The Journal of Experimental Education

Virginia, C., Alibali, M., & Nathan, M. (2016). Learning About Posterior Probability: Do Diagrams and Elaborative Interrogation Help?, The Journal of Experimental Education, 84:3, 579-599, DOI: 10.1080/00220973.2015.1048847

ISSN: 0022-0973 (Print) 1940-0683 (Online) Journal homepage: http://www.tandfonline.com/loi/vjxe20

Learning About Posterior Probability: Do Diagrams and Elaborative Interrogation Help?

Virginia Clinton, Martha W. Alibali & Mitchell J. Nathan

THE JOURNAL OF EXPERIMENTAL EDUCATION, *84*(3), 579–599, 2016

ISSN: 0022-0973 print/1940-0683 online
DOI: 10.1080/00220973.2015.1048847

Routledge
Taylor & Francis Group

LEARNING, INSTRUCTION, AND COGNITION

Learning About Posterior Probability: Do Diagrams and Elaborative Interrogation Help?

Virginia Clinton
University of North Dakota

Martha W. Alibali and Mitchell J. Nathan
University of Wisconsin-Madison

To learn from a text, students must make meaningful connections among related ideas in that text. This study examined the effectiveness of two methods of improving connections—elaborative interrogation and diagrams—in written lessons about posterior probability. Undergraduate students ($N = 198$) read a lesson in one of three questioning conditions (read twice, embedded questioning, and elaborative interrogation) and one of three diagram conditions (text only, diagram without redundant text, and diagram with redundant text). Elaborative interrogation negatively affected learning from the lesson, relative to reading the lesson twice. One possible explanation for this finding is that the quality of answers to the elaborative interrogations was poor. When the lesson was read twice, diagrams helped learning from the lesson relative to text only. Implications of these findings for instruction in probabilistic reasoning are discussed.

Keywords *diagram, elaborative interrogation, multimedia principle, redundancy principle*

PROBABILISTIC REASONING IS challenging for undergraduate students. However, it is important that students learn about probabilistic reasoning so they can be more informed citizens in modern society (Garfield & Ben-Zvi, 2008). One type of probabilistic reasoning task that students find particularly difficult is calculating posterior probability (i.e., updating prior probability using Bayes' theorem, such as when determining a test's positive predictive value; Kahneman & Tversky, 1972; Stanovich & West, 1998; Zieffler et al., 2008). This task involves considering both the probability of a particular characteristic or disease in a given population (i.e., the base

Address correspondence to Virginia Clinton, 319 Harvard St., Grand Forks, ND 58202, USA. E-mail: virginia.clinton@und.edu

rate) and the probability that the identification of that characteristic or disease is accurate (i.e., the test accuracy).

The integration of base rate and test accuracy information can be challenging, presumably because of the complexities in connecting the two probabilities correctly (Bar-Hillel, 1980; Garcia-Retamero & Hoffrage, 2013; Konheim-Kalkstein, 2008). Therefore, techniques to improve connections between the two probabilities could help students learn about posterior probability. Moreover, students learn more when meaningful connections are made throughout a lesson (Kintsch, 1998; McNamara, Levinstein, & Boonthum, 2004). The purpose of this study is to examine the effects of two techniques that may promote students' making connections among ideas within a written lesson: elaborative interrogation and diagrams.

Elaborative Interrogation

One technique that may foster connections is *elaborative interrogation*. Elaborative interrogation typically involves "how" or "why" questions intended to prompt students to integrate an idea from a written lesson with other ideas from the lesson or with background knowledge (e.g., McDaniel & Donnelly, 1996; Pressley, Symons, McDaniel, Snyder, & Turnure, 1988; see Dunlosky, Rawson, Marsh, Nathan, & Willingham, 2013, for discussion). Elaborative interrogation has been found to promote learning from texts, likely because of the greater number of connections students make within the text or between the text and their background knowledge (Rouet & Vidal-Abarca, 2002). Thus, it seems likely that elaborative interrogation would improve learning from reading a lesson about posterior probability. Indeed, other techniques that are intended to foster connection making in ways similar to elaborative interrogation, such as self-explanation, have been shown to benefit learning. For example, self-explanation prompts designed to encourage connections have been found to promote learning from probability lessons in which problems are explicitly solved step by step (i.e., worked examples; Atkinson, Renkl, & Merrill, 2003).

It is important to note the similarities and differences between elaborative interrogation and self-explanation in order to understand the findings related to these techniques. Elaborative interrogation is a specific form of questioning designed to prompt connections within the text or between the text and the student's background knowledge (Pressley et al., 1998; Ozgungor & Guthrie, 2004). In contrast, the term "self-explanation" encompasses a broad range of prompts that vary across the literature. In some studies, the prompts are instructions to explain the material to oneself (Chiu & Chi, 2014) or explain how a proposed answer is correct or incorrect (Rittle-Johnson, 2006). In other studies, the prompts are open-ended questions for the student to answer (Leppink, Broers, Imbos, van der Vleuten & Berger, 2012). Another type of self-explanation prompt is designed to prompt connections in a manner similar to elaborative interrogations (Berthold & Renkl, 2009). Because the term *self-explanation* has been used in such varied ways, we use the term *elaborative interrogation* to describe the technique we used in the current study because the term is more precise than self-explanation (see Dunlosky et al., 2013, for a discussion).

In one form of elaborative interrogation, students are asked to explain why a fact presented in the text is true (e.g., Martin & Pressley, 1991; Pressley, McDaniel, Turnure, Wood, & Ahmad, 1987; Smith, Holliday, & Austin, 2010; Woloshyn, Pressley, & Schneider, 1992). This type of elaborative interrogation is intended to prompt students to connect facts with their background knowledge (Willoughby, Waller, Wood, & MacKinn, 1993). Hence, students with higher levels of background knowledge typically accrue greater benefits from this approach than do their peers with lower levels of background knowledge (e.g., Pressley et al., 1987; Willoughby et al., 1993). For students with lower levels of background knowledge, a different form of elaborative interrogation may be more effective—namely, questions that can be answered using information in the text (Ozgungor & Guthrie, 2004). In support of this idea, one study showed that students with low levels of background knowledge benefited from self-explanation prompts in a statistics lesson only when the answers to the prompts were provided in the lesson (Leppink et al., 2012). These findings imply that, for a topic such as calculating posterior probability, for which undergraduate students typically lack background knowledge (Beyth-Marom, Fidler, & Cumming, 2008; Keeler & Steinhorst, 2001), elaborative interrogation that guides students to connect information within the text may be more appropriate than elaborative interrogation that guides students to connect information to background knowledge.

Diagrams

Another technique that may promote students' making connections within a lesson is the use of diagrams. According to the *multimedia principle*, learning is greater when texts are accompanied by relevant visuals (Mayer, 2009), because students develop a verbal mental model based on the text and a visual mental model based on the visuals. Students are more likely to make connections among different ideas when the information is represented in two distinct mental models (Mayer, 1999; Mayer & Gallini, 1990). These connections between the two mental models prompt deeper processing of the text and promote learning (Mayer, 2002).

Diagrams may be particularly useful in lessons about probability. In one previous study, performance on probability problems was improved through instruction on constructing diagrams (Beitzel & Staley, 2015). In addition, people solve posterior probability problems more accurately when diagrams are included, suggesting that diagrams may help people comprehend the problems (Garcia-Retamero & Hoffrage, 2013). One study in particular has suggested that students benefit from *frequency tree* diagrams in which probabilities are displayed in frequencies when learning how to solve posterior probability problems (Sedlmeier & Gigerenzer, 2001). In this study, students worked through examples of calculating posterior probability using a computer tutorial. Students who received this instruction with diagrams showed remarkable gains in accurately calculating posterior probability. In contrast, students who received instruction without diagrams showed little improvement. One reason for this could be the connections afforded by the presence of verbal and visual representations leading to better learning (Mayer, 2009).

