Contemporary Educational Psychology 38 (2013) 281–288

Contents lists available at SciVerse ScienceDirect

Contemporary Educational Psychology

journal homepage: www.elsevier.com/locate/cedpsych

The relative benefits of learning by teaching and teaching expectancy

Logan Fiorella *, Richard E. Mayer

University of California, Santa Barbara, United States

ARTICLE INFO

Article history:
Available online 7 June 2013

Keywords:
Learning by teaching
Multimedia learning
Explanation
Generative learning
Teaching expectancy

ABSTRACT

The purpose of this study was to explore the hypothesis that learning is enhanced through the act of teaching others. Specifically, two experiments aimed to disentangle the relative effects of teaching expectancy (i.e., preparing to teach) and actually teaching (i.e., explaining to others for instructional purposes) on learning. Some participants studied a lesson on the Doppler Effect without the expectation of later teaching the material and then took a comprehension test on the material (control group). Other students studied the same lesson with instructions that they would later teach the material; of those expecting to teach, some participants actually taught the material by presenting a brief video-recorded lecture before being tested (teaching group), whereas others only prepared to teach before being tested (preparation group). Results of Experiment 1 indicated that both the preparation group and teaching group significantly outperformed the control group on an immediate comprehension test (Teaching vs. Control: $d = 0.82$; Preparation vs. Control: $d = 0.59$). However, when the same test was given following a one-week delay (Experiment 2), only the teaching group significantly outperformed the control group (Teaching vs. Control: $d = 0.79$; Preparation vs. Control: $d = 0.24$). Overall, these findings suggest that when students actually teach the content of a lesson, they develop a deeper and more persistent understanding of the material than from solely preparing to teach.

1. Introduction

Imagine that a beginning graduate teaching assistant is asked to present a lecture on the Doppler Effect to a large introductory physics class. The assistant knows the material well but does not yet feel confident in his ability to *teach* it to others. He decides to spend his preparation time devoting more effort towards selecting the most relevant information from the textbook and organizing it into a meaningful representation that he can later retrieve while teaching. During his lecture, he explains to the class how the Doppler Effect works by elaborating on the representation he constructed during preparation and by drawing on his prior knowledge to provide concrete examples. Finally, by the end of the class the teaching assistant has developed a much deeper understanding of how the Doppler Effect works. This is an example of what can be called *learning by teaching*–learning new material more deeply through teaching it to others.

It is sometimes said that teaching others is a powerful way to learn, but how does learning by teaching actually work? For instance, in the above example, at which point did the teaching assistant acquire a strong understanding of the material: after preparing to teach or after actually presenting his lecture? Alternatively, it may be that his understanding developed across both stages, receiving additive benefits from preparing for and actually teaching. There is a considerable amount of research indicating that teaching others can be an effective way to learn (Allen & Feldman, 1973; Bargh & Schul, 1980; Chi, Siler, Jeong, Yamauchi, & Hausmann, 2001; Cohen, Kulik, & Kulik, 1982; Coleman, Brown, & Rivkin, 1997; Ehly, Keith, & Bratton, 1987; Ginsburg-Block & Fantuzzo, 1997; Graesser, Person, & Magliano, 1995; Griffin & Griffin, 1998; Morgan & Toy, 1970; Palinscar & Brown, 1984; Robinson, Schoefield, & Steers-Wentzell, 2005; Rohrbeck, Ginsburg-Block, Fantuzzo, & Miller, 2003; Roscoe & Chi, 2007; Webb, 1982); however, the reasons for this effect are much less clear (Galbraith & Winterbottom, 2011; Ploetznerl, Dillenbourg, Praier, & Traum, 1999; Rohrbeck et al., 2003; Roscoe & Chi, 2007). In particular, this ambiguity is likely due to the diversity of learning by teaching approaches available (e.g., cross-age tutoring, reciprocal tutoring, teachable agents), the many teaching-related activities potentially responsible for learning (e.g., preparing to teach, explaining to others, providing feedback), as well as other aspects of teacher-student interactions that are not necessarily unique to teaching (e.g., answering questions). Consequently, research on learning by teaching—particularly within peer tutoring contexts—has produced quite mixed results (Rohrbeck et al., 2003) and overall learning effects that are relatively underwhelming (Cohen et al., 1982; Roscoe & Chi, 2007). Given learning by teaching is at the core of

* Corresponding author. Address: Department of Psychological and Brain Sciences, University of California, Santa Barbara, CA 93106, United States.

