Expecting to teach enhances learning and organization of knowledge in free recall of text passages

**John F. Nestojko · Dung C. Bui · Nate Kornell ·
Elizabeth Ligon Bjork**

Abstract The present research assessed the potential effects
of expecting to teach on learning. In two experiments, partic-
ipants studied passages either in preparation for a later test or
in preparation for teaching the passage to another student who
would then be tested. In reality, all participants were tested,
and no one actually engaged in teaching. Participants
expecting to teach produced more complete and better orga-
nized free recall of the passage (Experiment 1) and, in general,
correctly answered more questions about the passage than did
participants expecting a test (Experiment 1), particularly ques-
tions covering main points (Experiment 2), consistent with
their having engaged in more effective learning strategies.
Instilling an expectation to teach thus seems to be a simple,
inexpensive intervention with the potential to increase learn-
ing efficiency at home and in the classroom.

When trying to learn new information, we typically do so with
goals or expectations for how we will use that information in
the future. Students, for example, typically have the goal of
maximizing their performance on a later test when learning
new material. In contrast, teachers presumably have the goal of
being able to effectively communicate the new material they
are learning to their students. The present research investigated

whether these different orientations—expecting a test or
expecting to teach—changes how new material is learned.

General effects of expectancy on learning and memory

Learners' expectations about how they will be tested can produce
both quantitative and qualitative effects on their later memory
performance. For example, Szpunar, McDermott, and Roediger
(2007) demonstrated that participants expecting a final cumula-
tive test achieved higher free recall test scores (quantitative
advantage) and greater organization of recall output (qualitative
advantage) on a final cumulative test than did participants not
expecting a final cumulative test. Additionally, in research on so-
called test-expectancy effects, participants expecting a free recall
test typically performed better on both free recall and recognition
tests than did participants expecting a recognition test (e.g.,
Lundeberg & Fox, 1991). Furthermore, while exploring a pro-
cess that he dubbed cognitive tuning, Zajonc (1960) asked
participants to read a letter, telling one group ("transmitters") that
they would have to relay the information in the letter to another
person, while telling another group ("receivers") that they would
receive additional details from another person about the letter. On
a later test about the contents of the letter, transmitters
outperformed receivers on multiple measures, including one
reflecting organization of information. In sum, learners' expec-
tations as to how they will later need to utilize to-be-learned
information can apparently affect how they encode and recall that
information.

Is learning affected by expecting to teach versus expecting to be tested?

Prior research has shown that when a student teaches/tutors
one or more other students, the teaching student often shows

J. F. Nestojko (✉) · D. C. Bui
Washington University in St. Louis, One Brookings Drive,
Campus Box 1125, St. Louis, MO 63130, USA
e-mail: nestojko@artsci.wustl.edu

E. L. Bjork
University of California, Los Angeles, Los Angeles, CA, USA

N. Kornell
Williams College, Williamstown, MA, USA

learning gains (e.g., Cohen, Kulik, & Kulik, 1982; Rohrbeck, Ginsburg-Block, Fantuzzo, & Miller, 2003; Roscoe & Chi, 2007). The research presented herein addressed a different question: Does having an expectancy to teach, without actually teaching, enhance learning, as compared with having an expectancy to be tested? Prior research on this question has produced inconsistent results. Some studies have reported advantages of expecting to teach over expecting a test (e.g., Bargh & Schul, 1980; Benware & Deci, 1984; Coleman, Brown, & Rivkin, 1997; Ehly, Keith, & Bratton, 1987; Fiorella & Mayer, 2013), whereas others have reported no differences between these conditions (Renkl, 1995; Ross & DiVesta, 1976).

In examining the past literature, we decided upon five criteria as being necessary for a study to comprise an adequate assessment of an effect on learning owing to a test expectation versus a teaching expectation: (1) Instructions manipulating learning goals must be given prior to exposure to the to-be-learned material; (2) duration of exposure to the to-be-learned material must be held constant across conditions; (3) all conditions must be given the same opportunities and suggestions about how to learn (e.g., if one group is allowed or encouraged to take notes, the other group must also be allowed or encouraged to take notes); (4) the retention interval between studying and the final criterion test must be held constant; and (5) no actual teaching can take place after study and prior to measurements of learning. Many of the existing reports involving a teaching-expectancy condition failed to meet all five criteria, making it difficult to attribute advantages seen in teaching-expectancy conditions solely to expectancy (e.g., Benware & Deci, 1984; Coleman et al., 1997; Ehly et al., 1987; Gregory, Walker, McLaughlin, & Peets, 2011).

Four reports met our five criteria. In a laboratory-based experiment by Bargh and Schul (1980), participants were instructed—prior to reading a text passage—either to study for a test or to prepare to teach another student. On the later criterion tests, participants instructed to prepare to teach performed better than did participants instructed to study for a test. Fiorella and Mayer (2013) found that participants expecting to teach a lesson on the Doppler effect outperformed control subjects on an immediate comprehension test, but not on a delayed comprehension test. This finding should be interpreted with caution, however, because participants were allowed to take notes while studying the passage, and note quantity is positively correlated with recall even when participants are not allowed to study their notes afterward (e.g., Aiken, Thomas, & Shennum, 1975; Bui, Myerson, & Hale, 2013). Fiorella and Mayer did not report the rate of note taking, but if expecting to teach prompted more note taking than did expecting to take a test, then it is not clear whether expecting to teach had a direct benefit on learning or only an indirect one by increasing note taking.

Ross and DiVesta (1976) manipulated whether participants expected (1) to deliver an oral explanation or summary of the to-be-studied material to another student or (2) to take both a short-answer and multiple-choice test on both general concepts and specific details of the to-be-studied materials. The participants expecting to give an explanation or summary (who were never actually required to perform either task but, instead, were given both types of tests) performed no better on these tests, as compared with those participants who had expected to take those specific tests. Finally, in an experiment by Renkl (1995) in which participants were trained to perform probability calculus via examination of worked-out problems, those instructed before training that they would later have to explain similar problems to another student were no better at solving similar problems on a later test than were the participants instructed that they would later have to solve similar problems. Thus, the four prior studies that met our criteria were evenly split: Two showed that expecting to teach enhanced learning, as compared with expecting a test, and two did not.