Sedlmeier and Gigerenzer's (2001) instructional method with diagrams was effective, but it involved individualized instruction with feedback (e.g., computer tutorials or one-on-one tutoring). At the postsecondary level, instructors expect students to learn from reading independently (Kim & Anderson, 2011). Reading independently does not provide the opportunity for practice and feedback on performance that a computer tutorial can. Therefore, there is a need to develop effective lessons from which students can learn when reading independently. Given that making meaningful connections can improve learning from written lessons (Kintsch, 1998; McNamara et al., 2004) and that diagrams are most effective when they can be connected to relevant text (Atkinson, Derry, Renkl, & Wortham, 2000; Tarmizi & Sweller, 1988), techniques for improving connections between diagrams and texts while students are reading independently should be examined.

Making connections between the text and diagram could be easier if there were similar surface features in both the text and the diagram (Renkl, 2005). In the case of frequency tree diagrams, this would involve having the frequencies in the diagram also provided in the text. Indeed, previous work has shown that adding text to diagrams can help guide connections between the verbal information and visual information, thereby enhancing learning (Johnson & Mayer, 2012; Mason, Pluchino, & Tornatora, 2013; Mayer & Johnson, 2008). On the other hand, according to the *redundancy principle*, information presented in the diagram should not also be presented in the text (Sweller, Ayres, & Kalyuga, 2011). This is because presenting information in both text and diagram increases the amount of information the student has to process, which may overload working memory and diminish learning (e.g., Chandler & Sweller, 1991). Furthermore, other findings have indicated that, although diagrams with redundant text may benefit memory for the lesson, there does not seem to be a benefit for accurate application of the ideas in the lesson (McCrudden, Hushman, & Marley, 2014; Ortegren, Serra, & England, 2015).

It is possible that the redundant information would actually be helpful for lessons about probabilistic reasoning with diagrams. Previous studies in which the redundant information led to less learning used complex diagrams for scientific concepts with detailed information that may be difficult to communicate through text (e.g., Holliday, 1976). In contrast, it is possible that having the simple, numeric information in the text that is redundant with the diagram would serve a function similar to labels and could improve connections between the text and diagram. In this way, it is possible that text redundant with the diagram could increase learning from the lesson.

Elaborative Interrogation with Diagrams

It stands to reason that diagrams might be particularly beneficial when elaborative interrogation is included in the lesson. Students often attend to visual representations less than they attend to text (Schüler, Scheiter, Rummer, & Gerjets, 2012; Schwonke, Berthold, & Renkl, 2009), and without sufficient attention to visual representations, students cannot connect visual and verbal information. Because connections between visual and verbal information are critical for meaningful learning (Cromley, Snyder-Hogan, & Luciw-Dubas, 2010; Mayer, 2009), methods that prompt students to integrate verbal and visual information in text may promote learning. Elaborative interrogation may be particularly useful for guiding students to attend to the visual representations and then to integrate visual and verbal information (Aleven & Koedinger, 2002). This could be accomplished by posing elaborative interrogations that encourage students to connect information in the diagram and the text, in addition to elaborative interrogations that prompt connections within the text. This idea is supported by previous work, indicating that self-explanation prompts can foster students making connections between relevant information in text and diagrams in worked examples (Berthold & Renkl, 2009; Berthold, Eysink, & Renkl, 2009). If elaborative interrogation is beneficial for learning from a lesson on posterior probability, this benefit might be enhanced when elaborative interrogation is coupled with diagrams.

The Current Study

The purpose of this study was to examine the effects of elaborative interrogation and diagrams on learning about posterior probability from a lesson that students read independently. Learning was assessed with posterior probability problems on which the logic presented in the lesson could be applied. We predicted that students would learn more with elaborative interrogation.

As in past work (e.g., Ozgungor & Guthrie, 2004; Pease, 2012; Smith et al., 2010), we included two additional conditions to allow us to better assess the effects of elaborative interrogation. To control for time on task (i.e., time with the lesson), one group of students read the lesson twice. To control for the effects of simply answering questions (e.g., McMaster et al., 2012), a second group of students answered simple "who" or "what" questions about specific ideas explicitly stated in the lesson. These questions could be accurately answered with only shallow processing of the lesson material. By comparing performance in the elaborative interrogation condition against performance in these two other conditions, we could be confident that any observed effects of elaborative interrogation were due to these questions prompting deeper processing, rather than time with the lesson or the simple act of answering questions.

In line with previous research on visual representations (e.g., Beitzel & Staley, 2015; Mayer, 2009), we also predicted that students would learn more if their assigned lesson included a diagram. However, we had competing predictions as to whether learning would be greatest if the information presented in the diagram were also presented redundantly in the text, or if that information were not presented redundantly in the text. Based on previous work with adding text to diagrams (Johnson & Mayer, 2012; Mason et al., 2013), including text that is redundant with the diagram may enhance learning. In contrast, based on the redundancy principle (Sweller et al., 2011), including text redundant with the diagram could decrease learning. However, we expected that diagrams would enhance learning relative to text alone, regardless of the presence of redundant text.

If we find that elaborative interrogation and diagrams are beneficial for student learning, we might find that combining elaborative interrogation with diagrams would enhance their effects. Some of the elaborative interrogations in the lessons with diagrams were designed to prompt connections between visual information in the diagram and verbal information in the text. Connecting information in visual and verbal representations has previously been found to be beneficial for learning (Cromley et al., 2010; Mayer, 2009; Mayer, Steinhoff, Bower, & Mars, 1995). Hence, we predicted that the lesson that combined diagrams and elaborative interrogations would lead to optimal learning.

We were also interested in how difficult the lessons were for students. Difficulty was assessed through ratings as well as through time spent on the lesson. Because answering elaborative interrogations likely requires more effort than reading twice or answering simple questions, we anticipated that lessons with elaborative interrogations would be more difficult. However, we expect that elaborative interrogation would be a *desirable difficulty* leading to increased learning (McDaniel & Butler, 2011; Richland, Bjork, Finley, & Linn, 2005). Knowledge of how difficult the lessons were for students could be useful in interpreting our findings for redundant text. Recall that the redundancy principle states that information should be presented in the diagram *or* the text, because including information in both the diagram and the text would be extra, unnecessary information for students to process (Sweller et al., 2011). Hence, this principle would be supported if students report that lessons with redundant text were more difficult than lessons without redundant text. In contrast, previous work has shown that adding redundant text to diagrams through labeling may make it easier to connect the text and diagram (Johnson & Mayer, 2012). Therefore, students may find that diagrams with redundant text are less difficult than diagrams without redundant text. Overall, the difficulty ratings in conjunction with performance on the learning measures could inform our understanding of the effects of elaborative interrogations and diagrams.

METHODS

Participants

Participants were 248 undergraduates who earned extra credit in their introductory psychology course for participation. Three participants left the study before completing the posttest; their data were excluded. Of the remaining 245 participants, there were 158 females and 86 males (1 participant did not report gender) with an average reported age of 19.05 years ($SD = 2.89$ years). English was reported as the native language of 164 participants; 80 participants reported a native language other than English (1 participant did not report a native language). Regarding ethnicity, 156 were Caucasian, 64 were Asian, 9 were Latino, 5 were African American, 1 was Native American, and 5 were biracial or multiracial (5 participants did not report ethnicity).

Materials

The lesson was adapted from a textbook by Heuer (1999) and incorporated an example from Sedlmeier and Gigerenzer's (2001) experiments. There were two examples of scenarios in which base rate and specificity probabilities were presented, and the posterior probability was calculated using natural frequencies. For the diagram conditions, there was a frequency tree diagram for each of the two examples. For the diagram-without-redundant-text condition, some of the information in the diagram was not also presented in the text. Specifically, the frequency information provided in the diagram was not in the text. For the diagram-with-redundant-text condition, the frequency information in the diagram was also presented in the text (see Appendix for lesson excerpts from each diagram condition). The texts for the text-only and diagram-with-redundant-text conditions were identical, with the exception that the diagram-with-redundant text condition contained a sentence directing the readers' attention to the diagram. The text-only condition was 987 words long, with a Flesch-Kincaid Grade Level of 12.0 and Flesch Reading Ease of 40.9. The text for the diagram-without-redundant-text condition was 979 words long, with a Flesch-Kincaid Grade Level of 12.0 and Flesch Reading Ease of 41.7. The diagram-with-redundant-text condition was 1,003 words long with a Flesch-Kincaid Grade Level of 12.0 and Flesch Reading Ease of 41.3.