E-mail addresses: fiorella@psych.ucsb.edu (L. Fiorella), mayer@psych.ucsb.edu (R.E. Mayer).

0361-476X/$ - see front matter
http://dx.doi.org/10.1016/j.cedpsych.2013.06.001

many popular educational practices, including small group discussions (e.g., Webb, 1982), reciprocal teaching (e.g., Palincsar & Brown, 1984), and cooperative learning environments (e.g., Slavin, 1983), it is important for research to identify which features of the teaching process contribute to learning.

Another limitation of past research is its focus on learning by teaching in the absence of teaching expectancy—that is, prior to learning participants may not be told that they will later teach the material (e.g., Roscoe & Chi, 2008). In other words, the effects of learning by teaching and learning by preparing to teach have largely been examined separately (e.g., Bargh & Schul, 1980; Benware & Deci, 1984; Roscoe & Chi, 2008; see Annis, 1983 for the one exception). This divergence causes difficulty in determining the role of preparing to teach within the broader spectrum of learning by teaching. In particular, little is known about the relative benefits of preparing to teach and actually teaching, whether these activities reflect the same or different cognitive processes, and whether or to what extent actually teaching goes beyond simply preparing to teach (and under which conditions). In short, research is needed to disentangle the potential contributions of preparing for and actually teaching to learning.

The goal of the present study is to begin addressing these issues by directly testing the relative learning benefits of preparing to teach and actually teaching academic content to others. In the current study, students in two experiments were asked to study a brief paper-based multimedia lesson on the Doppler Effect. Some students studied the lesson without the expectation of later teaching the material and then took a comprehension test on the material (control group). Other students studied the same lesson with instructions that they would later teach the material; of those expecting to teach, some participants actually taught the material before being tested (teaching group), whereas others only prepared to teach before being tested (preparation group). The teaching activity consisted of students providing a brief video-recorded lecture explaining how the Doppler Effect works as if the recording was to be used to teach another student with no prior knowledge of the subject. The rationale for using this mode of teaching was to isolate and systematically test the effects of the essential feature of learning by teaching—explaining content to others for instructional purposes—and examine its relationship with preparing to teach and studying normally.

Performance was compared across groups on a paper-based comprehension test that required participants to explain important concepts related to the Doppler Effect. In Experiment 1, participants were tested immediately following completion of their respective learning activity (i.e., normal studying, preparing for teaching, or preparing for and actually teaching), whereas in Experiment 2, participants were tested after a one-week delay following their respective learning activity. Overall, the two experiments aimed to examine the relative effects of preparing to teach and actually teaching, and further, to determine whether any potential benefits gained by preparing for or actually teaching remain persistent over time. In short, the goal of this research is to contribute towards a research-based understanding of how learning by teaching works.

1.1. Learning by preparing to teach

In their classic article, Bargh and Schul (1980) postulated that the cognitive benefits experienced through teaching others are a product of the preparation that takes place prior to teaching, the initial presentation of information to students, and the subsequent interactions with students (e.g., answering questions, providing feedback). Thus, they provided an early framework for distinguishing between different stages of learning by teaching, as well as for studying the cognitive mechanisms underlying each stage. In particular, their article proposed that the mere expectation of teaching others may change the way students study material, compared to studying normally for one's self (e.g., for a test). In other words, studying with the expectation of later teaching may alter one's cognitive processing during learning by priming students to devote more resources toward selecting the most relevant material and organizing it into a meaningful representation. In one experiment, Bargh and Schul provided early support for this potential difference in processing. Participants were given verbal material to study with the expectation of either answering questions afterwards or teaching the material to another person. The results found that those expecting to teach performed better on a subsequent retention test than those expecting to answer questions. Surprisingly few studies have further explored the effects of teaching expectancy (e.g., Annis, 1983; Benware & Deci, 1984; Renkl, 1995). For example, following a similar design to that of the Bargh and Schul (1980) study, Benware and Deci (1984) found that students expecting to teach performed better on a measure of conceptual learning after studying an article on brain functioning than did students who did not expect to teach. Overall, these studies provide some preliminary support for a teaching expectancy effect (although see Renkl, 1995 for an exception), presumably because preparing to teach helps students better select and organize important information from a lesson.