Predictions with respect to the present research question

As we started the present research, we had several reasons to predict that expecting a test would produce better later performance than would expecting to teach. First, students at universities (who served as our participants) are arguably more experienced at preparing to take an exam than they are at preparing to teach and, most likely, possess minimal developed skills related to teaching. Second, we thought it possible that expecting to teach might have a negative impact on learning because public speaking is a major source of anxiety for many people (Motley, 1988; Richmond & McCroskey, 1995; Ruscio et al., 2008) and anxiety is also associated with having to teach (e.g., Ameen, Guffey, & Jackson, 2002; Gardner & Leak, 1994). Because high levels of anxiety can be detrimental to performance on many tasks (Yerkes & Dodson, 1908), the expectation of having to teach could produce high anxiety in some participants, disrupting their learning and later performance. Third, violating participants' expectations—as happens when participants are first told that they will have to teach and then told that, in fact, they are going to take a test instead—could lead to a "double-cross" effect in which participants lower their effort on a final test in the condition in which their expectations were violated, although it might also be the case that they would look upon not having to teach as a relief and, thus, not be so affected.

In contrast to reasons for why expecting a test might produce better later performance, we also had reasons for predicting the opposite outcome—namely, that expecting to teach would produce better performance on a later test than would expecting to be tested. Teaching typically involves activities that are known to enhance learning, such as summarizing the critical points of the to-be-learned information,

identifying key concepts, seeking relationships among ideas, and mentally organizing the information (e.g., McKeachie, Pintrich, Lin, & Smith, 1986). If expecting to teach causes students to implement these or other effective learning strategies, their learning could be enhanced whether or not they actually go on to teach the material.

It is, of course, also possible that the two types of expectancy could yield similar outcomes, particularly if students expecting a test have learned to prepare for tests by focusing on main points, mentally organizing the to-be-learned material, and, in general, by engaging in the very encoding strategies we have suggested are likely to emerge from a teaching-expectancy mental set.

The present research

In the present research, we sought to answer two main questions. First, does expecting to teach enhance later test performance, as compared with expecting a test? Owing to the equivocal pattern of results emerging from the literature relevant to our research questions, we first attempted to replicate the previous observed benefits of expecting to teach (e.g., Bargh & Schul, 1980).

Second, does expecting to teach cause individuals to change the way they process information? We conjectured that expecting to teach might encourage beneficial encoding activities such as organizational processing (for similar arguments, see Bargh & Schul, 1980; Gartner, Kohler, & Riessman, 1971). Thus, in Experiment 1, we examined the previously untested prediction that expecting to teach would lead participants to produce relatively well-organized free recall responses. We also thought it possible that expecting to teach would cause participants to remember main points especially well, on the basis of observations that teachers often focus on key concepts in learning material (McKeachie et al., 1986). To assess this potential difference, Experiments 1 and 2 included test questions relating to main points, as well as detail points about the passage. In sum, we used measures of output organization and comparisons of memory for main and detail points with the assumption that such measures could provide insights into whether and, if so, how expecting to teach alters the study strategies and cognitive processes used by individuals during learning. The general design we employed in both Experiment 1 and Experiment 2 followed the approach introduced by Bargh and Schul (1980). All participants were given a passage to read for a specified amount of time. Prior to reading, participants were told either that they would later take a test on the contents of the passage or that they would later teach the contents of the passage to another student. Participants given the expectancy to teach, however, never actually did so; instead, they were given the same test as the participants given the expectancy of taking a test.

Experiment 1

After being told that they would either teach or take a test, participants read a passage and then took a free recall test. This test allowed us to quantify the degree to which participants' responses were coherently organized. Additionally, we measured participants' focus on key ideas in the passage by using a passage in which idea units had previously been categorized as main points, important details, or unimportant details (Rawson & Kintsch, 2005). After the free recall test, participants took a short-answer test consisting of questions about important and unimportant details.

Method

Participants and design

A total of 56 undergraduate students (34 females and 22 males; average age = 20.8 years, $SD = 3.27$) from the University of California, Los Angeles, served as participants for course credit. Participants were asked two questions at the start of the experiment to assess their prior knowledge of the to-be-learned materials. None reported having seen the movie *The Charge of the Light Brigade.* Although 2 participants (both in the test-expectancy condition) reported having knowledge of the Crimean War, their performance did not differ from that of their peers; thus, their data were not excluded from our analyses.

We employed a 2×3 mixed design with expectancy instructions (teaching-expectancy vs. test-expectancy) manipulated between subjects and information type (main points, important details, unimportant details) manipulated within subjects. Multiple measures of performance—proportion correct recall, organization, efficiency, and retention of different types of information—were derived from the free recall and short-answer tests we administered to all participants.

Materials and measures

We employed a 1,541-word passage (acquired from Rawson & Kintsch, 2005) comparing the depiction of the Crimean War in the movie *The Charge of the Light Brigade* (Curtiz, 1936) with the actual historical events of that war as a demonstration of the tendency of movies to portray historical events inaccurately. We chose this passage because Rawson and Kintsch had identified 125 idea units (IUs) in the passage, including a subset of 39 that they identified as main points (8 IUs), important details (15 IUs), and unimportant details (16 IUs; for additional details, see Rawson & Kintsch, 2005, p. 74). We used the full set of 125 IUs for most analyses, and the subset of 39 IUs to analyze levels of knowledge.

We employed two measures to assess our hypothesis regarding organization of free recall. First, we measured the

organization of our participants' recall in terms of IUs by adapting the adjusted ratio of clustering measure (ARC; Roenker, Thompson, & Brown, 1971) to our purposes. The ARC was originally intended for use in the context of learning lists of items belonging to different taxonomic categories (e.g., *fruits*, *trees*), where it indicates the extent to which learners cluster their recall output in terms of the categorical memberships of the list items. Broadly speaking, the ARC measures the degree to which a learner's recall patterns reflect the conceptual organization of the studied material. Typically, the bases for computing ARC measures is to count the number of repetitions occurring in the recall output, where a repetition is defined as two items belonging to the same category being output sequentially (e.g., *orange*, *banana*; for a complete explanation of how to compute ARC scores, see Roenker et al., 1971). With the present materials, we considered each IU to correspond to an "item" (i.e., category member) in the ARC analysis and each paragraph within the original passage to correspond to a "category." Thus, a repetition occurred when a participant contiguously recalled two IUs that had been presented in the same paragraph of the passage. To illustrate, if a paragraph in the original passage contained six IUs and a participant output those six units together (even if not in the same order as in the original passage), the ARC score for that participant would be higher than the ARC score for a participant who output the same exact information scattered throughout his or her recall output. Perfect overlap between the structure of the original text passage and the participants' output produces ARC = 1.0. Thus, this ARC measure reflects the degree to which participants' output of IUs corresponds to how those units were originally organized into paragraphs in the source passage.

