

THE EFFECTS OF EXPECTING TO TEACH LEARNED INFORMATION
ON STUDENTS' SELF-REGULATED LEARNING

by

Megan Thielman

Submitted in partial fulfillment of the
requirements for Departmental Honors in
the Department of Psychology
Texas Christian University
Fort Worth, Texas

May 7, 2018

THE EFFECTS OF EXPECTING TO TEACH LEARNED INFORMATION
ON STUDENTS' SELF-REGULATED LEARNING

Project Approved:

Supervising Professor: Sarah "Uma" Tauber, Ph.D.

Department of Psychology

Naomi Ekas, Ph.D.

Department of Psychology

Peter Worthing, Ph.D.

Department of History

ABSTRACT

Much of students' learning happens outside of the classroom when they make decisions regarding what and how long to study. These decisions are part of self-regulated learning. Research on self-regulated learning has not yet examined how different evaluation expectations affect students' study decisions. As such, my primary goal was to investigate the impact of expecting to teach versus expecting a test on students' study decisions and actual learning. To do this, I measured changes in self-regulated learning decisions by recording self-paced study latency. Students were randomly assigned to receive instructions that they would take a test on the material or teach the material to someone. Both groups studied a short physics lesson then took a test on the material. Neither group actually taught the material. I found that students who were expected to teach material, compared to students who expected to take a test, studied for twice as long and performed better on the direct recall portion of a final test.

TABLE OF CONTENTS

INTRODUCTION	1
The Impact of Expecting to Teach on Students' Learning	1
Self-Regulated Learning	3
Overview of Current Research	6
METHOD	6
Participants and design	6
Materials	7
Procedure	10
Data Scoring	12
RESULTS	13
DISCUSSION	18
REFERENCES	22

INTRODUCTION

While teaching happens within the classroom, much learning occurs outside of the classroom as students prepare for class evaluations such as exams and presentations. Students are responsible for their own study behavior in order to prepare for such assessments. When students make decisions regarding how they should study, they are engaging in self-regulated learning. One factor that may influence students' study decisions is how they expect to be evaluated. In this study, I will examine how students might change their self-regulated learning behavior when they expect to teach compared to when they expect to take a test. I will monitor changes in behavior by measuring self-paced study time as an index of self-regulated study. Because self-regulated learning is ubiquitous, it is important to understand factors that influence it.

The Impact of Expecting to Teach on Students' Learning

Several researchers have examined how evaluation expectation (e.g., expecting a test versus expecting to teach to others) influences subsequent memory performance. For example, Nestojko, Bui, Kornell, and Bjork (2014) hypothesized that students who studied a passage and expected to teach the material would perform better on a later memory test than would students who studied the same passage and expected to take a test on the material. They also predicted that students expecting to teach would write better-organized responses that contained more main points than would students expecting a test. In their research, participants were told either that they would teach or take a test. Participants then read a passage for 10 min and did a 25 min distractor task before the final assessment. Participants in the teaching expectation group did not teach the lesson; instead, they participated in the same 25 min distractor task as did participants in the test expectation group. Both groups then took a free-response and short answer test. Thus, the groups were identical with the exception of the expectation they received before the study—

teaching or taking a test. The researchers found that participants who expected to teach composed longer, more efficient, better-organized answers than students expecting a test. Moreover, participants who expected to teach out-performed participants who expected to take a test on the short answer component of the evaluation (Nestojko et al., 2014).

Hoogerheide, Deijkers, Loyens, Heijltjes, and van Gog (2016) conducted two experiments to test the impact of students' assessment expectations on their learning. Most relevant, in Experiment 2, participants were placed in one of three groups. One group studied with a test-expectation, engaged in a recall activity, then restudied the material. Another group studied with an explanation-expectation and then explained the material in writing. A third group studied with an explanation-expectation and then explained the material on video. All participants studied the same text for 12 min during learning phase I. Each group was given 6 min for learning phase II—restudying, writing an explanation, or creating a video explanation. All participants then took an immediate posttest and a delayed posttest one week later. Results from the immediate and delayed posttests revealed that explaining in a video benefitted participants' learning more than did restudy, but not significantly benefit learning more than explaining in writing. Hoogerheide et al. (2016) concluded that explaining learned material on video benefitted learning due to participants' enhanced feelings of social presence and the conscious belief that others may see their work.