Annis (1983) conducted the only study found in which the effects of teaching expectancy and actually teaching were directly compared. In the experiment, participants either read a history passage with the expectation of later being asked to recall the material or the expectation of tutoring another student on the material. Of those expecting to tutor, some participants actually tutored the material to someone else, whereas some participants only prepared to tutor someone else. The results provided some evidence that expecting to tutor may enhance learning beyond studying normally, and further, that tutoring another student enhanced learning beyond only preparing to tutor. Although this finding provides early evidence of an added benefit for tutoring, there are important limitations of Annis' study that are in need of further investigation. From a theoretical standpoint, one limitation is that the students who tutored interacted with another student (e.g., answering questions, providing and receiving feedback). According to Bargh and Schul (1980), interactions with students represent an additional stage of learning by teaching beyond only explaining to others. Therefore, it is unclear whether the added benefits of tutoring can be attributed to explaining material to another student or the various interactions that take place with the other student. Thus, the goal of the current study is to isolate the effects of two essential components of teaching—preparing to teach and explaining content to others—to test whether explaining content to others offers additional learning gains beyond preparing to teach. In addition to isolating and disentangling these two stages of teaching, the current study extends previous work by including both immediate tests (Experiment 1) and delayed tests (Experiment 2). Delayed tests are particularly relevant in light of a growing research base on desirable difficulties, showing that the effects of learning strategies that create difficulties during learning such as self-testing can best be distinguished from other effects on a delayed test (Dunlosky, Rawson, Marsh, Nathan, & Willingham, 2013; Roediger & Karpicke, 2006). Thus, any added benefit of explaining to others over preparing to teach may be most pronounced on a delayed test.

In summary, the existing literature provides some support for teaching expectancy effects, suggesting that studying with the expectation of later teaching may play a critical role in determining the overall effects of learning by teaching. In other words, the prospect of teaching others may help students better select the most important material from the lesson and organize it into a meaningful representation (Bargh & Schul, 1980). However, research has yet to determine whether the act of explaining to others for instructional purposes (i.e., teaching) provides unique contributions to learning beyond preparing to teach. Addressing this issue is the primary goal of the present set of experiments.

1.2. Learning by teaching

The learning by teaching literature is dominated by research on the benefits peer tutoring provides for the tutor (Cohen et al., 1982; Cook, Scruggs, Mastropieri, & Casto, 1986; Greenwood, Delquadri, & Hall, 1989; Mastropieri, Spencer, Scruggs, & Talbott, 2000; Mathes & Fuchs, 1994; Rohrbeck et al., 2003; Roscoe & Chi, 2007). Several studies have shown that peer tutoring can lead to learning gains for both the tutor and the tutee (Rohrbeck et al., 2003; Roscoe & Chi, 2007). In fact, some have found peer tutoring to be just as beneficial for the tutor as the student being tutored (e.g., Chi et al., 2001; Cloward, 1967; Fantuzzo, King, & Heller, 1989; Graesser et al., 1995; Kafai & Harel, 1991). Overall, research suggests that peer tutoring can result in tutor learning across different formats (e.g., cross-age, same-age, fixed roles, reciprocal roles), as well as for various academic disciplines, such as math, reading, and science (Roscoe & Chi, 2007). However, because of the diversity of peer tutoring environments, as well as the complex nature of interactions between tutors and tutees, what makes individual programs effective is largely unclear (Roscoe & Chi, 2007; Topping & Ehly, 2001). Accordingly, the results of studies investigating tutor learning vary considerably (Rohrbeck et al., 2003; Roscoe & Chi, 2007), producing a relatively small overall effect of approximately $d = .35$ (Cohen et al., 1982; Rohrbeck et al., 2003; Roscoe & Chi, 2007). Further, examining variables such as format type (e.g., cross-age, reciprocal), population (e.g., elementary school, high school), or subject area (e.g., math, science) do not appear to offer any significant insight toward understanding when and how peer tutoring environments support tutor learning (Roscoe & Chi, 2007).