For the two question conditions, (i.e., the elaborative interrogation and embedded questioning conditions), there was one question after approximately every paragraph, for a total of six questions. For the diagram and questioning conditions, two of the six questions were asked after the diagrams in the lesson. For the diagram-without-redundant-text condition, answering the two questions after diagrams correctly required using information from the diagram. However, for the diagram-with-redundant-text condition, the questions could be answered using information from either the text or the diagram (or both).

Elaborative interrogation questions were complex "how" and "why" questions (e.g., "Why does including the base rate make the probability that a patient with fever, chills, and skin lesions has sepsis 47%?" and "Why is the base rate fallacy such a common mistake?"). Following Ozgungor and Guthrie (2004), the answers to elaborative interrogations were implicit in the text; the participant had to connect different ideas presented in the text and/or diagrams. Therefore, participants did not need to use background knowledge to answer the elaborative interrogation prompts. In contrast, embedded questions were simple (e.g., "What is the base rate?" and "How many patients with sepsis do not have high fever, chills, and skin lesions?") and their answers were explicitly stated in the text and/or clearly presented in the diagrams. At the beginning of each lesson was a text reminder that the participant would be asked to answer questions based on the information in the lesson after reading the lesson. The lesson in the read-twice condition also began with the text reminder that participants needed to read the lesson twice. At the end of the read-twice condition lesson, participants were reminded that if they had only read the lesson once, they needed to reread the lesson starting at its beginning.

Measures

Pretest

Prior to reading the lesson, participants were asked to solve the following posterior probability problem adapted from Bramwell, West, and Salmon (2006): "All medical tests have error. For example, the serum test screens pregnant women for fetuses with Down syndrome. The test is a very good one, but not perfect. Roughly 100 fetuses out of 10,000 have Down syndrome. Of these 100 fetuses with Down syndrome, 90 pregnant women will have a positive test result. Of the remaining 9,900 unaffected fetuses, 99 pregnant women will still have a positive test result. What is the probability a pregnant woman who has a positive result on the test actually has a fetus with Down syndrome?" The problem was presented only as text; no visuals were included.

Difficulty Self-Report

Following the lesson, participants were asked to rate their level of agreement with the statement, "The lesson I just read was difficult," on a Likert scale from 1 to 7.

Comprehension Assessment

We developed a comprehension assessment based on previously used sentence verification (Royer, 2001) and inference verification (Wiley & Voss, 1999) techniques. The assessment consisted of 13 sentences. Seven of the sentences were correct or incorrect paraphrases (i.e., contained or contradicted information explicitly stated in the lesson) and six of the sentences were correct or incorrect inferences (i.e., based on information in the lesson that was not explicitly stated). These paraphrases and inferences were based on ideas that were evenly distributed throughout the lesson. Participants were asked to indicate whether each sentence was consistent or inconsistent with what they had just read. Internal consistency for the comprehension assessment was unacceptable for the paraphrases (Cronbach's $\alpha = .34$), unacceptable for the inferences (Cronbach's $\alpha = .30$), and poor for the overall measure (Cronbach's $\alpha = .50$). Because of the low internal consistency, the comprehension assessment was not used as a learning measure and it is not discussed further.

Posttest

A posttest was used to assess learning of calculating posterior probability. The posttest consisted of four posterior probability problems similar to the one on the pretest. The problems were presented in text only; no visuals were included. Internal consistency for the posttest was good (Cronbach's $\alpha = .87$).

Procedure

Participants were randomly assigned to one of nine conditions, reflecting a 3 (diagram condition: diagram with redundant text, diagram without redundant text, text only) x 3 (questioning condition: elaborative interrogation, embedded questions, read twice) factorial design. One to six participants came to the testing room for a given experimental session, and they engaged in the experimental tasks individually. All participants in each small group were in the same condition. All tasks were using paper and pencil (i.e., nothing was administered on a computer). After providing informed consent, participants were given one posterior probability problem to solve as a pretest. Then, if they were in one of the questioning conditions (i.e., elaborative interrogation or embedding questioning), they were given the lesson and an answer sheet for their questions. The experimenter instructed participants to read the lesson carefully because they would be asked to answer questions and solve problems based on its information. Participants in the read-twice condition were instructed to read the lesson twice. The experimenter monitored the participants to ensure that they followed directions and read the lesson twice. Participants in the questioning conditions were instructed to answer the numbered questions in the lessons in the appropriate places on the answer sheet. Participants were permitted to look back to previously read text, but they were not specifically instructed to do so. When participants finished reading the lesson, they returned it to the experimenter and completed the self-report of lesson difficulty. Then, they solved 21 simple multiplication and division problems as a distracter task, and afterward completed a comprehension assessment followed by the posttest. Finally, participants self-reported their demographic information. All of the tasks were completed in a single session, which was approximately 45 to 60 minutes in length. The experimenter recorded the time participants received the lesson as well as the time they returned the lesson to the experimenter.

Scoring

Pre- and Posttests

Pre- and posttests were scored for accuracy. An accurate answer was the number of true positives over the combined number of true positives and false positives for any given problem. Pretest scores were either 0 or 1, as the pretest had only one item. The maximum posttest score was four.

Difficulty

The participant's response on the Likert scale was used as a measure of perceived difficulty.

Answers to Elaborative Interrogation Questions

Answers to elaborative interrogation questions were scored using a rubric that contained the following categories: ideal, adequate, inadequate, circular, wrong, and missing. Elaborative interrogation (EI) questions were designed to prompt integration of concepts within the lesson.

TABLE 1
Descriptive Statistics (Means and Standard Errors) By Condition for Pretest Scores

	Text only M(SE)	*Diagram without redundant text* M(SE)	*Diagram with redundant text* M(SE)	*Total* M(SE)
Read twice	.07 (.07)	.39 (.08)	.17 (.06)	.21 (.04)
Embedded questioning	.21 (.07)	.39 (.11)	.16 (.07)	.25 (.05)
Elaborative interrogation	.16 (.07)	.13 (.08)	.22 (.08)	.17 (.05)
Total	.14 (.04)	.29 (.05)	.18 (.04)	

Therefore, an answer was considered ideal if it correctly included three or more concepts presented in the lesson. Adequate answers correctly included two concepts. Inadequate answers correctly included only one concept. Circular answers repeated the information in the question (e.g., EI: "Why is the base rate important when calculating probability?" Circular Answer: "Because it is important information you need to use to calculate probability."). Wrong answers provided incorrect or inappropriate information. Missing answers included responses such as "I have no idea" in addition to blank responses. Two research assistants coded answers to elaborative interrogations. Twenty-five percent of the answers were coded in common, with excellent reliability (Cohen's $\kappa = .84$). Disagreements were resolved through discussion.

To convert the rubric to scores, ideal answers were given a score of "3," adequate answers were given a score of "2," and inadequate answers were given a score of "1." Missing, incorrect, and circular answers were given a score of "0." The scores for all six of the elaborative interrogations were summed for a measure of answer quality. The maximum possible score was 18.

Answers to Embedded Questions

Because embedded questions were much simpler than elaborative interrogations, answers to embedded questions were scored dichotomously as correct or incorrect. As with the elaborative interrogations, two research assistants coded answers to embedded questions. Twenty-five percent of the answers were coded in common, with excellent reliability (Cohen's $\kappa = .98$). Disagreements were resolved through discussion. Correct answers were given a score of "1"; and incorrect answers, "0." The number of correct answers was the total score. The maximum possible score was 6.