Fiorella and Mayer (2013) conducted two experiments to disentangle the immediate and delayed learning effects of expecting to teach and actually teaching. Specifically, they directly compared the learning benefits of students who prepared to teach but did not teach and students who prepared to teach and did teach. In their experiments, participants were divided into three groups. The control group studied a lesson on the Doppler Effect then took a test. The teaching

group studied the lesson in preparation to teach, had 5 min to teach a lesson on the Doppler Effect while being videotaped, and then took a test. The preparation group studied the lesson in preparation to teach, did not actually teach the material, and took a test. During the study phase, all participants had 10 min to study the same short lesson on the Doppler Effect. In Experiment 1, participants took a comprehension test immediately after studying the material. In Experiment 2, participants took a comprehension test one week after studying the material. The results from Experiment 1 revealed that the teaching group and preparation group performed better on the comprehension test than did the control group. Teaching did not significantly improve participants' scores relative to preparing to teach (Fiorella & Mayer, 2013).

For Experiment 2, the researchers followed the same procedure used for Experiment 1 but included a delayed comprehension test one week later (Fiorella & Mayer, 2013). The researchers discovered that the teaching group scored significantly higher on the comprehension test than did the preparation group. The preparation group did not significantly outperform the control group. Through their two experiments, Fiorella and Mayer (2013) showed that while actually teaching has some longer-lasting benefits, preparing to teach (without actually teaching) has a positive effect on learning on an immediate test. Thus, expecting to teach to others has tangible learning benefits (Fiorella & Mayer, 2013; Hoogerheide et al., 2016; Nestojko et al., 2014).

Self-Regulated Learning

Despite the growing body of knowledge on evaluation expectations, no prior research has examined the impact of different expectations on students' self-regulated study decisions. Examining self-regulated learning is important because it relates to the choices students make regarding what, how, and how long to study. Two basic elements self-regulated learning are

monitoring and control (Bjork, Dunlosky, & Kornell, 2013). Students make judgments to monitor how much they think they know, which can influence their actions (Bjork et al., 2013). In simpler terms, the decisions students make when studying can provide information about how they think about their own learning.

One framework of self-regulated learning is referred to as the discrepancy reduction model (Bjork et al., 2013). The discrepancy reduction model suggests that students have a goal that they are trying to reach when they are learning. Learning goals provide students with a final state of knowledge they desire to reach, such as mastery of the learning material (Dunlosky, Ariel, & Thiede, 2011). Goals may be implicit, such as curiosity about a subject, or explicit, such as a researcher telling a student to study material in preparation for teaching the material later on. A student's behavior is aimed to reduce the gap, or discrepancy, between his or her current level of knowledge and intended final state of knowledge (Bjork et al., 2013).

A more comprehensive model for understanding self-regulated learning was established with the agenda-based regulation (ABR) model (Dunlosky et al., 2011). This model focuses on learners' use of agendas or plans in determining how to study. Similar to the discrepancy reduction model, ABR considers self-regulated studying to be a goal-oriented process (Dunlosky et al., 2011). Students create and act on agendas that allow them to study in a manner that will help them achieve goals efficiently. The use of agendas is dependent upon the interplay between monitoring and control. Students monitor their learning to inform themselves about what information should get more attention, allowing them to control their behavior and selectively give attention to that information (Dunlosky et al., 2011). Thus, study time is determined by monitoring while study decisions are executed as a function of students' control.

Both top-down and bottom-up processes affect students' allocation of study time (Dunlosky et al., 2011). For example, a student who is not goal-oriented or has little motivation may construct an agenda with the goal of obtaining a minimally accepted score on a test (i.e., more bottom-up driven). In comparison, a student who is highly goal-oriented and motivated may construct an agenda in order to maximize the chance that he or she will learn and retrain the information in the long term (i.e., more top-down driven). Individual characteristics such as motivation, interest, and self-efficacy alter students' agendas, which affects their behavior (Dunlosky et al., 2011). One person can have many agendas, which are also influenced by contextual environmental factors. Researchers have examined some environmental cues or task constraints; for instance, time limitations, the type of material being studied, and performance goals (Dunlosky et al., 2011).