One explanation is that the effectiveness of individual peer tutoring programs depends on several different factors associated with the design and implementation of the program, including the nature of tutor training, the level of autonomy granted to tutors and tutees, and the duration of the program. However, Roscoe and Chi (2007, 2008) argue that the effectiveness of peer tutoring may mostly depend on the quality of tutor-tutee interactions, such as the nature of the explanations, answers, and feedback provided by the tutor. This is consistent with the idea that learning gains take place because of the activities occurring during the teaching process, such as explaining, answering questions, and providing feedback (Cohen, 1986; Gartner, Kohler, & Riessman, 1971; King, Staffieri, & Adelgais, 1998). In particular, the research of Roscoe and Chi suggests that tutors' gains are to some extent dependent on whether tutors engage in *reflective knowledge building*—that is, the degree to which tutors reflect upon their own understanding of the material and build upon their own prior knowledge while explaining to and answering questions from tutees. On the other hand, tutor learning is unlikely to occur when tutors engage in *knowledge telling*, in which they simply summarize the material without integrating it with their prior knowledge. In other words, the effects of learning by teaching appear to depend on the extent to which the 'teacher' engages in generative processing (Mayer, 2005, 2009; Wittrock, 1989)—that is, cognitive processing devoted towards organizing the material into a meaningful representation and integrating this newly acquired information with prior knowledge.

Overall, research on tutor learning has progressed from examining factors related to the design of peer tutoring programs to factors related to the specific behaviors of tutors and tutees. While this research has contributed to a better understanding of the processes involved during the tutoring process, it is not aimed at disentangling the relative contribution of teaching expectancy to learning. In contrast, the purpose of the current study is to contribute to the study of learning by teaching by systematically determining the relative effects of preparing to teach and actually teaching on both immediate and delayed measures of comprehension.

1.3. Theory and predictions

A learning strategy is an activity in which the learner engages that is intended to promote learning (Weinstein & Mayer, 1985). A *generative* learning strategy is intended to promote meaningful learning (i.e., learning for understanding) by priming deep cognitive processes during learning, such as priming the learner to organize the material into a coherent structure and integrate it with relevant prior knowledge (Mayer, 2008; Mayer & Wittrock, 1996, 2006; Wittrock, 1989). In the present study, we consider the act of teaching (which can be called learning by teaching) to be a generative learning strategy and seek to determine the extent to which any benefits of teaching can be attained simply by preparing to teach.

According to the cognitive theory of multimedia learning, deep learning occurs when learners engage in appropriate cognitive processing during learning, including selecting relevant material from a lesson for further processing, mentally organizing it into a coherent structure in working memory, and integrating it with relevant prior knowledge activated from long-term memory (Mayer, 2009, 2011). Furthermore, the cognitive theory of multimedia learning proposes—similar to cognitive load theory (Sweller, 2005; Sweller, Ayres, & Kalyuga, 2011)—that working memory is limited in capacity. Thus, a generative learning strategy such as learning by teaching can prime the learner to engage in *generative processing*—that is, processing aimed at making sense of the material such as organizing and integrating—but the stress of teaching can create *extraneous processing*—that is, processing that does not serve an instructional goal and thereby wastes limited processing capacity. If the learner is able to minimize extraneous processing and engage in generative processing during learning, the learner should be able to construct a learning outcome that can support comprehension test performance over time.