Latent semantic analysis (LSA; Landauer, Foltz, & Laham, 1998) was the second measure of organization employed. LSA measures the sentence-to-sentence coherence of a passage of text by comparing the relations among terms within a sentence with a large database of terms, then by comparing pairs of sentences in sequences from text (see the Sentence Comparison tool at www.lsa.colorado.edu). Perfect overlap between all sequential pairs of sentences produces LSA = 1.0. The LSA sentence-to-sentence coherence score for the source passage used in the present experiment is .22. LSA sentence-to-sentence coherence scores were computed for each participant's recall output, and the means across participants were averaged to obtain the test-expectancy and teaching-expectancy group means. LSA provides a measure of the coherence of recall output that is independent of the structure of the passage participants read.

Our short-answer test (also acquired from Rawson & Kintsch, 2005) consisted of 8 questions asking about important details of the passage and 10 asking about unimportant details. One question was presented per page in a Microsoft Word document.

Procedure

Experiment 1 consisted of four phases: study, distractor task, free recall test, and short-answer test.

Study phase After collecting demographic information, we told participants that they would have 10 min to read a passage. They were told that they could not take notes or highlight or underline items but that they could read the passage at their own pace and return to previously read parts of the passage whenever they wanted. Additionally, participants were told that they would not have access to this passage once the 10-min reading time expired. Finally, before handing out the passage, participants in the test-expectancy condition were told that they would later be given a test on the material in the passage, whereas participants in the teaching-expectancy condition were told that they would be teaching the material in the passage to another participant, who would then be asked to take a test on the passage.

Distractor task Following the study phase, participants engaged in a distractor task for 25 min (a separate memory experiment using categorized word lists with no overlap with the present materials).

Test phase Following the distractor phase, participants in the test-expectancy condition were told—consistent with their expectation—that they were now going to be tested on the previously read passage. Participants in the teaching-expectancy condition were asked to take a test—inconsistent with their expectation—because the student they were supposed to teach had failed to show up.

Free recall test All participants first received a free recall test for the studied passage. Specifically, they were asked to type as much information from the passage as they could recall onto a blank document in Microsoft Word. They were given unlimited time to recall the information in any format (e.g., paragraph form, bullet point, etc.) and were told to inform the experimenter when they were done. For each participant, the experimenter recorded the amount of time spent on this test.

Short-answer test Immediately following the free recall test, the experimenter opened a Microsoft Word document containing the short-answer test. All participants were instructed to type their responses directly into the document, to proceed through the test at their own pace, and to inform the experimenter when they were finished.

Results

Effect sizes for comparisons of means are reported as Cohen's *d* calculated using the pooled standard deviation of the groups

being compared (Olejnik & Algina, 2000, Box 1, Option B). Effect sizes for analyses of variance (ANOVAs) are reported as $\omega^2_{partial}$ calculated using the formulae provided by Maxwell and Delaney (2004).

Free recall test

Four aspects of performance were assessed via the free recall data: amount of correct output (i.e., proportion correct recall), output efficiency, output organization, and type of information recalled.

Two independent raters, blind to conditions, scored the free recall tests; reliability was high between the two raters ($\alpha = .81$). Participants were given a full point for recall of an entire idea unit, half a point for partial recall of that idea unit, and zero for no recall. Discrepancies in scoring were resolved by a third rater who was also blind to the conditions.

Proportion of IUs correctly recalled As is indicated in Fig. 1a, participants expecting to teach produced a greater proportion of correct IUs ($M = .17$, $SD = .07$) than did participants

expecting a test ($M = .13$, $SD = .08$), and this difference was significant, $t(54) = 2.08$, $p = .043$, $d = 0.56$, $CI_{mean\ difference}$ [0.001, 0.079].

Efficiency of free recall Participants initially instructed to prepare to teach spent slightly less time typing their output ($M = 15.59$ min, $SD = 7.15$) than did participants instructed to prepare for a test ($M = 17.00$ min, $SD = 10.83$), but this difference did not reach significance, $t < 1.00$. In addition to measuring the total time taken by each participant in recalling the passage, we also measured the efficiency of each participant's recall by dividing the number of idea units recalled by the total time spent recalling. This measure (# IUs/min), which is plotted in Fig. 1b, revealed that the teaching-expectancy group recalled information more efficiently ($M = 1.50$, $SD = 0.64$) than did the test-expectancy group ($M = 1.02$, $SD = .57$), $t(54) = 2.94$, $p = .005$, $d = 0.79$, $CI_{mean\ difference}$ [0.151, 0.802].

Organization of free recall As indicated in Fig. 1c, expecting to teach enhanced the organization of participants' free recall. An independent samples t-test indicated a higher ARC score
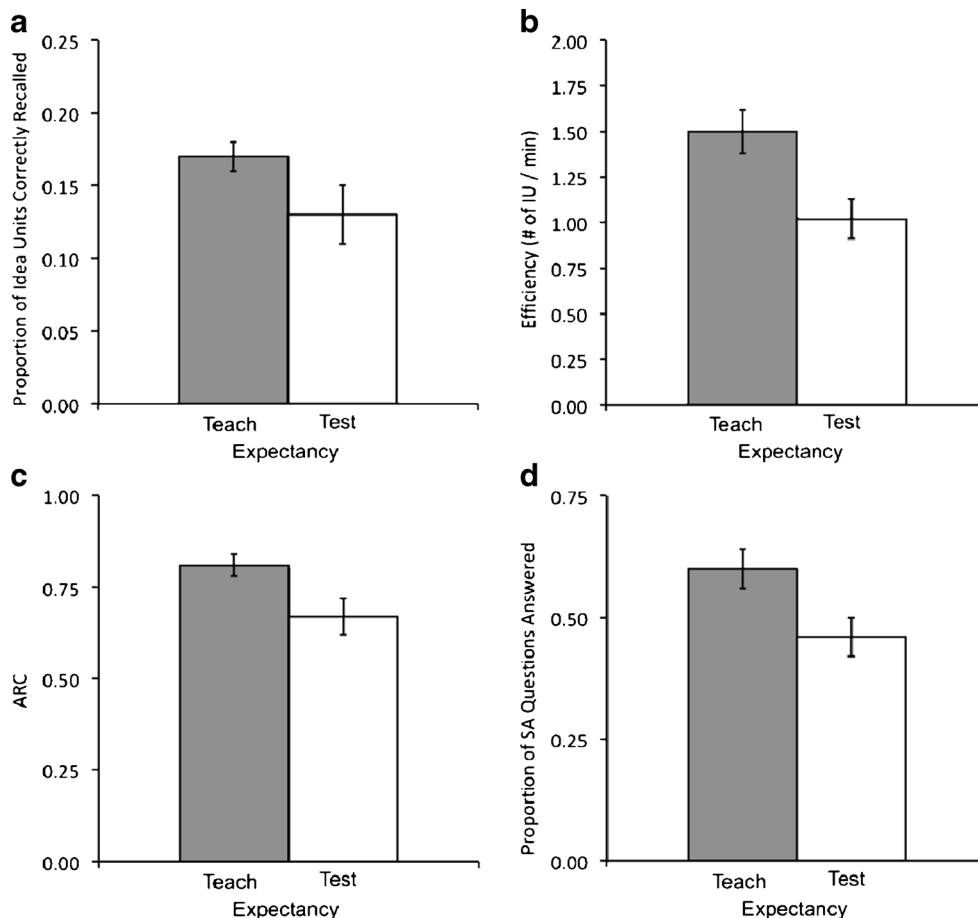