Evaluation expectation is another example of an environmental cue because it changes the nature of a given task. It is reasonable to expect that telling a student that he or she will have to teach information, rather than take a test, will cause him or her to set a different agenda. The task of teaching involves a different type of evaluation and the additional pressure of being socially affected. If altering evaluation format results in students changing their agendas, it is reasonable to assume that their behavior will change as a result because agendas affect learning behavior (Dunlosky et al., 2011). Prior research on ABR and self-regulated learning has used tests as the main format of evaluation (Dunlosky et al., 2011), so little is known about self-regulation with different types of assessment. This model could extend to help explain how students prepare for evaluations other than tests. My study aims to explore the degree to which ABR generalizes to other situations, specifically preparing to teach. To do this, in the current study, I examine the effects of expecting to teach on students' self-regulated learning behaviors.

Overview of Current Research

In this experiment, participants were given unlimited time to study a short multimedia lesson on the Doppler Effect and decided for themselves when they were finished studying (materials provided by Fiorella & Mayer, 2013). We chose self-paced study as a primary measure of self-regulated learning because it is commonly used and is typically related to measures of monitoring accuracy (Dunlosky et al., 2011). Half of the participants served in the test-expectancy group and were told that they would take a test after studying. Half of the participants served in the teaching-expectancy group and were told that they would be videotaped teaching a lesson after studying. Despite these instructions, participants in the teaching-expectancy group did not actually teach any material. All participants were tested directly after studying then took a brief post-questionnaire. Based on the ABR model (Dunlosky et al., 2011), I hypothesized that participants in the two groups would differ in self-paced study time, with students in the teaching-expectancy group studying longer than would students in the test-expectancy group. Additionally, I expected the teaching-expectancy group to outperform the test-expectancy group on the comprehension test (cf. Fiorella & Mayer, 2013; Nestojko et al., 2014). Self-regulated study time and test performance are my two main measures of interest. Other measures including recall latency and questions in the post-questionnaire, such as perceived effort invested, are secondary interests included to gather a more complete understanding of participants' thoughts.

METHOD

Participants and design

Eighty undergraduate students from Texas Christian University were recruited for psychology course extra credit. Forty participants were randomly assigned to each group. Data

from two participants are excluded (one from each group) from the final study because the participants neglected to answer all questions on the post-test due to advancing past questions before answering them. The final obtained sample size was 78 participants, with 39 serving in each group. Evaluation expectation (test-expectancy or teaching-expectancy) was manipulated between-participants. The groups were made up of participants identifying from similar ethnic groups. Specifically, the teaching-expectancy group was composed of 30 Caucasian students, 6 Hispanic students, and 2 Asian/Pacific Islander students. The test-expectancy group consisted of 25 Caucasian students, 6 Hispanic students, 2 Asian/Pacific Islander students, and 5 African American students. The ethnic distribution was similar although significantly different between groups, $\chi^2(8) = 119.91, p = .00$. The difference was driven by the fact that due to random assignment, the teaching-expectancy group had 0 African American participants, while the test group had 5 African American participants. The gender distribution was similar although there was a significant difference between groups, with 8 men randomly assigned to the teaching-expectancy group and 12 men randomly assigned to the test-expectancy group, $\chi^2(4) = 113.32, p = .00$. Both groups were predominantly women. Finally, the groups did not differ in terms of age, $t < 1; p = 0.70$. The average age of participants was 20.63 years ($SD = 2.72$) in the teaching-expectancy group and 20.42 years ($SD = 2.05$) in the test-expectancy group. One participant in each group did not report demographic data.

Materials

The computer materials consisted of a consent form, a demographic questionnaire, a prior knowledge assessment, instructions, a multi-media lesson on the Doppler Effect, a test, and a post-test questionnaire (experimental materials provided by Fiorella & Mayer, 2013). The consent form was presented electronically and indicated that participants should type their full

name to sign the form if they chose to give consent. The form used was a standard consent. The demographic questionnaire included questions about age, gender, and ethnicity. First, to assess perceived prior knowledge, participants rated their knowledge of the Doppler Effect on a scale from 1 (very low) to 5 (very high). They were then given a free-response question asking them to explain the Doppler Effect. This question was worth 13 possible points and was used to evaluate whether or not participants had existing prior knowledge of the Doppler Effect. Then, following the Fiorella and Mayer (2013) protocol, to measure self-efficacy, participants rated how well they thought they would perform on a test after studying a lesson on the Doppler Effect on a scale of 1 (very poorly) to 5 (very well). To gauge general knowledge of physics, participants were then asked to select “yes” or “no” for each of the following items: “I know what Hz means,” “I have used an oscilloscope,” “I know what relative motion is,” “I know what the red shift is,” and “I know what a sine curve is.” Given our interest in self-regulated learning, these questions were self-paced so participants could take as much time as they desired to answer each.