RESULTS

Table 1 presents descriptive statistics for pretest scores by condition. Prior to testing the hypotheses, we examined the a priori distribution of pretest scores by condition. Although the distribution of pretest scores did not differ significantly as a function of diagram condition, χ^2 (2, $N = 243$) = 5.04, $p = .08$, or questioning condition, χ^2 (2, $N = 243$) = .62, $p = .73$, the difference in pretest scores by condition could still confound results. For this reason, participants who answered the pretest question correctly ($N = 47$) were removed from the analyses. To test the effects of diagram and questioning on learning from the lesson, we used logit mixed models, as recommended by Jaeger (2008; see Snijders & Bosker, 2012, for more information about multilevel models,

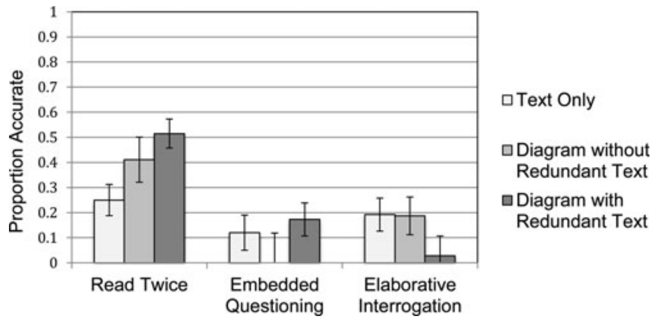

FIGURE 1 Posttest-item accuracy by condition (means with standard error bars; $N = 198$).

including logit mixed models), implemented using the package lme4 in the *R* statistical software (Bates, 2010; Bates, Maechler, & Bolker, 2012). Pairwise corrections were not necessary because of the use of multilevel modeling through lme4 (Gelman, Jill, & Yajima, 2012). For all analyses, we set the Type I error rate at $\alpha = .05$. We report Type III Wald chi-square tests of the parameter estimates against zero.

Did Questioning and/or Diagrams Affect Posttest Accuracy?

We first examined the effects of the questioning and diagram conditions on posttest accuracy. Based on previous research findings (Ozgungor & Guthrie, 2004; Pressley et al., 1987; Smith et al., 2010), we expected that elaborative interrogation would improve learning. Based on the multimedia principle, we expected that diagrams would improve application of the lesson content, as indicated by accuracy on posttest questions, relative to lessons with only text. We had competing predictions regarding the effects of redundant text with the diagram. Based on the redundancy principle, it is possible that redundant text would add unnecessary information to process, thereby, decreasing learning. However, it is also possible, based on previous work with labeling diagrams (Johnson & Mayer, 2012) that text redundant with the diagram would facilitate connections between the text and the diagram, thereby, increasing learning.

To address these hypotheses, we used a mixed-effects model with questioning and diagram condition as fixed factors, problem and participant as random factors, and accuracy on each item as the dependent variable. Figure 1 presents the mean posttest-item accuracy by condition. In our analysis, we first considered the effects of *diagram presence*, by collapsing the two conditions that involved diagrams and comparing them to the text-only condition. We then compared the two conditions that included diagrams (diagram with redundant text and diagram without redundant text) to one another.

Contrary to hypotheses, there was no main effect of diagram presence on posttest accuracy, Wald $\chi^2(2, N = 198) = 1.43$, $p = .15$. However, there was a significant interaction of diagram presence and questioning condition, Wald $\chi^2(4, N = 198) = 6.89$, $p = .03$. In the read-twice condition, participants whose lessons included diagrams performed better on the posttest than participants whose lessons did not include diagrams, $B = 3.11$, Wald $Z = 2.29$, $p = .02$. However, diagram presence did not affect posttest performance for participants in the embedded questioning condition, $B = .19$, Wald $Z = .12$, $p = .91$, or the elaborative interrogation condition, $B = .86$, Wald $Z = .28$, $p = .60$.

Questioning condition had a main effect on posttest-item accuracy, Wald $\chi^2(2, N = 198) = 13.73$, $p < .001$. Unexpectedly, however, participants in the elaborative interrogation condition had *lower* posttest-item accuracy than did participants in the read-twice condition, $B = 3.75$, Wald $Z = 13.81$, $p < .001$. There were no differences between the elaborative interrogation and embedded questioning conditions, $B = .81$, Wald $Z = .50$, $p = .62$.

We next examined whether the redundancy of the diagram with the text influenced posttest performance. To address this question, we used a mixed-effects model to compare the diagram with redundant text and diagram without redundant text conditions (i.e., excluding the text-only condition). There was no main effect of diagram condition, Wald $\chi 2(2, N = 198) = .78$, $p = .44$ and no interaction with questioning condition, Wald $\chi 2(4, N = 198) = 1.30$, $p = .52$. Thus, performance was not affected by the redundancy of the text and diagram; participants performed similarly with and without redundant text.

The negative effect of questioning condition was also evident in this analysis, Wald $\chi 2(2, N = 198) = 16.27$, $p < .001$. However, in this subsample, the benefit of reading twice compared to elaborative interrogation was not reliable, $B = 3.33$, Wald $Z = 1.68$, $p = .09$. The elaborative interrogation and embedded questioning conditions did not differ, $B = -.89$, Wald $Z = 00$, $p = .99$.

In brief, contrary to what was predicted, elaborative interrogation was not beneficial for learning how to solve posterior probability problems. Indeed, reading twice yielded higher scores on the posttest than elaborative interrogation. Diagrams also had limited effects; they helped only when participants read the lesson twice, and not when they answered questions. Redundant text did not appear to affect posttest accuracy for students whose lessons had diagrams.

Did Diagrams and Questioning Affect Lesson Difficulty and Time with the Lesson?

To better understand the nature of the posttest findings, we also examined the effects of diagrams and elaborative interrogations on lesson difficulty ratings and time with the lesson (i.e., the number of minutes spent with the lesson). The number of minutes spent with the lesson was positively skewed, so a square-root transformation was applied prior to analysis (Osborne, 2002). Because there was only one rating of difficulty or measure of time with the lesson, the multilevel modeling analyses conducted with posttest item accuracy were not appropriate (Snijders & Bosker, 2012). Instead, we conducted ANOVAs with diagram condition and questioning condition as fixed factors and with lesson-difficulty ratings and minutes spent on the lesson as the dependent variables. Bonferonni corrections were applied to pairwise comparisons.

As with posttest accuracy, we examined the effect of a diagram in two ways. The first was to examine diagram presence by collapsing the two diagram conditions and comparing them to text alone. The second was to compare the two diagram conditions to one another. We expected that diagrams would make the lessons easier to comprehend, thereby, yielding lower difficulty ratings and less time with the lesson. We did not have specific predictions regarding the effects of redundancy of text and diagram. In addition, we expected that elaborative interrogation would increase the amount of effort exerted on the lesson, thereby, yielding higher difficulty ratings and more time spent with the lesson.

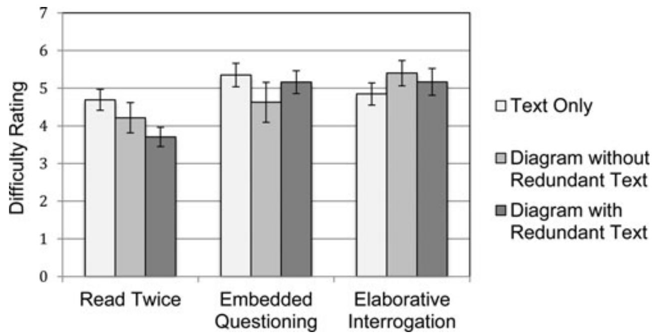

FIGURE 2 Difficulty ratings by condition (means with standard error bars; $N = 198$).

Lesson Difficulty

Figure 2 presents the mean difficulty ratings by condition. We first examined whether diagram presence and questioning condition affected difficulty ratings. There was no main effect of diagram presence on lesson-difficulty ratings, $F(1, 196) = 1.16$, $p = .28$. However, there was a significant interaction, $F(2, 192) = 3.04$, $p = .05$. In the read-twice condition, participants whose lesson included a diagram rated the lesson as easier than those whose lesson did not include a diagram, $F(1, 75) = 5.76$, $p = .02$, Cohen's $d = .55$. In the embedded questioning and elaborative interrogation conditions, presence of the diagram did not affect difficulty ratings, $F(1, 54) = .65$, $p = .43$ and $F(1, 62) = 1.24$, $p = .27$, respectively.