Fig. 1 Performance (means) on the free recall test (**a**, **b**, and **c**) and the short-answer test (**d**) in Experiment 1. Error bars are standard error of the mean (*SEM*). (Performance shown in panels **a**, **b**, and **c** is based on the total set of 125 IUs; performance shown in panel **d** is based on a subset of these IUs (8 important and 10 unimportant)

for the teaching-expectancy group ($M = .81$, $SD = .18$) than for the test-expectancy group ($M = .67$, $SD = .25$), $t(54) = 2.42$, $p = .019$, $d = 0.65$, $CI_{mean\ difference}$ [0.024, 0.255], suggesting that expecting to teach led participants to organize their encoding and/or their recall in a way that reflected the structure of the source passage.

In contrast, however, the two expectancy conditions did not differ in their mean sentence-to-sentence coherence scores. An independent samples t-test confirmed that the LSA scores for the teaching-expectancy group ($M = .24$, $SD = .09$) and the test-expectancy group ($M = .22$, $SD = .09$) did not differ from one another, $t(54) = 0.89$, $p = .380$, $d = 0.24$, $CI_{mean\ difference}$ [−0.025, 0.064], suggesting that the sentence-to-sentence coherence of recall output was similar for the two groups.

Information type Table 1 depicts the relationship between expectancy instructions and memory for IUs of different types of information. These means are based on only the 39 idea units identified as main points (8 IUs), important details (15 IUs), and unimportant details (16 IUs). The results based on these 39 IUs appear consistent with the three analyses already reported, which utilized the full set of 125 IUs, in that expecting to teach enhanced recall of each type of information.

We analyzed these data using a 2 (expectancy instruction: teaching-expectancy vs. test-expectancy) × 3 (information type: main point vs. important detail vs. unimportant detail) mixed-design ANOVA. Consistent with the previous analyses on the full set of 125 IUs, the group expecting to teach recalled more idea units ($M = .22$, $SD = .13$) than the group expecting a test ($M = .17$, $SD = .14$), $F(1, 54) = 3.81$, $MSE = .035$, $p = .056$, $\omega^2_{partial} = .05$. There was also a main effect of information type, $F(1, 54) = 29.2$, $MSE = .012$, $p < .001$, $\omega^2_{partial} = .17$. Post hoc tests revealed that main points ($M = .28$, $SD = .17$) were better recalled than were important details ($M = .21$, $SD = .13$), $t(55) = 3.12$, $p = .003$, $d = 0.47$, $CI_{mean\ difference}$ [0.026, 0.116], which, in turn, were better recalled than unimportant details ($M = .12$, $SD = .13$), $t(55) = 4.31$, $p < .001$, $d = 0.69$, $CI_{mean\ difference}$ [0.047, 0.130]. Post hoc tests also

Table 1 Mean proportions of free recall idea units correctly recalled at specific types of information for the different instruction conditions in Experiment 1

Instructions	Level of information		
	Main (SD)	Important detail (SD)	Unimportant detail (SD)
Teaching expectancy	0.33 (0.14)	0.23 (0.11)	0.14 (0.14)
Test expectancy	0.24 (0.18)	0.19 (0.15)	0.11 (0.12)

Note. Proportion correct recall is based on 8 main point IUs, 15 important detail IUs, and 16 unimportant detail IUs. Standard deviations (SDs) are in parentheses.

revealed a teaching-expectancy advantage over test-expectancy for recall of main points, [$t(54) = 2.16$, $p = .035$, $d = 0.56$, $CI_{mean\ difference}$ [0.007, 0.181], but not for important details, $t(54) = 1.198$, $p = .236$, $d = 0.30$, $CI_{mean\ difference}$ [−0.028, 0.111] or unimportant details, $t(54) = 0.982$, $p = .330$, $d = 0.23$, $CI_{mean\ difference}$ [−0.035, 0.109]. Importantly, however, despite the numerical suggestion of an interaction, the interaction between expectancy instruction and information type was not significant, $F(1, 54) = 1.22$, $MSE = .012$, $p = .299$, $\omega^2_{partial} = .09$, a point to which we return in the Discussion section for Experiment 1.

Short-answer test

Average correct recall performance on the short-answer test for the 8 questions about important details and the 10 questions about unimportant details is shown in Fig. 1d. The apparent superior performance for participants in the teaching-expectancy group was confirmed by the results of a 2 (expectancy instruction: teaching-expectancy vs. test-expectancy) × 2 (information type: important details vs. unimportant details) mixed-design ANOVA, which revealed a significant main effect of expectancy instruction, $F(1, 54) = 5.04$, $MSE = .104$, $p = .03$, $\omega^2_{partial} = .07$, indicating better overall performance for the teaching-expectancy group than for the test-expectancy group. No effect of information type emerged, however, $F(1, 54) = 0.06$, $MSE = .023$, $p = .81$, $\omega^2_{partial} = .00$. Additionally, and consistent with the pattern observed in the free recall data, a significant interaction between expectancy instruction and information type was not obtained, $F(1, 54) = 0.88$, $MSE = .023$, $p = .35$, $\omega^2_{partial} = .00$.