For the primary task, the instructions informed participants about how they would be assessed after studying a lesson on the Doppler Effect. The test-expectancy group was told that they would be asked to take a test over the material they studied:

You will now study a short lesson on how the Doppler Effect works. You will then take a test about what you learned. So, you should study as if you were studying for an exam in class. You have as long as you need to study the lesson. You can take your time, revisit the material, and make sure that you know the information well. If you have any questions, please ask the experimenter.

The teaching-expectancy group was told that they would be asked to teach the material they studied:

You will now study a short lesson on how the Doppler Effect works. You will then teach the material that you learned and your lesson will be videotaped. Specifically, you will make the best video you can to clearly explain the Doppler Effect to someone who has no prior knowledge of it so that they can learn it well. To familiarize you with the

videotaping procedure, the researcher will give you a media release form and will show you where your presentation will be filmed.

You have as long as you need to study the lesson. You can take your time, revisit the material, and make sure that you know the information well. If you have any questions, please ask the experimenter.

Participants in the teaching-expectancy group each signed a paper copy of a media release form agreeing to be recorded during their participation in the research. By signing the form, participants agreed to have their media records studied for use in this project, played to other subjects in this research study, and used for scientific publications or scholarly conferences. They were then shown a Cannon Vixia HF R42 camcorder that was set up on a tripod that would be used to video tape their lesson.

Participants in both groups read a roughly 600-word lesson on the Doppler Effect accompanied by five images (Fiorella & Mayer, 2013). The entire lesson was presented on one computer screen. The images were on one side of the screen and the text was on the other side with a scroll bar. The lesson included a basic explanation of the Doppler Effect, explained with a concrete example (i.e., a fire truck with sirens approaching a listener). Next, the lesson contained detailed information about sound waves, along with accompanying definitions and illustrations of frequency and wavelength. Lastly, the lesson provided another concrete example (i.e., a bug creating ripples by moving on the surface of water) and reiterated the first example. Participants were given unlimited time to study the lesson.

The test consisted of six free-response questions (Fiorella & Mayer, 2013). The first question, "Explain how the Doppler Effect works," was designed to test participants' knowledge of the lesson content. The other five questions required participants to think critically as the answers to these questions were not directly present in the text. The following questions were included: "How could you increase the intensity of the Doppler Effect?" "Would the Doppler

Effect occur if the source was stationary and the observer was moving? Why or why not?” “What would happen to the Doppler Effect if the observer was moving the same speed and in the same direction as the source? Explain your answer.” “How would an observer experience sound if the speed of the source were traveling faster than the speed of sound?” and “What would happen to the Doppler Effect if the source and observer were both moving towards each other at on a parallel path at a constant speed? Explain your answer.” Participants were given unlimited time to answer the free-response questions.

Following the Fiorella & Mayer (2013) protocol, the post-test questionnaire was included to gather additional information even though these questions were not our primary interest. It consisted of the following five statements: “I felt that the subject matter was difficult,” “I found the lesson about the Doppler Effect to be useful to me,” “I felt stressed while I was learning about the Doppler Effect,” “I feel like I have a good understanding of how the Doppler Effect works,” and “Please rate the amount of mental effort you put into learning about the Doppler Effect.” The post-test questionnaire included an attentional measure, “I will select seven to indicate I am answering carefully.” Participants answered these questions based on how much they agreed with each of them on a scale from 1 (strongly disagree) to 7 (strongly agree). Each of these questions was self-paced.

Procedure

Participants were randomly assigned to either the test-expectancy or teaching-expectancy group. Participants were individually tested, and the experiment was conducted in any one of three small rooms. Before the experiment, participants gave their informed consent by signing an electronic form and completed a brief demographic questionnaire. Participants were free to abstain from reporting any demographic information. To begin the experiment, experimenters

first read the instructions on the computer to each participant before they began studying to ensure that the participant knew what type of assessment to expect. Second, for the teaching-expectancy group, experimenters had participants sign a media release form and showed them a video camera set up in a separate room before studying took place. The experimenters told participants “We will film your presentation here and it will be recorded by this video camera. Do you have any questions?” If a participant asked about the duration of the video, the experimenter answered, “Your video can be as long as you need it to be. There is no set duration for your presentation.” If a participant asked about visual aids, the experimenter told them “You will not have access to any visual aids.” If the participant asked any other question, the experimenter answered vaguely or told the participant that s/he would look into the inquiry while the participant studied. Third, participants in both groups were told that they could study the material for as long as they desired. Lastly, participants were told to select the “finished” button on the screen when they finished studying. When the participants decided they were done studying and selected “finish,” the computer program prompted them to locate an experimenter. Participants then opened their individual study room door to do so.