Based on this theoretical framework, we predict that learning by teaching will improve students' comprehension of to-be-learned material on both immediate and delayed tests. Research on generative learning strategies such as self-testing (Dunlosky et al., 2013; Roediger et al., 2006) shows that the positive effects of generative learning strategies relative to other learning approaches are most pronounced on delayed tests, which indicate that a deep strategy such as learning by teaching should have lasting effects on what is learned as compared to more passive strategies. Simply preparing to teach may prime some of the initial cognitive processes needed for deep learning such as selecting relevant material and mentally representing it (which is called *essential processing* in the cognitive theory of multimedia learning). Having a newly constructed mental representation of the presented material should be helpful on an immediate test, but the minimal amount of deeper processing during learning (i.e., generative processing that creates internal coherence and external context) may not be sufficient for the representation to persist over time. Thus, we predict that preparing to teach will improve comprehension test performance on immediate tests but not necessarily on delayed tests.

2. Experiment 1

The purpose of Experiment 1 was to test the extent to which actually teaching the content of a to-be-learned lesson improves learning and whether preparing to teach offers some or all of the same benefit. Students studied a brief paper-based multimedia lesson on the Doppler Effect and were tested immediately following studying normally (control group), preparing to teach (preparation group), or preparing for and actually teaching (teaching group). Those who taught the material were asked to give a brief video-recorded lecture on how the Doppler Effect works as if they were teaching another student. In short, this study sought to determine whether each stage of learning by teaching offers additive learning gains over normal studying when students are given an immediate comprehension test.

2.1. Method

2.1.1. Participants and design

The participants were 93 undergraduate students recruited from the Psychology Subject Pool at the University of California, Santa Barbara. Thirty served in the teaching group, 32 served in the preparation group, and 31 served in the control group. The mean age was 19.0 (SD = 1.13), and there were 26 men and 67 women. The groups did not differ significantly in terms of mean age or proportion of men and women. Participants had low prior knowledge of the Doppler Effect, as reported on a questionnaire with a possible score of 13 (α = .70), and there was no difference in prior knowledge across conditions (Teaching: M = 3.8; SD = 2.2; Preparation: M = 4.5; SD = 2.7; Control: M = 5.3; SD = 2.4). Participants also did not differ across conditions in their ratings of self-efficacy (i.e., how well they thought they would perform on a test after studying a lesson on the Doppler Effect), as indicated on a scale ranging from 1 (very poor) to 5 (very well) (Teaching: M = 2.8; SD = .84; Preparation: M = 2.8; SD = .74; Control: M = 2.6; SD = .77).

2.1.2. Materials

The paper-based materials consisted of a consent form, a demographics form, an instructions sheet, a paper-based lesson on the Doppler Effect, a comprehension test, and a post-questionnaire. The consent form described the details of the study, informed participants they may be videotaped during the experiment and that their privacy was protected, and included a place for them to sign. The demographics questionnaire asked participants to provide their age, gender, and SAT scores. They were also asked to rate their knowledge of the Doppler Effect on a scale from 1 (very low) to 5 (very high) and to place a check mark next to each of the following items that applied to them: "I have taken a college course in physics," "I know what Hz means," "I have used an oscilloscope," "I know how radar works," "I know the basic characteristics of sound waves," "I know what relative motion is," "I know what the red shift is," and "I know what a sine curve is." The self-reported rating of prior knowledge and the checklist served as a measure of participants' prior knowledge of the Doppler Effect, with a combined possible score of 13 (i.e., one point for each checked item on the list plus 1 through 5 based on which level was checked on the rating scale with 5 = very high and 1 = very low). Finally, the demographics form asked participants to report their self-efficacy by rating how well they thought they would perform on a test after studying a lesson on the Doppler Effect on a scale from 1 (very poorly) to 5 (very well), which was used to determine if the groups were equated on reported self-efficacy.

