Seductive Details

Stratovolcanoes are the typical volcanoes most people envision and are usually located near convergent plate boundaries where subduction is occurring, particularly around the Pacific basin. *Some stratovolcanoes are covered in black volcanic sand making them popular sites for tourist activities such as sand boarding.* The magma produced by subduction is generally very thick and does not allow gas to readily escape from the magma. When the magma reaches the vent of the volcano, gas bubbles begin to form and to grow. The rapid expansion of the gas tears the magma apart, and the volcano erupts violently, producing great volumes of ash. *After the eruption of Mt. Galunggung in Indonesia, an airplane flew*

through an ash cloud resulting in the failure of all four engines. The plane descended rapidly and the engines restarted only minutes before impacting the ground. If enough gas escapes, the volcano can produce a sticky, slow-moving lava flow, which may only travel a short distance from the vent before solidifying. *When Mount St. Helen's erupted it created a landslide that carried mud and debris down the mountain at speeds of over 100 miles per hour for more than 3 miles.* Earthquakes can also be caused when oceanic and continental plates collide. Further, the movement of magma in subduction zones can also trigger deep earthquakes and rarely, large earthquakes can trigger volcanic eruptions.

Appendix B

Student Drawing Examples

For the sketches depicting the plate interactions that cause stratovolcanoes, 5 points were possible. (See Figures B1 and B2.)

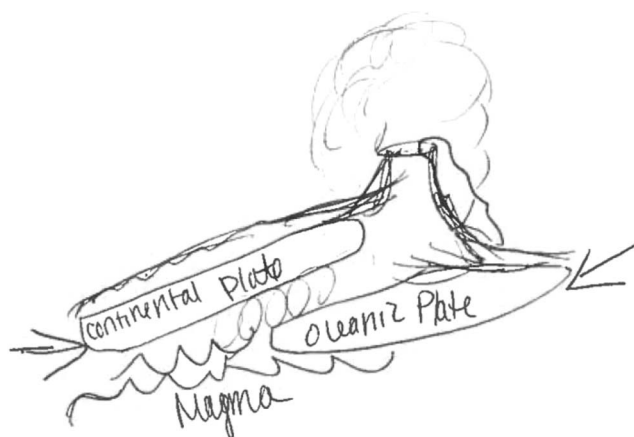


Figure B1. Example of a sketch that received five points for depicting the plate interaction that creates stratovolcanoes.



Figure B2. Example of a sketch that received zero points for depicting the plate interaction that creates stratovolcanoes.

(Appendices continue)

Appendix C

Correct Sketch and Summary

Correct Sketch (see Figure C1)

Correct Summary

Stratovolcanoes are formed when an oceanic plate converges with a continental plate. The oceanic plate is subducted under the

continental plate because the oceanic plate is denser. At the subduction site, a trench forms. As the oceanic plate is subducted further into the mantle, it experiences high pressure and temperatures which causes it to melt. This melted crust produces magma which is very thick and gaseous. As the magma builds up, it rises to the surface and erupts in the form of a volcano.

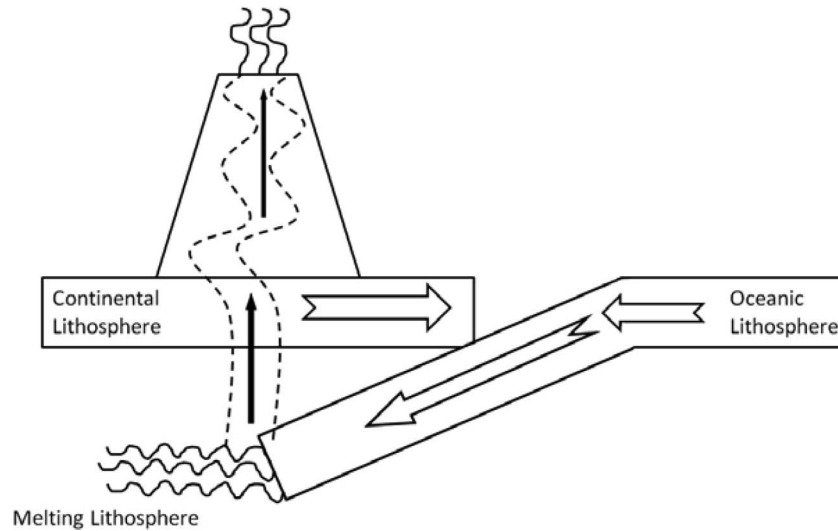


Figure C1. Correct sketch depicting the plate interaction that creates stratovolcanoes.

Received August 4, 2017

Revision received September 29, 2017

Accepted November 9, 2017 ■