There was also a main effect of questioning condition, $F(2, 195) = 7.3$, $p = .001$. Participants in the elaborative-interrogation condition rated the lesson as more difficult than did participants in the read-twice condition, $p = .001$, Cohen's $d = .59$. In addition, participants in the embedded-questioning condition also rated the lesson as more difficult relative to the ratings of participants in the read-twice condition, $p = .01$, Cohen's $d = .63$.

Next, we examined whether redundant text affected difficulty ratings for lessons with diagrams. As with posttest-item accuracy, there was no effect of text redundancy on difficulty ratings by students whose lessons had diagrams, $F(1, 119) = .05$, $p = .82$. In addition, there was no interaction between text redundancy and questioning, $F(2, 114) = .90$, $p = .41$. Again, there was a main effect of questioning condition, $F(2, 117) = 7.65$, $p = .001$. Participants in the elaborative-interrogation condition rated their lessons as more difficult relative to the ratings of participants in the read-twice condition, $p = .001$, Cohen's $d = .88$. Participants in the embedded questioning also rated their lessons as more difficult than participants in the read-twice condition, but this difference was not reliable ($p = .06$).

Time with Lesson

Figure 3 presents the mean number of minutes spent on the lesson by condition (means reported are not square-root transformed in order to provide the reader with meaningful descriptive statistics). There was no effect of diagram presence on the number of minutes spent with the lesson, $F(1, 196) = 1.40$, $p = .24$, and no interaction between diagram presence and questioning

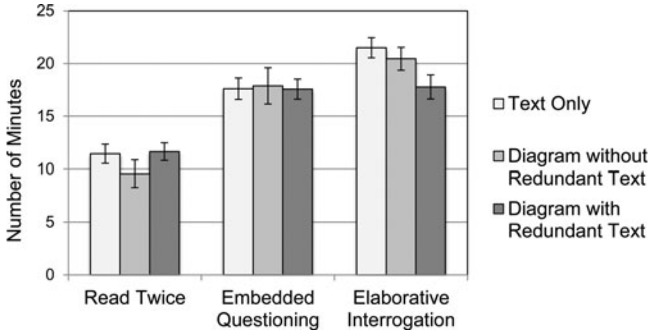

FIGURE 3 Number of minutes with the lesson by condition (means with standard error bars; $N = 198$).

condition, $F(2, 192) = .97$, $p = .38$. However, there was a main effect of questioning condition ($F(2, 195) = 68.21$, $p < .001$). Participants in the elaborative-interrogation condition spent more time with the lesson than did participants in either the read-twice condition ($p < .001$, Cohen's $d = 1.48$), or the embedded-questioning condition ($p = .001$, Cohen's $d = .40$). In addition, participants in the embedded-questioning condition spent more time with the lesson than did participants in the read-twice condition ($p < .001$, Cohen's $d = 1.08$).

When examining the effects of redundant text on the time spent with lessons with diagrams, the findings are similar to those for difficulty ratings. Redundant text did not affect time spent with the lesson ($F(1, 119) = .09$, $p = .77$). In addition, there was not a reliable interaction between redundancy and questioning ($F(2, 114) = 2.95$, $p = .06$). As with previous analyses, there was main effect for questioning ($F(2, 117) = 41.41$, $p < .001$). Participants in the elaborative-interrogation condition spent more time with their lessons than either participants in the read-twice condition ($p < .001$, Cohen's $d = 1.41$) or participants in the embedded-questioning condition ($p < .001$, Cohen's $d = .25$). In addition, participants in the embedded-questioning condition spent more time with their lessons than did participants in the read-twice condition ($p = .001$, Cohen's $d = 1.16$).

Overall, based on difficulty ratings and time spent with the lessons, diagrams had little effect on making the lessons easier to understand. However, the lesson-difficulty ratings and number of minutes spent on the lesson clearly indicate that answering elaborative interrogations while reading was difficult and time consuming for students. In addition, reading twice, although intended to control for time with the lesson to a certain degree relative to elaborative interrogation, was clearly not an appropriate control for time on task. Had elaborative interrogation yielded a benefit relative to reading twice, time with the lesson would have been an important variable to include in the analyses of learning measures.

Answer Quality for Questioning Conditions

There were two important findings regarding questioning. The first was that reading twice was more beneficial for learning how to calculate posterior probability than was elaborative interrogation. The second was that elaborative interrogation did not foster learning more than embedded questions. To better understand these findings, we examined the quality of the answers to the questions in the elaborative interrogation and embedded questioning conditions (see Scoring section for scoring criteria).

Elaborative Interrogations

Participants received an average score of 4.80 ($SD = 2.27$), out of a maximum possible score of 18, for the answers to elaborative interrogations. Approximately 37% of the answers were missing, circular, or wrong; 51% of the answers were inadequate; 9% of the answers were adequate; and 4% were ideal. We used a Spearman's rank-order correlation to test for associations between accuracy of answers to the elaborative interrogations and the sum of accurate answers on the posttest. There was a positive correlation between elaborative interrogation answer scores and posttest scores ($\rho(62) = .40$, $p = .001$).

Embedded Questions

Participants answered an average of 3.08 embedded questions correctly ($SD = 1.70$) out of 6 questions total. As with the elaborative interrogations, we used a Spearman's rank-order correlation to test for an association between the accuracy of the answers to embedded questions and the posttest scores. There was no association between embedded question answer accuracy and posttest scores ($\rho(54) = .11$, $p = .44$).

DISCUSSION

The purpose of this study was to examine the effects of elaborative interrogations and diagrams on learning from a written lesson on calculating posterior probability. Both elaborative interrogation and diagrams were expected to improve the meaningful connections students made in the lesson, thereby, increasing learning.

Effects of Elaborative Interrogation on Learning

We expected that elaborative interrogation would benefit learning from the lesson, because of an anticipated increase in meaningful connections among ideas in the lesson (e.g., McDaniel & Donnelly, 1996; Ozgungor & Guthrie, 2004). Contrary to this expectation, answering elaborative interrogation questions led to less learning than reading the lesson twice. One possible explanation for this finding is that the quality of the answers students provided was generally poor. According to Jiang and Elen (2011), answering questions while reading improves learning from text only if students answer the questions as they were intended. For the elaborative interrogations in this study, students generally did not provide answers indicative of the connections among ideas in the lesson these questions were intended to stimulate. The overwhelming majority of the answers to the elaborative interrogation questions were missing, circular, wrong, or inadequate. Given these findings regarding answer quality, it is not surprising that elaborative interrogation did not promote learning. These findings are consistent with previous work in which incorrect elaborations for self-explanations while learning were associated with poor problem-solving performance (Berthold & Renkl, 2009).

We also expected that elaborative interrogation would add difficulty to the lesson; however, we anticipated that this difficulty would be desirable and increase learning (Richland et al., 2005). As expected, we found that students reported higher difficulty ratings and spent more time with lessons that incorporated elaborative interrogations. Based on the learning measures, however, this increased difficulty did not appear to be desirable. Thus, the level of difficulty imposed by the elaborative interrogations may be another explanation for our findings. The logical reasoning behind posterior probability is quite complex; students may have been overwhelmed by the elaborative interrogations.

Effects of Diagrams on Learning

In this study, we sought to improve student learning on posterior probability problems by incorporating diagrams into lessons. Based on the multimedia principle, which states that learning is improved when relevant visuals are incorporated into lessons (Mayer, 2009), we expected that including diagrams in the lesson would improve learning. We found a benefit of diagrams—but only for the read-twice condition. These findings are somewhat consistent with the multimedia principle in that the benefits of diagrams were limited to a particular condition. Although the benefits of a diagrams were evident only when reading twice, this benefit was likely due to the connections students generated between the verbal information in the text and the visual information in the diagrams.