The instructions sheet described the task to participants that they would be expected to do after studying a short lesson on the Doppler Effect. Participants received one of two versions of the instructions sheet depending on the condition to which they were randomly assigned. The control group was told that they would be expected to answer questions on the material they studied:

> *In this experiment, you will study a short lesson on how the Doppler Effect works and then be asked to answer some questions about what you learned. You will have 10 minutes to study the lesson. You may take notes on the lesson itself and/or use the blank sheet of paper provided to you by the experimenter. Please continue working until your time has expired. If you have any questions, please ask the experimenter now. He/she will not be able to answer questions once you have begun.*

The preparation and teaching groups were told they would be expected to teach the material they studied:

> *In this experiment, you will study a short lesson on how the Doppler Effect works and then be asked to teach the material that you learned. Specifically, you will be expected to provide a short lecture explaining how the Doppler Effect works as if you were teaching the material to someone else. Your lecture will be videotaped, and you will have the option of using a whiteboard to illustrate your explanation. You will have 10 minutes to study the lesson and prepare to teach the material. You may take notes on the lesson itself and/or use the blank sheet of paper provided to you by the experimenter. If you have any questions, please ask the experimenter now. He/she will not be able to answer questions once you have begun.*

The Doppler Effect lesson consisted of two single-sided pages of both illustrations and text. The text was single-spaced and consisted of approximately 600 words. First, a concrete example of the Doppler Effect was provided in which the reader was asked to imagine a fire-truck approaching and passing by them. The siren's perceived pitch is described as getting higher as the fire truck gets closer to them and lower as the fire truck moves further away. Second, a description of the basic characteristics of sound waves was provided, including a definition and illustration of wave frequency and wavelength. Third, an analogy was described comparing the Doppler Effect to a bug jiggling on the surface of a pond. Specifically, if the bug is stationary, the water waves it produces will occur at the same frequency and wavelength in all directions; if the bug moves to the right, the waves will occur more frequently to the right of the bug (i.e., in the direction of the bug's movement) and less frequently to the left of the bug. Finally, the lesson returned to the original scenario of an approaching fire truck and explained how the fire truck's motion influences the behavior of sound waves and how as a result, sound is perceived differently as the fire truck approaches and as it passes an observer. Participants were given 10 min to study the lesson; pilot testing indicated that this was a sufficient amount of time.

The comprehension test consisted of six free-response questions designed to assess participants' ability to explain key concepts related to sounds waves and motion (i.e., each question targeted the *comprehension level* of Bloom's Taxonomy; Anderson et al., 2001): "Explain how the Doppler Effect works," "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 at 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 on a parallel path, at a constant speed? Explain your answer." Participants were given 3 min to complete the first question (i.e., "Explain how the Doppler works") and 2 min to complete each of remaining five questions. Pilot testing of the questions indicated that sufficient time was provided to write a response to each question.

Regarding scoring, the first comprehension question was scored by awarding one point for each of thirteen possible components extracted from the lesson, which make up the explanation of how the Doppler Effect works. For example, participants earned points for describing how motion changes the way sound is perceived by an observer, for correctly defining wavelength and frequency, for describing the relationship between frequency, wavelength, and pitch, and for describing how the motion of a sound source influences the behavior of sound waves (i.e., frequency and wavelength) and the way in which the sound is perceived by an observer (i.e., pitch). For the remaining five questions, one point was assigned for each correct component of the answer, with questions varying in value from 2 to 3 possible points. For example, regarding the second comprehension question (i.e., "How could you increase the intensity of the Doppler Effect?"), correct responses included increasing the velocity of the sound source toward the observer or moving the observer toward an approaching sound source. The total possible score on the comprehension test was 25 points.

The post-questionnaire was used primarily for exploratory purposes and asked participants to report how much they agreed with each of seven statements on a seven-point scale ranging from 1 (strongly disagree) to 7 (strongly agree): "I felt the subject matter was difficult," "I enjoyed learning about the Doppler Effect," "I would like to learn this way in the future," "I feel like I have a good understanding of how the Doppler Effect works," "After this lesson, I would be interested in learning more about the Doppler Effect," "I found the lesson about the Doppler Effect to be useful to me," and "I felt stressed while I was learning about the Doppler Effect." It also asked participants to rate the amount of mental effort they invested while learning about the Doppler Effect on a scale ranging from 1 (very low effort) to 7 (very high effort). Finally, the post-questionnaire asked participants to write any additional comments they had about the study.

The apparatus consisted of one Cannon FS400 camcorder used to record the lectures given by participants assigned to the teaching group.