Discussion

In Experiment 1, multiple measures of participants' responses converged to support the claim that expecting to teach promotes learning in ways that expecting a test does not. First, expecting to teach enhanced the amount and efficiency of output in free recall. Additionally, expecting to teach enhanced the match of organization of free recall output to the structure of the source passage (ARC scores), although there was not a teaching-expectancy advantage in the sentence-to-sentence coherence measure (LSA scores). Finally, expecting to teach also produced better performance on short-answer questions. These findings suggest that participants processed information differently, and more effectively, when they expected to teach than when they expected to take a test.

Experiment 1 did not provide strong support for the hypothesis that expecting to teach would specifically enhance recall of important information (as indicated by the lack of a significant interaction between expectancy instructions and information type). As can be seen in Table 1, however, the numerical pattern of our results does appear to indicate that the

recall advantage of the teaching-expectancy group over the test-expectancy group was greater for main points than for the other two types of information. Specifically, the numerical advantage for expecting to teach diminished across levels of importance of information type (9 %, 4 %, and 3 % in the main points, important details, and unimportant details, respectively), and this difference was statistically significant only in the main points, which had the largest effect size of the three comparisons. Thus, in Experiment 2, we further explored the effect of expecting to teach on how different types of information are processed and recalled.

Experiment 2

The primary purpose of Experiment 2 was to test the hypothesis that teaching-expectancy would selectively enhance memory for the central ideas of a to-be-learned passage. We took a different approach to testing this hypothesis in Experiment 2, however, by choosing a passage that we considered to contain mostly important information, with the assumption that—given such a passage—participants expecting to teach would concentrate even more on finding and learning the main points of the passage. That is, because the passage contained a greater number of discrete, important facts than participants would be able to learn in the time provided, participants expecting to teach might be motivated to strategically select main points for conveying to the students they expected to teach. Finally, by using a cued-recall test (fill-in-the-blank), we would be able to assess memory for information that might be available but not necessarily accessible via free recall testing (Tulving & Pearlstone, 1966), which could provide additional insight into the effects of teaching-expectancy.

Method

Participants and design

A total of 44 undergraduate students (32 females and 12 males; average age = 19.8, SD = 2.32) from the University of California, Los Angeles, served as participants for course credit. Experiment 2 employed a 2 × 2 mixed design, with expectancy instructions (teaching-expectancy vs. test-expectancy) as a between-subjects variable and information type (main point vs. detail point) as a within-subjects variable. At the end of the experiment, participants were asked, "Prior to this experiment, what was your knowledge of the topic of the passage you read?" and were given a 10-point scale on which to make their ratings (1 = *no prior knowledge*, 10 = *mastery*). The participants assigned to the two conditions did not differ in their background knowledge of the passage, F = 1.61.

Along with the 44 students who participated in the experiment proper, an additional 44 undergraduate students (23 females and 21 males; average age = 19.23, SD = 1.10) from Washington University in St. Louis were used in a pilot study to categorize test questions as main points and detail points. Details about the administration of this classification task are also presented in the Procedure subsection below.

Materials

To approximate materials used in educational settings, we used a 1,300-word scientific passage titled *Growth and Development of the Brain Reflect the Interaction of Intrinsic and Extrinsic Factors,* which was adopted from a university-level textbook on psychobiology (Rosenzweig, Breedlove, & Leiman, 2002, pp. 194–197).

Procedure

Experiment 2 consisted of several phases: study, distractor, fill-in-the-blank test, and a test-item judgment task.

Study phase The study phase for Experiment 2—including the wording of the prestudy instructions used to create the teaching-expectancy and test-expectancy groups—was identical to that of Experiment 1, with the single exception that participants in the present experiment were given 9 min, rather than 10 min, to study the passage.

Distractor phase After study, participants performed an unrelated distractor task (a separate memory experiment using categorized word lists with no overlap with the present materials) for 30 min before beginning the testing phase.

Test phase Immediately before the test, participants in the test-expectancy condition were told that they would now take a test on the passage. Participants in the teaching-expectancy condition were asked to take a test—inconsistent with their expectation—after being told that the student they were to teach had failed to show up. Both types of participants then took a 27-item fill-in-the-blank test, the answers to which could be found verbatim in the source passage. Each test question was presented on the computer screen, along with a text response box. The test was self-paced such that, after responding, participants could move on to the next question by hitting the Enter key; if 10 s elapsed with no response, the computer automatically advanced to the next question.

Test-item-judgment task: defining information type In the separate pilot study for categorizing test questions as main points and detail points, participants read the passage and were then shown the test questions that were used in the experiment proper. Each question was shown one at a time, but the

information that would be missing for experimental participants was presented in bold, and pilot participants were told that this information would be missing from the items when used as actual fill-in-the-blank test questions in a future experiment. Pilot participants were asked to label each test question as either a *main point* or a *detail point,* and they were told that main points were "statements about the overall idea of the passage" and detail points were "statements about a specific detail." For the analysis of memory performance in the experiment proper, test questions were classified as main points if more than 50 % of the pilot participants rated them as main points and as details if more than 50 % rated them as detail points. Using this "majority rules" criterion, one question that received an even split of votes was dropped from analysis. One other question was dropped because only 2 participants answered it correctly. Among the remaining 25 items, 13 were classified as main-point questions (average rater agreement = 76 %; range of rater agreement = 57 % – 91 %) and 12 as detail-point questions (average rater agreement = 86 %; range of rater agreement = 64 % –100 %).