We also examined the effect of including text redundant with the diagram. Based on the *redundancy* principle, which states that information should be presented in the text or diagram but not by both (Sweller et al., 2011), one prediction was that omitting text redundant with the diagram would enhance learning. Alternatively, adding text redundant with the diagram could assist students in making connections between the text and diagram in a manner similar to labeling (Mason et al., 2013). Hence, a competing prediction was that including text redundant with the diagram would improve learning more than omitting text redundant with the diagram. However, we did not find a reliable difference for text redundancy for either learning or difficulty. One reason could be that the possible benefits of redundant text evened out the possible negative effects of redundant text, thereby, causing no overall difference.

Elaborative Interrogation with Diagrams

The elaborative interrogation questions were designed to promote connections between the text and diagrams, thereby, increasing learning. However, the elaborative interrogation questions appear to have attenuated any benefit the diagram may have provided. Students may have been too distracted by generating answers to the elaborative interrogation questions to develop meaningful connections between the visual and verbal information.

Pedagogical Implications

Given that some of the findings are inconsistent with previous work (e.g., Ozgungor & Guthrie, 2004; Smith et al., 2010), it is premature to propose implications. However, if this study would be supported by future work with similar findings, there would be implications for educational practice, both for probability instruction and more generally. At a general level, our findings indicate that relevant visual information, such as diagrams, in texts *can* improve learning—though it does not do so under all instructional conditions. In this study, diagrams fostered students' application of the lesson content relative to text alone only if students read the lesson twice and not if they answered questions while reading. Taken together, these findings suggest that visual representations may be beneficial but only if students have adequate resources available to process them. Teachers may find it beneficial to incorporate relevant visual representations in their lessons and to encourage students to construct visual representations when solving problems. However, teachers may also need to consider the resource demands of working with visual representations.

Our findings suggest that elaborative interrogation does not necessarily prompt meaningful connections among ideas in the text that would improve learning of material; however, the results indicate that better answers to elaborative interrogations were associated with better learning from the lesson. Therefore, teachers may find it useful to model appropriate answers and monitor answer quality when using elaborative interrogation with their students.

Limitations and Future Directions

Some limitations of this study should be noted—namely, we did not address the possible influence of background knowledge because the topic in this study, posterior probability, is one with which this population typically has little background knowledge (Morsanyi & Handley, 2012). For this reason, the answers to the elaborative interrogations were specifically designed to not require background knowledge (see Ozgungor & Guthrie, 2004, for similar methodology). However, background knowledge may have provided support to students to help them understand the material and effectively answer the elaborative interrogations. In other words, students may have had the cognitive capacity to provide suitable answers to the elaborative interrogations if the lesson were about a topic about which they had a good amount of background knowledge (Renkl, Atkinson, & Groβe, 2004). A future study on a topic about which students have stronger background knowledge may find that the combined use of elaborative interrogation and diagram enhances learning. In addition, there was only one reliable measure of learning in this study that assessed how to solve the problem (i.e., procedural knowledge) but not why the problem should be solved a particular way (i.e., conceptual knowledge). Given the importance of conceptual knowledge (Crooks & Alibali, 2014), a study in which measures of conceptual learning are included would be informative.

Overall, answer quality for elaborative interrogation questions was quite poor in this study. We propose two approaches to improve answer quality in pedagogical contexts and in future studies. The first is practice training, in which students receive examples of what types of answers are expected of them. This practice training would provide a model for the students, which could improve their answer quality (see Dornisch, Sperling, & Zeruth, 2011, for discussion). The second is to provide students with feedback on their answers—either directly, through their peers while working collaboratively (after training), or through computer-based tutoring. Such feedback has been shown to enhance the effectiveness of answering questions while reading a lesson (Andre & Thieman, 1988; García-Rodicio, 2014; Peverly & Wood, 2001). Moreover, feedback could provide motivation for students to construct high-quality answers to elaborative interrogations

(Biggs & Tang, 2011). If these ideas could effectively improve answer quality, positive effects of elaborative interrogation might be seen.

CONCLUSION

Posterior probability, like many probabilistic reasoning concepts, is frequently challenging for students to learn. In this study, we sought to improve student learning with regard to posterior probability problems by incorporating diagrams and elaborative interrogations in lessons that students read independently. Our findings regarding elaborative interrogations indicate that the difficulty of the material should be considered when using this technique. We believe that poor answer quality is one reason for the negative effects of elaborative interrogation that we observed, and we have suggested some approaches that might improve answer quality. Our finding that diagrams promoted learning is consistent with the multimedia principle (Mayer, 2009). In sum, this work provides some guidance concerning how to improve student learning about probabilistic reasoning, while also highlighting the challenges inherent in this complex domain.

ACKNOWLEDGMENTS

We thank Brittany Ewert and Joanne Xiong for collecting and coding data. We greatly appreciate Pooja Sidney's assistance with data analyses.

FUNDING

This research was supported by the Institute of Education Sciences, U.S. Department of Education, through Grant R305C100024 to the University of Wisconsin–Madison. The opinions expressed are those of the authors and do not represent views of the Institute or of the U.S. Department of Education.

AUTHOR NOTES

Virginia Clinton, Ph.D., is an Instructor in the Department of Psychology at the University of North Dakota. **Martha W. Alibali** is a Professor in the Department of Psychology at the University of Wisconsin-Madison. **Mitchell J. Nathan** is a Professor in the Departments of Educational Psychology, Psychology, and Curriculum and Instruction at the University of Wisconsin-Madison.

REFERENCES

Aleven, V. A., & Koedinger, K. R. (2002). An effective metacognitive strategy: Learning by doing and explaining with a computer-based cognitive tutor. *Cognitive Science*, *26*(2), 147–179. doi: 10.1016/S0364-0213(02)00061-7

Andre, T., & Thieman, A. (1988). Level of adjunct question, type of feedback, and learning concepts by reading. *Contemporary Educational Psychology*, *13*(3), 296–307. doi: 10.1016/0361-476X(88)90028-8

Atkinson, R. K., Derry, S. J., Renkl, A., & Wortham, D. (2000). Learning from examples: Instructional principles from the worked examples research. *Review of Educational Research*, *70*(2), 181–214. doi: 10.3102/00346543070002181

Atkinson, R. K., Renkl, A., & Merrill, M. M. (2003). Transitioning from studying examples to solving problems: Effects of self-explanation prompts and fading worked-out steps. *Journal of Educational Psychology*, *95*(4), 774. doi: 10.1037/0022-0663.95.4.774

Bar-Hillel, M. (1980). The base-rate fallacy in probability judgments. *Acta Psychologica*, *44*(3), 211–233. doi: 10.1016/0001-6918(80)90046-3

Bates, D. M. (2010). *lme4: Mixed-effects modeling with R*. Retrieved from http://lme4.r-forge.r-project.org/book

Bates, D., Maechler, M., & Bolker, B. (2012). *lme4: Linear mixed-effects models using S4 classes*. Retrieved from http://cran.R-project.org/package=lme4.

Beitzel, B. D., & Staley, R. K. (2015). The efficacy of using diagrams when solving probability word problems in college. *Journal of Experimental Education*, *83*(1), 130–145. doi: 10.1080/00220973.2013.876232

Berthold, K., Eysink, T. H., & Renkl, A. (2009). Assisting self-explanation prompts are more effective than open prompts when learning with multiple representations. *Instructional Science*, *37*(4), 345–363. doi: 10.1007/s11251-008-9051-z

Berthold, K., & Renkl, A. (2009). Instructional aids to support a conceptual understanding of multiple representations. *Journal of Educational Psychology*, *101*(1), 70. doi: 10.1037/a0013247

Beyth-Marom, R., Fidler, F., & Cumming, G. (2008). Statistical cognition: Towards evidence-based practice in statistics and statistics education. *Statistics Education Research Journal*, *7*(2), 20–39. Retrieved from https://www.stat.auckland.ac.nz/~iase/publications.php?show=serj

Biggs, J., & Tang, C. (2011). *Teaching for quality learning at university*. London, UK: McGraw-Hill International.