2.1.3. Procedure

Participants were randomly assigned to a treatment group. There were up to three participants in each session, with each participant seated in an individual cubicle out of sight from the other participants. Those assigned to the teaching condition participated individually. First, the experimenter provided a brief verbal introduction to the experiment, passed out the consent form for participants to sign, and collected the signed consent forms. Second, the experimenter passed out the demographics questionnaire and collected them when the participants were finished. Third, the experimenter passed out the appropriate instructions sheet. Those in the control group were informed that they would have 10 min to study a lesson on the Doppler Effect and then would be expected to answer questions on the material. Those in the preparation and teaching groups were informed that they would have 10 min to study a lesson on the Doppler Effect and then would be expected to teach the material. Specifically, they were told they would be expected to provide a brief lecture of up to 5 min explaining how the Doppler Effect works as if they were teaching the material to someone else. All groups were informed they could take notes on the lesson itself and that they could take notes on a blank sheet of paper, which was provided to them.

Following a 10-min study period, the experimenter collected the lesson and any notes taken by participants. Participants in the control and preparation groups were then given the comprehension test, whereas participants in the teaching group were asked to teach the material before taking the comprehension test. Specifically, the teaching group was given up to 5 min to explain how the Doppler Effect works as they were video-recorded via camcorder, with instructions to teach as if the recording was to be used to teach another student. Participants in the teaching group were asked to stand while giving their lecture, and they were given the option of using a whiteboard on the wall if they wished to illustrate their explanation.

Upon completion of their respective task (i.e., studying normally, preparing to teach, or actually teaching), participants completed the comprehension test. Participants in the preparation group were told before the tests were administered that they would answer some questions on the material before providing their video-recorded lecture. Each question on the comprehension test was timed and completed individually; 3 min were provided for the first question, and 2 min were provided for each of the remaining five questions. After completing the comprehension test, participants then completed the post-questionnaire. Finally, all participants were debriefed and thanked for their participation. The total duration of the experiment was approximately 35 min for the control and preparation groups and approximately 40 min for the teaching group.

Table 1
Mean comprehension test score (and SD) for three groups in Experiment 1 (immediate test).

Group	*n*	Comprehension score		Effect size (*d*)
		M	*SD*	
Control	31	6.2	3.3	
Preparation	32	7.9*	2.4	.59
Teaching	30	8.7*	2.8	.82

* Significantly different from control group at $p < .05$.

2.1.4. Results

The primary purpose of this study was to test the relative effects of preparing to teach and actually teaching on the comprehension test. Table 1 shows the means and standard deviations of each group on the comprehension test. The two authors, blind to experimental conditions, scored all responses for each participant ($\alpha = .94$), and discrepancies between scores were settled by consensus. A one-way analysis of variance (ANOVA) indicated that comprehension performance differed significantly across groups, $F(2, 90) = 5.92$, $p = .004$, $\eta^2 = .116$. Dunnett post hoc tests at $\alpha < .05$, with the control group serving as the reference group, revealed that the teaching group and preparation group each significantly outperformed the control group (Teaching vs. Control: $d = .82$; Preparation vs. Control: $d = .59$). Finally, a Z-test comparing the two effect sizes (see Rosenthal, 1991, pp. 71–72 for calculation) revealed that the teaching effect was not significantly different from the teaching expectancy effect ($Z = 1.61$, $p > .05$). Overall, this finding is consistent with the hypothesis that actually teaching does not offer additional learning benefits beyond preparing to teach on an immediate measure of comprehension.[1]

3. Experiment 2

The purpose of Experiment 2 was to investigate whether actually teaching had an effect on comprehension test score beyond preparing to teach when the comprehension test was administered after a one-week delay. Therefore, Experiment 2 consisted of two parts. In Part 1, students were given the lesson on the Doppler Ef-

[1] For exploratory purposes, we also analyzed data from the post-questionnaire (which assessed students' effort, perceived difficulty, and satisfaction with the learning environment). A one-way ANOVA indicated no significant differences across groups; however, this may be due to the time at which students answered completed the questionnaire (i.e., following the comprehension test, rather following the lesson).

fect and either studied the material normally (control group), expected to teach but did not teach the material (preparation group), or expected to and actually taught the material (teaching group). In Part 2, students returned one week later to complete a comprehension test on the material. Experiment 2 was intended to assess the pure effects from delayed testing (i.e., without any influence of taking an immediate test before taking the delayed test); thus, only a delayed test was administered. Overall, this study sought to determine whether the equivalent effects of preparing to teach and actually teaching found on an immediate comprehension test in Experiment 1would persist on a delayed comprehension test in Experiment 2.