Before beginning the test, participants in the test-expectancy group were told, “In order for you to complete the study and earn credit, you are going to take a brief test on the material you studied.” Participants in the teaching-expectancy group were told that experimenters were having problems with the video camera, so they would not be able to create their video. They were then told, “In order for you to complete the study and earn credit, you are going to take a brief test on the material you studied instead.” At that point, the experimenter clicked a hidden button on the computer to advance participants in both groups to the test. Participants had unlimited time to complete the free-response questions. After the test was completed,

participants completed the post-test questionnaire. Finally, experimenters debriefed participants, thanked them for participating, and granted them credit.

Data Scoring

The final comprehension test was worth 25 points. Two raters blind to group assignment scored participants' responses. The raters had high inter-rater reliability, as evidenced by strong positive correlations (See table 1). I used Pearson's r correlation to assess inter-rater reliability because doing so is common with continuous measures as used in this study (e.g., Tauber et al., in press). Agreement between the two raters was high, indicated by relatively high correlations ($r = .61$ to $r = .96$). Kappa is also included for interested readers. The Kappa values are to be interpreted with caution, as the measures used in the current study were not categorical.

Table 1 Interrater statistics

	Correlation	Significance
Pearson Correlation		
Pre-test Question	.96	.00
Question 1	.88	.00
Question 2	.66	.00
Question 3	.93	.00
Question 4	.61	.00
Question 5	.94	.00
Question 6	.66	.00
Kappa		
Pre-test Question	.70	.00
Question 1	.34	.00
Question 2	.52	.00
Question 3	.46	.00
Question 4	.25	.00
Question 5	.82	.00
Question 6	.28	.00

To calculate a participant's final score, any discrepancies between scores from the two raters were averaged. Participants could earn a total of thirteen points on the first question by

incorporating certain key components of the lesson into their answer. For example, a participant would earn one point for stating that wavelength and frequency are the two characteristics of sound waves. The remaining five questions were worth two to three points each, with one point awarded for each correct element the participant included. For example, the fourth comprehension question asks, “How would an observer experience sound if the speed of the source were traveling faster than the speed of sound?” and participants earned one point for each of the following three parts that they included: 1. Sound would be delayed, 2. A sonic boom would occur, and 3. Sound waves would be heavily compressed.

RESULTS

Prior to experiencing any experimental material, I assessed participants’ prior knowledge of the Doppler Effect in multiple ways. Most participants reported not taking college courses in physics, which did not differ between the groups, $\chi^2(1) = .14, p = .71$. A total of 5 participants or 12.8% of the teaching-expectancy group and 6 participants or 15.8% of the test-expectancy group had taken college physics. Participants had little expertise on the Doppler Effect based on their self-ratings on a scale from 1 to 5 (Teaching-expectancy: $M = 1.62; SD = .81$; Test-expectancy: $M = 1.37; SD = .63$), and scores did not differ between groups, $p = .14$. Participants from both groups gave predictions for test performance on a scale from 1 to 5 (Teaching-expectancy: $M = 2.95; SD = 1.05$; Test-expectancy: $M = 2.92; SD = 1.08$), and predictions did not differ between the groups, $p = .91$. Finally, participants were asked to explain the Doppler Effect on a free response question worth 13 points. Participants had minimal knowledge of the Doppler Effect, and the two groups did not differ in prior knowledge (Teaching-expectancy: $M = .02; SD = .05$; Test-expectancy: $M = .01; SD = .05; t < 1; p = .37$). I chose to leave all

participants in the sample because the scores on the free response question indicated that no participants were able to explain the Doppler Effect.

My primary interest was to evaluate self-paced study latency. An independent samples t-test indicated that study latency differed significantly between the groups, $t(75) = 4.62, p < .001, d = 1.05$. As evident in Figure 1, on average participants in the teaching-expectancy group studied for twice as long as did participants in the test-expectancy group.