Results and discussion

Correct performance on the fill-in-the-blank test is shown in Fig. 2, and, as indicated there, orientation instructions and information type appear to interact, with an expectancy to teach selectively enhancing performance on questions related to main points of the passage. To evaluate this observation, a 2 (expectancy instruction: teaching-expectancy vs. test-expectancy) × 2 (information type: main vs. detail) mixed-model ANOVA was conducted, with passage background knowledge included as a covariate. A significant main effect of information type emerged, $F(1, 41) = 8.24$, $MSE = 0.009$, $p = .006$, $\omega^2_{partial} = .02$, indicating that main points ($M = .51$, $SD = .20$) were better recalled than detail points ($M = .33$, $SD = .17$). The main effect of expectancy instruction was not significant, $F(1, 41) = 0.76$, $MSE = 0.054$, $p = .389$, $\omega^2_{partial} = .00$, indicating no difference in the average overall performance of participants expecting to teach ($M = .44$, $SD = .16$) and that of participants expecting a test ($M = .39$, $SD = .17$). Critically, however, a significant interaction between expectancy instruction and information type was revealed, $F(1, 41) = 7.48$, $MSE = 0.009$, $p = .009$, $\omega^2_{partial} = .02$. Follow-up comparisons (again using passage background knowledge as a covariate) showed that the advantage of expecting to teach over expecting a test was marginally significant for the recall of main points, $F(1, 41) = 3.15$, $MSE = 0.034$, $p = .083$, $d = 0.45$, $CI_{mean\ difference}$ [−0.014, 0.211], and that there was no difference between expectancy groups for the recall of details, $F(1, 41) = 0.06$, $MSE = 0.029$, $p = .814$, $d = 0.11$, $CI_{mean\ difference}$ [−0.116, 0.092]. Thus, the pattern of results obtained in Experiment 2 was consistent with our hypothesis that an expectancy to teach would induce learners

to focus on the processing of information related to the central topic of the passage.[1]

General discussion

In two experiments, participants who expected to teach learned more from a passage than did participants who expected to take a test. In addition to overall memory performance, this difference was reflected by two additional measures. First, relative to test-expectancy, teaching-expectancy produced greater organization of output on a later free recall test (Experiment 1). Second, it selectively enhanced memory for main points, whereas recall for detail points did not significantly differ between the two expectancy conditions (Experiment 2; a similar but nonsignificant finding was obtained in Experiment 1). Thus, a relatively simple instructional manipulation led to a meaningful enhancement in how people processed the information to be learned.

In Experiment 1, organization of recall was measured with a modified version of the ARC and using LSA. The ARC, in the present context, measured the degree to which a participant's recall output clustered IUs similarly to how they were organized into paragraphs in the source passage. Thus, one explanation for our finding of higher ARC scores for participants expecting to teach is that teaching-expectancy led participants to pay more attention to the structure of the passage during study, perhaps because they believed that using the passage's structure would help them better teach their prospective student. Alternatively, a teaching-expectancy may have enhanced participants' encoding of the passage's organization in other ways; for example, they may have been trying to think of ways to improve on the passage's organization. The LSA scores suggest that the latter explanation may not be the case, however. The sentence-to-sentence coherence measured by LSA is an index of the coherence of participants' recall output, independent of the structure of the source passage. The average LSA scores were similar for the teaching- and test-expectancy groups in Experiment 1, suggesting that overall coherence of output was not different between the two conditions. This result is not particularly surprising, however, because the LSA scores for both groups were similar to the LSA score for the source passage, indicating that all of the participants in this study recalled information in a fairly

---

[1] In the interest of full disclosure, we note that we also conducted this analysis using a different criterion for classifying main points and detail points: Specifically, questions were labeled as main points or detail points only if 66 % of the pilot participants agreed in their ratings. Using this criterion, the interaction between expectancy instructions and information type was not statistically significant, although the pattern of data was essentially the same as reported using the "majority rules" criterion reported here. The 66 % criterion led to six questions being dropped from analysis, which reduced the power of the ANOVA to detect differences, and so we interpret the observed null effect with caution.
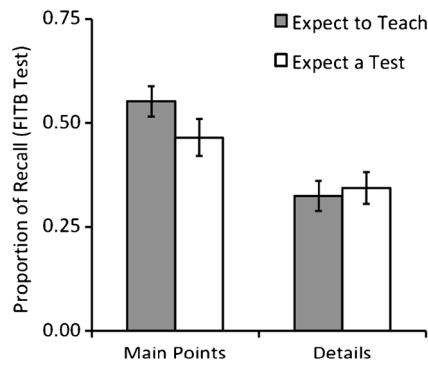
Fig. 2 Means for fill-in-the-blank test in Experiment 2, displayed as a function of instruction condition and information type. Using the "majority rules" criterion of rater agreement, 13 questions were classified as main points, and 12 questions were classified as detail points. Error bars are standard error of the mean (*SEM*)

organized structure. A potentially promising future direction for research on this topic might be to investigate whether teaching-expectancy can lead learners to reorganize an initially poorly structured passage into one with a better structure, given the expectation of having to teach it to another student.

The evidence in the present research regarding the hypothesis that expecting to teach selectively promotes memory for important information was less conclusive. In both experiments, expecting to teach versus expecting a test appeared to convey an advantage for the learning of main points, but not detail points. Expecting to teach versus expecting to take a test produced enhanced cued recall of main points, but not of detail points, in Experiment 2. An analogous pattern for free recall was observed in Experiment 1, but the interaction between level of information and expectancy did not obtain significance. Critically, the effect sizes for the difference between expectancy conditions was consistently greater for main points ($d = 0.56$ and 0.45 in Experiment 1 and Experiment 2, respectively) than for details ($d = 0.30$ and 0.23 for important and unimportant details, respectively, in Experiment 1, and 0.11 for details in Experiment 2). Consistent with recent arguments that effect sizes are an important source of information to consider when interpreting results (Cummings, 2013), we believe that the consistent pattern of results across the two experiments reported herein supports the notion that expecting to teach selectively enhances memory for main points. This issue is not settled, however, and further research is necessary before strong conclusions can be made. Another issue not completely resolved in the present research is whether the observed results reflect enhanced encoding processes, enhanced retrieval processes, or a combination of the two. One venue of future research, therefore, would be to take online measures (e.g., talk-aloud protocols) during encoding and/or retrieval, which might provide additional insights into the loci and nature of the processes underlying the observed benefits of expecting to teach.

Theoretical implications

Why does expecting to teach enhance organization of output and encoding of the main points of a passage? The explanation we currently favor is that participants expecting to teach put themselves into the mindset of a teacher, leading them to adopt certain effective strategies used by teachers when preparing to teach—such as organizing and weighing the importance of different concepts in the to-be-taught material, focusing on main points, and thinking about how information fits together. These teaching preparation techniques are parallel to encoding strategies that are known to be powerful learning or mnemonic processes—namely, relational (organizational) and item-specific processing strategies.