Bramwell, R., West, H., & Salmon, P. (2006). Health professionals' and service users' interpretation of screening test results: Experimental study. *BMJ*, *333*(7562), 284. doi: 10.1136/bmj.38884.663102.AE

Chandler, P., & Sweller, J. (1991). Cognitive load theory and the format of instruction. *Cognition and Instruction*, *8*(4), 293–332. doi: 10.1207/s1532690xci0804_2

Chiu, J. L., & Chi, M. T. H. (2014). Supporting self-explanation in the classroom. In V. A. Benassi, C. E. Overson, & C. M. Hakala (Eds.), *Applying science of learning in education: Infusing psychological science into the curriculum*. Retrieved from http://bit.ly/KbYLtG

Crooks, N. M., & Alibali, M. W. (2014). Defining and measuring conceptual knowledge in mathematics. *Developmental Review*, *34*(4), 344–377. doi: 10.1016/j.dr.2014.10.001

Cromley, J. G., Snyder-Hogan, L. E., & Luciw-Dubas, U. A. (2010). Cognitive activities in complex science text and diagrams. *Contemporary Educational Psychology*, *35*(1), 59–74. doi: 10.1016/j.cedpsych.2009.10.002

Dornisch, M., Sperling, R. A., & Zeruth, J. A. (2011). The effects of levels of elaboration on learners' strategic processing of text. *Instructional Science*, *39*(1), 1–26. doi: 10.1007/s11251-009-9111-z

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

Garcia-Retamero, R., & Hoffrage, U. (2013). Visual representation of statistical information improves diagnostic inferences in doctors and their patients. *Social Science & Medicine*, *83*, 27–33.

García-Rodicio, H. (2014). Support for learning from multimedia explanations: A comparison of prompting, signaling, and questioning. *Journal of Educational Computing Research*, *50*(1), 29–43. doi: 10.2190/EC.50.1.b

Garfield, J., & Ben-Zvi, D. (2008). *Developing students' statistical reasoning: Connecting research and teaching practice*. New York, NY: Springer.

Gelman, A., Hill, J., & Yajima, M. (2012). Why we (usually) don't have to worry about multiple comparisons. *Journal of Research on Educational Effectiveness*, *5*(2), 189–211.

Heuer, R. J. (1999). *Psychology of intelligence analysis*. Central Intelligence Agency: Center for the Study of Intelligence. Retrieved from https://www.cia.gov/library/center-for-the-study-of-intelligence/csi-publications/books-and-monographs/psychology-of-intelligence-analysis/art15.html

Holliday, W. G. (1976). Teaching verbal chains using flow diagrams and texts. *AV Communication Review*, *24*, 63–78. doi: 10.1007/BF02768332

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language*, *59*(4), 434–446. doi: 10.1016/j.jml.2007.11.007

Jiang, L., & Elen, J. (2011). Instructional effectiveness of higher-order questions: The devil is in the detail of students' use of questions. *Learning Environments Research*, *13*(3), 279–298. doi: 10.1007/s10984-011-9095-x

Johnson, C. I., & Mayer, R. E. (2012). An eye movement analysis of the spatial contiguity effect in multimedia learning. *Journal of Experimental Psychology: Applied*, *18*(2), 178. doi: 10.1037/a0026923

Kahneman, D., & Tversky, A. (1972). Subjective probability: A judgment of representativeness. *Cognitive Psychology*, *3*(3), 430–454. doi: 10.1016/0010-0285(72)90016-3

Keeler, C., & Steinhorst, K. (2001). A new approach to learning probability in the first statistics course. *Journal of Statistics Education*, *9*(3), 1–24. Retrieved from http://www.amstat.org/publications/JSE/v9n3/keeler.html

Kim, J. Y., & Anderson, T. (2011). Reading across the curriculum: A framework for improving the reading abilities and habits of college students. *Journal of College Literacy & Learning*, *37*, 29–40. Retrieved from http://www.j-cll.com/files/37_Kim_Anderson.pdf

Kintsch, W. (1998). *Comprehension: A paradigm for cognition.* New York, NY: Cambridge University Press.

Konheim-Kalkstein, Y. L. (2008). *Facilitation of Bayesian decision making* (Unpublished doctoral dissertation). University of Minnesota, Twin Cities, Minneapolis, MN.

Leppink, J., Broers, N. J., Imbos, T., van der Vleuten, C. P., & Berger, M. P. (2012). Self-explanation in the domain of statistics: An expertise reversal effect. *Higher Education*, *63*(6), 771–785. doi: 10.1007/s10734-011-9476-1

Martin, V. L., & Pressley, M. (1991). Elaborative-interrogation effects depend on the nature of the question. *Journal of Educational Psychology*, *83* (1), 113. doi: 10.1037/0022-0663.83.1.113

Mason, L., Pluchino, P., & Tornatora, M. C. (2013). Effects of picture labeling on science text processing and learning: Evidence from eye movements. *Reading Research Quarterly*, *48*(2), 199–214. doi: 10.1002/rrq.41

Mayer, R. E. (1999). Designing instruction for constructivist learning. In C. M. Reigeluth (Eds.), *Instructional-design theories and models: A new paradigm of instructional theory*, (Vol. 2, pp. 141–159). Mahwah, NJ: Erlbaum.

Mayer, R. E. (2002). Multimedia learning. *Psychology of Learning and Motivation*, *41*, 85–139. doi: 10.1016/S0079-7421(02)80005-6

Mayer, R. E. (2009). *Multimedia learning* (2nd ed.). New York, NY: Cambridge University Press.

Mayer, R. E., & Gallini, J. K. (1990). When is an illustration worth ten thousand words? *Journal of Educational Psychology*, *82*(4), 715. doi: 10.1037/0022-0663.82.4.715

Mayer, R. E., & Johnson, C. I. (2008). Revising the redundancy principle in multimedia learning. *Journal of Educational Psychology*, *100*(2), 380. doi: 10.1037/0022-0663.100.2.380

Mayer, R. E., Steinhoff, K., Bower, G., & Mars, R. (1995). A generative theory of textbook design: Using annotated illustrations to foster meaningful learning of science text. *Educational Technology Research and Development*, *43*(1), 31–41. doi: 10.1007/BF02300480

McCrudden, M. T., Hushman, C. J., & Marley, S. C. (2014). Exploring the boundary conditions of the redundancy principle. *Journal of Experimental Education*, *82* (4). 537–554. doi: 10.1080/00220973.2013.813368

McDaniel, M. A., & Butler, A. C. (2011). A contextual framework for understanding when difficulties are desirable. In A. S. Benjamin (Ed.), *Successful remembering and successful forgetting: A festschrift in honor of Robert A. Bjork* (pp. 175–198). New York, NY: Psychology Press.

McDaniel, M. A., & Donnelly, C. M. (1996). Learning with analogy and elaborative interrogation. *Journal of Educational Psychology*, *88*(3), 508. doi: 10.1037/0022-0663.88.3.508

McMaster, K. L., van den Broek, P., Espin, C. A., White, M. J., Rapp, D. N., Kendeou, P., Bohn-Gettler, C. M., & Carlson, S. (2012). Making the right connections: Differential effects of reading intervention for subgroups of comprehenders. *Learning and Individual Differences*, *22*(1), 100–111. doi: 10.1016/j.lindif.2011.11.017

McNamara, D. S., Levinstein, I. B., & Boonthum, C. (2004). iSTART: Interactive strategy training for active reading and thinking. *Behavior Research Methods, Instruments, & Computers*, *36*(2), 222–233. doi: 10.3758/BF03195567

Morsanyi, K., & Handley, S. (2012). Does thinking make you biased? The case of the engineers and lawyer problem. *Proceedings of the Annual Meeting of the Cognitive Science Society*, *34*, 2049–2054.