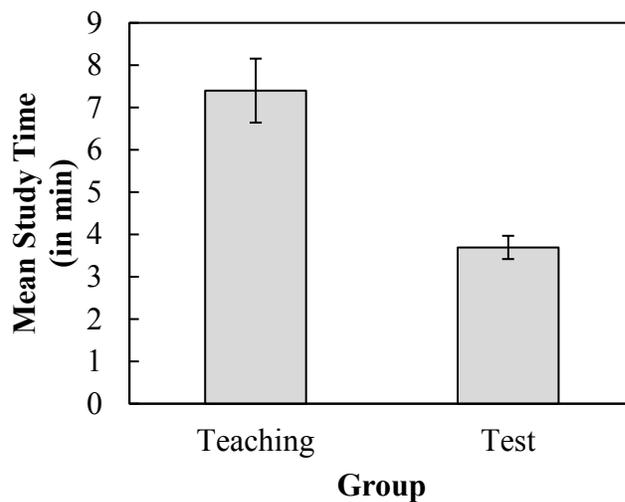


Fig. 1 Mean study latency. Error bars represent standard error.

My other main goal was to replicate prior research in which the teaching-expectancy group outperformed the test-expectancy group on a comprehension test (cf. Fiorella & Mayer, 2013; Nestojko et al., 2014). As evident in Figure 2, the teaching-expectancy group performed significantly better than the test-expectancy group on question 1 (which required direct recall of the Doppler effect), according to an independent samples t-test, $t(75) = 2.44, p < .05, d = .56$.

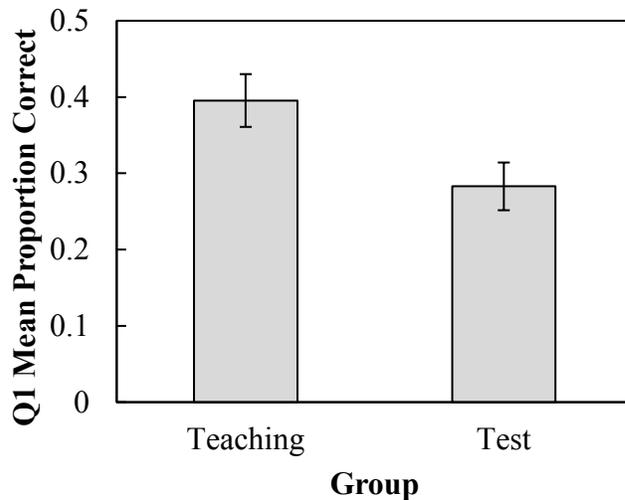


Fig. 2 Mean proportion correct on question 1. Error bars represent standard error.

In contrast with the proportion correct on question 1, the two groups performed similarly on questions 2 through 6. These were inference-based questions, which required critical thinking. Performance on these questions did not differ significantly between the groups (See top portion of Table 2).

Table 2 Proportion correct and recall latency on inference based questions

	Teaching Expectancy	Test Expectancy	<i>t</i>	<i>p</i>
Proportion Correct on Inference-Based Questions				
Question 2	21.79 (22.71)	22.70 (22.11)	.18	.86
Question 3	45.73 (34.03)	51.54 (31.41)	.78	.44
Question 4	49.68 (30.28)	55.26 (26.26)	.86	.39
Question 5	15.00 (17.45)	11.18 (29.97)	.88	.62
Question 6	20.19 (26.84)	17.43 (21.26)	.50	.37
Recall Latency for Inference-Based Questions				
Question 2	62.55 (74.78)	57.55 (37.57)	.37	.71
Question 3	72.51 (47.92)	81.45 (43.10)	.86	.39
Question 4	62.30 (41.03)	74.10 (45.46)	1.20	.24

Question 5	61.07 (32.72)	60.61 (32.72)	.06	.95
Question 6	70.50 (36.15)	80.33 (38.06)	1.16	.25

Note. Descriptive and inferential statistics provided for the teaching expectancy and test expectancy groups for the proportion correct on and recall latency for the inference-based questions (questions 2-6). Standard deviations are provided in parentheses.

For exploratory purposes I also measured recall latency, which is the amount of time participants took to respond to each test item. As evident in Figure 3, the teaching-expectancy group took significantly longer to respond to question 1 than did the test-expectancy group, based on an independent samples t-test, $t(75) = 3.82, p < .001, d = 0.87$. Recall latency did not differ significantly between groups on questions 2 through 6, as seen in Table 2.