According to basic research on human memory, relational and item-specific processing strategies are typically beneficial to memory (e.g., Hunt & Einstein, 1981). Relational processing—processing the relationships amongst units of information—is proposed to enhance recall by increasing the elements incorporated into memory traces and by allowing for an effective search strategy at the time of retrieval via generative, reconstructive means. The idea that relational processing promotes coherent, effective reconstruction fits with our interpretation of the present finding that the higher degree of output organization displayed by our teaching-expectancy participants reflects their greater relational processing at encoding. Item-specific processing—the encoding of single units of information—is believed to benefit memory by enhancing the distinctiveness of specific memory traces. Participants in the present research recalled a larger proportion of main points than detail points (Experiment 1), and this difference was exaggerated for participants expecting to teach (Experiment 2), indicating that teaching-expectancy enhances item-specific processing for specific subtypes of items—namely, the main points of the passage. Hunt and colleagues have claimed that although tasks often promote one type of processing more than the other, relational and item-specific processing are not mutually exclusive but, rather, often interact in ways that promote overall memory more than does either processing type alone. Consistent with this notion, the general pattern of our results indicates that participants expecting to teach may have benefited from both enhanced relational and item-specific processing.

Expecting to teach might also have triggered explanatory questioning, which enhances learning and memory (see, e.g., Dunlosky, Rawson, Marsh, Nathan, & Willingham, 2013; Roediger & Pyc, 2012). Explanatory questioning techniques such as elaborative interrogation (e.g., Pressley, McDaniel, Turnure, Wood, & Ahmad, 1987) and self-explanation (e.g., Berry, 1983; Chi, de Leeuw, Chiu, & LaVancher, 1994) typically involve prompting learners to generate an explanation for an explicitly stated fact or solution of a problem. Teachers often engage in explanatory questioning when preparing to

teach. Perhaps some of the gains seen in the current studies are due to participants in the teaching-expectancy conditions asking themselves explanatory questions while reading.

Finally, an important question to explore in future research is how prior knowledge might interact with the expectation to teach. Teachers typically start with some knowledge of the topic they intend to teach, so they are likely building upon their current state of understanding when they prepare to teach new material. Likewise, when participants do, in fact, possess relevant knowledge, it might be the case that those expecting to teach would utilize such knowledge more than would those expecting a test. Thus, the benefit of expecting to teach might be especially pronounced when learners have prior knowledge.

Educational implications

Finding effective educational interventions that can be implemented easily and inexpensively is becoming ever more critical, and it has been argued that one way to make progress toward this goal is to apply research findings from cognitive psychology to the classroom (e.g., Roediger & Pyc, 2012). Over the past century, psychologists have evaluated the efficacy of many study methods and, on the basis of a recent assessment of the available evidence (Dunlosky et al., 2013), Roediger and Pyc have suggested that teachers employ the use of distributed practice (see Cepeda, Pashler, Vul, Wixted, & Rohrer, 2006), retrieval enhanced learning (see Roediger & Karpicke, 2006), and explanatory questioning (e.g., Chi et al., 1994). On the basis of the present findings, we would argue that instilling students with an expectancy to teach may lead them to employ some of these techniques and, thus, might be a vehicle for bringing about learning gains in the classroom. Furthermore, given the relative ease of instilling such expectations, this technique could be one of the more easily deployed of such interventions. Whereas few students would know what to do if asked to engage in "explanatory questioning" or "organizational processing," asking them to "prepare to teach" appears to be an instruction they understand quickly and intuitively.

Admittedly, the deceptive version of teaching-expectancy employed in the present report would not work in classroom settings, because students would quickly catch on that they would not ever be required to teach. Perhaps informing the classroom that at least one student among them will be required to teach—but not revealing which student—would prompt all students to prepare as if they will have to teach. This "on the hook" approach improves learning following questions asked by teachers to students in classrooms, when the question is asked before revealing which student will be selected to answer it (Harris, Pashler, & Kang, 2014). We are currently conducting experiments to test whether participants who believe they *might* teach show the same learning gains as

those shown for participants expecting to teach in the present research.

Concluding comments

Teachers learn while they teach (Topping, 1996) and while they prepare to teach (e.g., Benware & Deci, 1984). Expecting to teach appears to encourage effective learning strategies such as seeking out key points and organizing information into a coherent structure. Our results suggest that students also turn to these types of effective learning strategies when they expect to teach. It is noteworthy, then, that when students instead expect to be tested, they underutilize these strategies, although our results clearly indicate that these strategies must be available to them and, furthermore, would better serve their presumed goal of achieving good test performance than do the strategies they instead adopt for this purpose. Students seem to have a toolbox of effective study strategies that, unless prodded to do so, they do not use.

This pattern of findings is in sync with much recent evidence that students do not necessarily employ activities that best foster learning, despite their many years of active involvement in both formal and informal learning activities (for a discussion of this view and the relevant research, see, e.g., Bjork & Bjork, 2011; Bjork, Dunlosky, & Kornell, 2013). In many situations, students appear to need to be guided in how to discover those strategies that are optimal for learning. The present research has demonstrated one way of doing so that promises to be easy to implement in various educational settings. We hope the present findings encourage future researchers to discover other such potentially easy-to-implement ways of leading students to adopt more effective learning strategies.

References

Aiken, E. G., Thomas, G. S., & Shennum, W. A. (1975). Memory for a lecture: Effects of notes, lecture rates, and informational density. *Journal of Educational Psychology, 67,* 430–444.

Ameen, E. C., Guffey, D. M., & Jackson, C. (2002). Evidence of teaching anxiety among accounting educators. *Journal of Education for Business, 78,* 16–22.

Bargh, J. A., & Schul, Y. (1980). On the cognitive benefits of teaching. *Journal of Educational Psychology, 72,* 593–604.

Benware, C. A., & Deci, E. L. (1984). Quality of learning with an active versus passive motivational set. *American Educational Research Journal, 21*(4), 755–765.

Berry, D. C. (1983). Metacognitive experience and transfer of logical reasoning. *Quarterly Journal of Experimental Psychology, 35A,* 39–49.

Bjork, E. L., & Bjork, R. A. (2011). Making things hard on yourself, but in a good way: Creating desirable difficulties to enhance learning. In M. A. Gernsbacher, R. W. Pew, L. M. Hough, & J. R. Pomerantz (Eds.), *Psychology and the real world: Essays illustrating fundamental contributions to society* (pp. 56–64). New York: Worth Publishers.

Bjork, R. A., Dunlosky, J., & Kornell, N. (2013). Self-regulated learning: Beliefs, techniques, and illusions. *Annual Review of Psychology, 64,* 417–444.

Bui, D. C., Myerson, J., & Hale, S. (2013). Note-taking with computers: Exploring alternative strategies for improved recall. *Journal of Educational Psychology, 105,* 299–309.