Ortegren, F. R., Serra, M. J., & England, B. D. (2015). Examining competing hypotheses for the effects of diagrams on recall for text. *Memory & Cognition*, *43*(1), 70–84. doi: 10.3758/s13421-014-0429-7

Osborne, J. (2002). Notes on the use of data transformations. *Practical Assessment, Research & Evaluation*, *8*(6), 1–8. Retrieved from http://PAREonline.net/getvn.asp?v=8&n=6

Ozgungor, S., & Guthrie, J. T. (2004). Interactions among elaborative interrogation, knowledge, and interest in the process of constructing knowledge from text. *Journal of Educational Psychology*, *96*, 437–443. doi: 10.1037/0022-0663.96.3.437

Pease, R. S. (2012). *Using elaborative interrogation enhanced worked examples to improve chemistry problem solving* (Unpublished doctoral dissertation). University of Maryland, College Park, MD.

Peverly, S. T., & Wood, R. (2001). The effects of adjunct questions and feedback on improving the reading comprehension skills of learning-disabled adolescents. *Contemporary Educational Psychology*, *26*(1), 25–43. doi: 10.1006/ceps.1999.1025

Pressley, M., McDaniel, M. A., Turnure, J. E., Wood, E., & Ahmad, M. (1987). Generation and precision of elaboration: Effects on intentional and incidental learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *13*(2), 291. doi: 10.1037/0278-7393.13.2.291

Pressley, M., Symons, S., McDaniel, M. A., Snyder, B. L., & Turnure, J. E. (1988). Elaborative interrogation facilitates acquisition of confusing facts. *Journal of Educational Psychology*, *80* (3), 268. doi: 10.1037/0022-0663.80.3.268

Renkl, A. (2005). The worked-out example principle in multimedia learning. In R. E. Mayer (Ed.), *Cambridge handbook of multimedia learning* (pp. 229–247). Cambridge, UK: Cambridge University Press.

Renkl, A., Atkinson, R. K., & Große, C. S. (2004). How fading worked solution steps works: A cognitive load perspective. *Instructional Science*, *32*(1–2), 59–82. doi: 10.1023/B:TRUC.0000021815.74806.f6

Richland, L. E., Bjork, R. A., Finley, J. R., & Linn, M. C. (2005). Linking cognitive science to education: Generation and interleaving effects. In B. G. Bara, L. Barsalou, & M. Bucciarelli (Eds.), *Proceedings of the twenty-seventh annual conference of the Cognitive Science Society* (pp. 1850–1855). Mahwah, NJ: Erlbaum.

Rittle-Johnson, B. (2006). Promoting transfer: Effects of self-explanation and direct instruction. *Child Development*, *77*(1), 1–15. doi: 10.1111/j.1467-8624.2006.00852.x

Rouet, J. F., & Vidal-Abarca, E. (2002). Mining for meaning: Cognitive effects of inserted questions in learning from scientific text. In J. Otero, J. Leon, A. C. Graesser (Eds.), *The psychology of science text comprehension* (pp. 417–436). Mahwah, NJ: Erlbaum.

Royer, J. M. (2001). Developing reading and listening comprehension tests based on the Sentence Verification Technique (SVT). *Journal of Adolescent & Adult Literacy*, *45*(1), 30–41. Retrieved from http://www.jstor.org/stable/40007629

Schüler, A., Scheiter, K., Rummer, R., & Gerjets, P. (2012). Explaining the modality effect in multimedia learning: Is it due to a lack of temporal contiguity with written text and pictures? *Learning and Instruction*, *22*(2), 92–102. doi: 10.1016/j.learninstruc.2011.08.001

Schwonke, R., Berthold, K., & Renkl, A. (2009). How multiple external representations are used and how they can be made more useful. *Applied Cognitive Psychology*, *23*(9), 1227–1243. doi: 10.1002/acp.1526

Sedlmeier, P., & Gigerenzer, G. (2001). Teaching Bayesian reasoning in less than two hours. *Journal of Experimental Psychology: General*, *130*, 380–400.

Smith, B. L., Holliday, W. G., & Austin, H. W. (2010). Students' comprehension of science textbooks using a question-based reading strategy. *Journal of Research in Science Teaching*, *47*(4), 363–379. doi: 10.1002/tea.20378

Snijders, T., & Bosker, R. (2012), *Multilevel analysis: An introduction to basic and applied multilevel analysis (*2nd *ed.)*. London, UK: Sage.

Stanovich, K. E., & West, R. F. (1998). Individual differences in rational thought. *Journal of Experimental Psychology: General*, *127*(2), 161. doi: 10.1037/0096-3445.127.2.161

Sweller, J., Ayres, P., & Kalyuga, S. (2011). *Cognitive load theory*. Dordrecht, The Netherlands: Springer.

Tarmizi, R. A., & Sweller, J. (1988). Guidance during mathematical problem solving. *Journal of Educational Psychology*, *80*(4), 424–436. doi: 10.1037/022-0663.80.4.424

Wiley, J., & Voss, J. F. (1999). Constructing arguments from multiple sources: Tasks that promote understanding and not just memory for text. *Journal of Educational Psychology*, *91*(2), 301. doi: 10.1037/0022-0663.91.2.301

Willoughby, T., Waller, T. G., Wood, E., & MacKinnon, G. E. (1993). The effect of prior knowledge on an immediate and delayed associative learning task following elaborative interrogation. *Contemporary Educational Psychology*, *18*, 36–36. doi: 10.1006/ceps.1993.1005

Woloshyn, V. E., Pressley, M., & Schneider, W. (1992). Elaborative-interrogation and prior-knowledge effects on learning of facts. *Journal of Educational Psychology*, *84*(1), 115. doi: 10.1037/0022-0663.84.1.115

Zieffler, A., Garfield, J., Alt, S., Dupuis, D., Holleque, K., & Chang, B. (2008). What does research suggest about the teaching and learning of introductory statistics at the college level? A review of the literature. *Journal of Statistics Education*, *16*(2), 1–23.

APPENDIX: LESSON EXCERPTS

Text Only

It's helpful to imagine 100 patients that come to the walk-in clinic. According to the base rate, we know that 10 of the patients will have sepsis and 90 will not. According to the specific case information, we know that 80%, or 8 of the 10 patients, with sepsis will have high fever, chills, and skin lesions and that 20%, or 2 of the 10 patients, will not. Also, of the 90 patients who don't have sepsis, 10%, or 9, will have high fever, chills, and skin lesions and 90%, or 81, will not have these symptoms.

This makes a total of 17 patients with high fever, chills, and skin lesions and 83 patients without these symptoms, of which 8 of the 17 actually have sepsis; the other 9 have the symptoms without actually having sepsis. Therefore, when a doctor sees a patient with the symptoms of sepsis (high fever, chills, and skin lesions), the probability that the patient actually has sepsis is only 8/17ths, or 47%.

Diagram without Redundant Text

It's helpful to imagine 100 patients that come to the walk-in clinic. According to the base rate, we know how many of these patients will have sepsis and how many will not. According to the specific case information, we know that 80% of patients with sepsis will have high fever, chills, and skin lesions and 20% of the patients with sepsis will not. Also, of the 90 patients who don't have sepsis, 10% will have high fever, chills, and skin lesions and 90% will not have these symptoms. This is shown in the diagram.

This makes a total of 17 patients with high fever, chills, and skin lesions; but not all of these patients have sepsis. Therefore, when a doctor sees a patient with the symptoms of sepsis (high fever, chills, and skin lesions), the probability that the patient actually has sepsis is only 8/17ths or 47%.

Diagram with Redundant Text (Redundant Frequency Information in Bold)

It's helpful to imagine 100 patients that come to the walk-in clinic. According to the base rate, we know that **10** of the patients will have sepsis and **90** will not. According to the specific case information, we know that 80%, or **8** of the **10** patients, with sepsis will have high fever, chills, and skin lesions and 20%, or **2** of the **10** patients, will not. Also, of the **90** patients who don't have sepsis, 10%, or **9**, will have high fever, chills, and skin lesions and 90%, or **81**, will not have these symptoms. This is shown in the diagram.

This makes a total of 17 patients with high fever, chills, and skin lesions and **83** patients without these symptoms, of which **8** of the **17** actually have sepsis; the other **9** have the symptoms without actually having sepsis. Therefore, when a doctor sees a patient with the symptoms of sepsis (high fever, chills, and skin lesions), the probability that the patient actually has sepsis is only 8/17ths, or 47%.