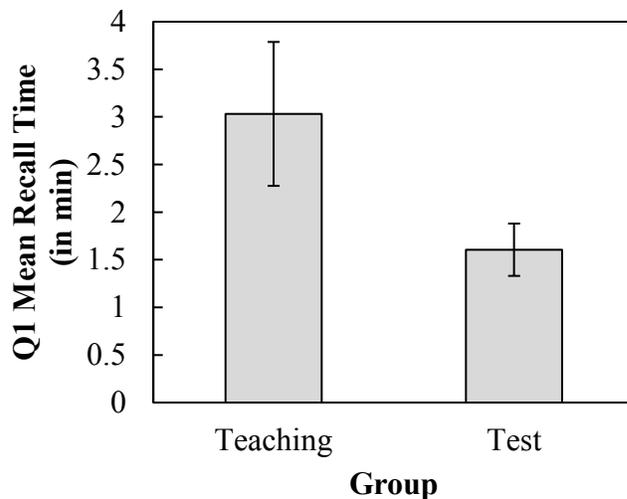


Fig. 3 Mean recall latency on question 1. Error bars represent standard error.

After the final test, participants completed a post-test questionnaire. Specifically, they rated the subject matter on difficulty and usefulness, how stressed they felt while learning, the degree to which they understand the information, and the amount of mental effort they put into learning. The post-test questionnaire also included an attention check. There was no difference between groups on the attention check, which asked participants to select “7” to indicate they

were paying attention (See Table 3). Responses from both groups indicated that they were paying attention. The groups did not differ in their ratings of lesson difficulty, lesson usefulness, or stress while learning (See Table 3). The groups differed significantly on their level of agreement with the statement “I feel like I have a good understanding of how the Doppler Effect works,” with participants in the teaching-expectancy group agreeing more than did participants in the test-expectancy group (See Table 3, $t(75) = 2.56, p < .001$). Further, participants who expected to teach the material self-reported a better understanding of the Doppler Effect relative to participants who expected to take a test. This result is interesting because it reflects participants’ awareness about how well they knew the material. Finally, the groups also differed on their rating of how much mental effort they put into learning about the Doppler Effect, with participants in the teaching-expectancy group rating their mental effort as significantly higher than participants in the test-expectancy group did (See Table 3, $t(75) = 3.42$). Participants who expected to teach self-reported putting more mental effort into learning than participants who expected to take a test did.

Table 3 Self-rated post-test measures

	Teaching Expectancy	Test Expectancy	<i>p</i>
Post-Test Questionnaire Results			
Difficulty	3.15 (1.29)	3.61 (1.41)	.15
Usefulness	4.5 (1.70)	3.98 (1.76)	.16
Stress	2.77 (1.51)	2.52 (1.20)	.44
Understanding	5.31 (1.49)	4.42 (1.55)	.01
Attention check	7.00 (.00)	6.95 (.32)	.31
Mental effort	5.62 (1.23)	4.58 (1.43)	<.001

DISCUSSION

In the current experiment, I sought to measure self-regulated study behavior, particularly self-paced study latency. Participants who expected to teach studied on average for twice as long as did participants who expected to take a test (See Figure 1). Differences in study latency reflect differences in self-regulated study behavior. Telling participants to prepare for a certain format of evaluation—teaching or a test—created different evaluation expectations between the groups. It is reasonable to assume that different evaluation expectations caused participants to set different agendas. The test-expectancy group may set an agenda with the goal of knowing the information well enough to recall the main points on a written short-answer test. The teaching-expectancy group might have constructed an agenda in order to know the information well enough to verbally explain it to others. Importantly, based on the ABR model, agendas affect learning behavior (Dunlosky et al., 2011). For example, if participants in the teaching-expectancy group created a higher criterion agenda than those in the test-expectancy group, that agenda should result in studying longer in order to achieve that goal. Even so, one experiment is not sufficient to understand how students alter their study behavior based on the type of evaluation they expect. If the differential agenda hypothesis holds, then it is possible that expecting to teach produces different levels of motivation than does expecting at test. To explore this possibility, in follow-up work we are including post-test questionnaires asking participants about the motivation behind their study behavior, such as why they chose to stop studying when they did. In additional follow-up experiments, we are currently investigating whether or not self-regulated study behavior changes based on the type of test participants expect (short answer, short essay, or multiple choice).