Cepeda, N. J., Pashler, H., Vul, E., Wixted, J. T., & Rohrer, D. (2006). Distributed practice in verbal recall tasks: A review and quantitative synthesis. *Psychological Bulletin, 132,* 354–380.

Chi, M. T. H., de Leeuw, N., Chiu, M.-H., & LaVancher, C. (1994). Eliciting self-explanations improves understanding. *Cognitive Science, 18,* 439–477.

Cohen, P. A., Kulik, J. A., & Kulik, C. C. (1982). Educational outcomes of tutoring: A meta-analysis of findings. *American Educational Research Journal, 19*(2), 237–248.

Coleman, E. B., Brown, A. L., & Rivkin, I. D. (1997). The effect of instructional explanations on learning from science texts. *The Journal of the Learning Sciences, 6*(4), 347–365.

Cummings, G. (2013). The new statistics: Why and how. *Psychological Science, 27,* 7–29.

Curtiz, M. (Director). (1936). *The charge of the light brigade* [Motion picture]. United States: Warner Bros.

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.

Ehly, S., Keith, T. Z., & Bratton, B. (1987). The benefits of tutoring: An exploration of expectancy and outcomes. *Contemporary Educational Psychology, 12,* 131–134.

Fiorella, L., & Mayer, R. E. (2013). The relative benefits of learning by teaching and teaching expectancy. *Contemporary Educational Psychology, 38,* 281–288.

Gardner, L. E., & Leak, G. K. (1994). Characteristics and correlates of teaching anxiety among college psychology teachers. *Teaching of Psychology, 21,* 28–32.

Gartner, S., Kohler, M., & Riessman. (1971). *Children teach children: Learning by teaching.* New York: Harper and Row.

Gregory, A., Walker, I., McLaughlin, K., & Peets, A. D. (2011). Both preparing to teach and teaching positively impact learning outcomes for peer teachers. *Medical Teacher, 33,* 417–422.

Harris, C., Pashler, H., & Kang, S. (2014). Learning on the hook: Effects of random questioning on retention of factual material. Manuscript in preparation.

Hunt, R. R., & Einstein, G. O. (1981). Relational and item-specific information in memory. *Journal of Verbal Learning and Verbal Behavior, 20,* 497–514.

Landauer, T. K., Foltz, P. W., & Laham, D. (1998). Introduction to latent semantic analysis. *Discourse Processes, 25,* 259–284.

Lundeberg, M. A., & Fox, P. W. (1991). Do laboratory findings on test-expectancy generalize to classroom outcomes? *Review of Educational Research, 61,* 94–106.

Maxwell, S. E., & Delaney, H. D. (2004). *Designing experiments and analyzing data: A model comparison perspective* (2nd ed.). Mahwah, New Jersey: Lawrence Erlbaum Associates.

McKeachie, W. J., Pintrich, P. R., Lin, Y. G., & Smith, D. (1986). *Teaching and learning in the college classroom: A review of the research literature.* Ann Arbor, MI: The Regents of the University of Michigan.

Motley, M. T. (1988). Taking the terror out of talk: Thinking in terms of communication rather than performance helps us calm our biggest fear. *Psychology Today, 22,* 46–49.

Olejnik, S., & Algina, J. (2000). Measures of effect size for comparative studies: Applications, interpretations, and limitations. *Contemporary Educational Psychology, 25,* 241–286.

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,* 291–300.

Rawson, K. A., & Kintsch, W. (2005). Rereading effects depend on time of test. *Journal of Educational Psychology, 97*(1), 70–80.

Renkl, A. (1995). Learning for later teaching: An exploration of meditational links between teaching expectancy and learning results. *Learning and Instruction, 5,* 21–36.

Richmond, V. P., & McCroskey, J. C. (1995). *Communication: Apprehension, avoidance, and effectiveness* (4th ed.). Scottsdale, AZ: Gorsuch Scarisbrick.

Roediger, H. L., III, & Karpicke, J. D. (2006). The power of testing memory: Basic research and implications for educational practice. *Perspectives on Psychological Science, 1,* 181–210.

Roediger, H. L., III, & Pyc, M. A. (2012). Inexpensive techniques to improve education: Applying cognitive psychology to enhance educational practice. *Journal of Applied Research in Memory and Cognition, 1,* 242–248.

Roenker, D. L., Thompson, C. P., & Brown, S. C. (1971). Comparison of measures for the estimation of clustering in free recall. *Psychological Bulletin, 76*(1), 45–48.

Rohrbeck, C. A., Ginsburg-Block, M. D., Fantuzzo, J. W., & Miller, T. R. (2003). Peer-assisted learning interventions with elementary school students: A meta-analytic review. *Journal of Educational Psychology, 95*(2), 240–257.

Roscoe, R. D., & Chi, M. T. H. (2007). Tutor learning: The role of explaining and responding to questions. *Instructional Science, 36,* 321–350.

Rosenzweig, M. R., Breedlove, S. M., & Leiman, A. L. (2002). *Biological psychology: An introduction to behavioral, cognitive, and clinical neuroscience.* Sunderland, MA: Sinauer Associates, Inc.

Ross, S. M., & DiVesta, F. J. (1976). Oral summary as a review strategy for enhancing recall of textual material. *Journal of Educational Psychology, 68,* 689–695.

Ruscio, A. M., Brown, T. A., Chiu, W. T., Sareen, J., Stein, M. B., & Kessler, R. C. (2008). Social fears and social phobia in the United States: Results from the national comorbidity survey replication. *Psychological Medicine, 38*(1), 15–28.

Szpunar, K. K., McDermott, K. B., & Roediger, H. L., III. (2007). Expectation of a final cumulative test enhances long-term retention. *Memory & Cognition, 35*(5), 1007–1013.

Topping, K. J. (1996). The effectiveness of peer tutoring in further and higher education: A typology and review of the literature. *Higher Education, 32*(3), 321–345.

Tulving, E., & Pearlstone, Z. (1966). Availability versus accessibility of information in memory for words. *Journal of Verbal Learning and Verbal Behavior, 5,* 381–391.

Yerkes, R. M., & Dodson, J. D. (1908). The relation of strength of stimulus to rapidity of habit-formation. *Journal of Comparative Neurology and Psychology, 18,* 459–482.

Zajonc, R. B. (1960). The process of cognitive tuning in communication. *Journal of Abnormal and Social Psychology, 61*(2), 159–167.