Participants who expected to teach outperformed those who expected to take a test on question 1, which directly tested the information present in the material that participants studied (See Figure 2). This finding replicates prior research (e.g., Fiorella & Mayer, 2013; Nestojko et al., 2014). Participants' performance has implications for what class activities may be beneficial. For example, students may benefit from preparing to teach information or from actually creating presentations about class concepts in comparison with expecting to take a test. Additionally, I recorded recall latency as an exploratory measure in order to be thorough. Not only did participants in the teaching-expectancy group outperform participants in the test-expectancy group on question 1, they also took significantly longer to answer question 1 than participants in the test-expectancy group took. This information would be useful for teachers if they ask students to expect to teach information. Those students may take longer on subsequent tests or quizzes over that material relative to students who were always expecting to take a test or quiz.

Participants who expected to teach and those who expected to take a test performed similarly on questions 2 through 6, which were inference-based questions (See Table 2). This finding replicates prior research (e.g., Fiorella & Mayer, 2013). Even so, more research is needed to discern why participants' results for the inference questions did not differ between groups and how performance could be improved for inference questions rather than solely for direct recall questions. Finally, there was no significant difference in recall latency between participants in the teaching-expectancy group and test-expectancy group on questions 2 through 6 (See Table 2). If future studies showed participants in the teaching-expectancy group improve on inference questions relative to the test-expectancy group, I would expect their reaction time increase as well because they may be thinking more deeply about the material. Research showing whether or not reaction times are related to accuracy for the two groups would be interesting.

Participants from both groups reported similar levels of stress while studying on the post-test questionnaire (See Table 3). Prior to conducting the study, I expected that there might be a difference in stress ratings. I anticipated that the teaching-expectancy group would report feeling more stressed than the test-expectancy group, potentially due to the added pressure of being videotaped. The teaching-expectancy group may have not reported additional stress because the post-test questionnaire was given after the experiment was over, when participants had already been told that they would not be making a video. While it is unlikely that the difference in learning between groups is due simply to stress, the measure used in this study is a self-report that may not accurately portray physiological stress. In order to better assess stress, future studies could use more careful measurements, such as skin conductance, and should consider assessing stress before participants in the teaching-expectancy group are told that they will not have to teach the material.

Taken together, the results of this study reveal that asking participants to teach information rather than take a test altered how they approached the learning task. Specifically, participants in the teaching-expectancy group studied longer, performed better on the direct recall question, and spent longer answering the direct recall question than participants in the test-expectancy group did. Changes participants in the teaching-expectancy group made in their self-regulated study behavior directly benefitted learning, as evidenced by their superior performance on test question 1 (the direct recall question). Based on the findings of this experiment, I would recommend that professors have students teach information more often as a tool to help students learn. Using this method of learning along with taking tests could enhance students' learning. While further research is needed to understand why participants who expect to teach outperform participants who are expecting a test on direct recall questions, the current experiment was the

first to show that participants expecting to teach information change their self-regulated study behavior by studying material for longer than participants expecting to take a test.

REFERENCES

- Bjork, R. A., Dunlosky, J., & Kornell, N. (2013). Self-regulated learning: Beliefs, techniques, and illusions. *Annual Review of Psychology*, *64*, 417-444. doi: 10.1146/annurev-psych-113011-143823
- Dunlosky, J., Ariel, R., & Thiede, K. W. (2011). Agenda-based regulation of study-time allocation. In P. A. Higham & J. P. Leboe (Eds.), *Constructions of Remembering and Metacognition: Essays in Honour of Bruce Whittlesea* (pp. 182-198). London: Palgrave Macmillan.
- Fiorella, L., & Mayer, R. E. (2013). The relative benefits of learning by teaching and teaching expectancy. *Contemporary Educational Psychology*, *38*(4), 281-288. doi: 10.1016/j.cedpsych.2013.06.001
- Hoogerheide, V., Deijkers, L., Loyens, S. M., Heijltjes, A., & Gog, T. V. (2016). Gaining from explaining: Learning improves from explaining to fictitious others on video, not from writing to them. *Contemporary Educational Psychology*, *44-45*, 95-106. doi: 10.1016/j.cedpsych.2016.02.005
- Nestojko, J. F., Bui, D. C., Kornell, N., & Bjork, E. L. (2014). Expecting to teach enhances learning and organization of knowledge in free recall of text passages. *Memory & Cognition*, *42*(7), 1038-1048. doi: 10.3758/s13421-014-0416-z
- Tauber, S. K., Witherby, A. E., Dunlosky, J., Rawson, K. A., Putnam, A. L., and Roediger, H. L. (in press). Does covert retrieval benefit learning of key-term definitions? *Journal of Applied Research in Memory and Cognition*, <http://dx.doi.org/10.1016/j.jarmac.2016.10.004>