3.1. Method

3.1.1. Participants and design

The participants were 75 undergraduate students recruited from the Psychology Subject Pool at the University of California, Santa Barbara. Twenty-five participants served in the teaching group, 25 served in the preparation group, and 25 served in the control group. The mean age was 19.4 ($SD = 1.2$), and there were 31 men and 44 women. The groups did not differ significantly in terms of mean age or proportion of men and women. As in Experiment 1, participants did not differ in performance on the prior knowledge test ($\alpha = .74$) across conditions (Teaching: $M = 4.4$; $SD = 2.1$; Preparation: $M = 5.3$; $SD = 3.1$; Control: $M = 5.4$; $SD = 2.8$). Also as in Experiment 1, participants did not differ across conditions in their self-efficacy ratings of how well they thought they would perform on a test after studying a lesson on the Doppler Effect (Teaching: $M = 2.7$; $SD = .74$; Preparation: $M = 3.2$; $SD = .65$; Control: $M = 3.0$; $SD = 1.1$).

3.1.2. Materials

The materials and apparatus were identical to those of Experiment 1.

3.1.3. Procedure

The procedure was identical to that of Experiment 1, except that the comprehension test was administered following a one-week delay. Participants completed the post-questionnaire upon completion of the comprehension test.

3.1.4. Results

The purpose of Experiment 2 was to test whether the effects found in Experiment 1 would persist if comprehension performance were measured following a one-week delay. Table 2 below shows the means and standard deviations of each group. A one-way ANOVA again indicated that comprehension performance differed significantly across groups, $F(2, 72) = 5.00$, $p = .009$, $\eta^2 = .122$. As in Experiment 1, Dunnett post hoc tests at $\alpha < .05$ revealed that the teaching group outperformed the control group ($d = .79$); however, in contrast to Experiment 1, the preparation group did not significantly outperform the control group ($d = .24$). Further, a Z-test comparing the two effects (Rosenthal, 1991) showed that the teaching effect was significantly higher than the teaching expectancy effect ($Z = 2.75$, $p < .05$). Therefore, the data from Experiment 2 are consistent with our hypothesis that the effects of actually teaching would remain strong on the delayed comprehension test, whereas the effects of preparing to teach would be considerably reduced.[2]

Table 2
Mean comprehension test score (and SD) for three groups in Experiment 2 (delayed test).

Group	n	Comprehension score		Effect size (d)
		M	SD	
Control	25	5.0	3.0	
Preparation	25	5.6	2.0	.24
Teaching	25	7.3*	2.8	.79**

* Significantly different from control group at $p < .05$.
** Teaching effect is significantly higher than teaching expectancy effect at $p < .05$.

4. Discussion

4.1. Empirical contribution

The goal of this study was to determine the relative benefits of learning by teaching and learning by preparing to teach (i.e., teaching expectancy) when compared to a control group that studied normally. Consistent with our hypotheses, the results show evidence of a teaching effect for both immediate (Experiment 1; $d = .82$) and delayed (Experiment 2: $d = .79$) comprehension performance. Further, the results show evidence for a significant teaching expectancy effect on the immediate comprehension test ($d = .59$) but not the delayed comprehension test ($d = .24$). Overall, the primary empirical contribution of this study is that actually teaching appears to improve student comprehension beyond preparing to teach on delayed measures of learning but not on an immediate test. In other words, explaining content to others appears to offer additional cognitive benefits than experienced through only preparing to teach.

These findings also suggest an important methodological consideration regarding the use of both immediate and delayed learning outcome measures. In particular, using only immediate measures in the current study would have led us to conclude that no considerable learning gains are acquired through the act of teaching beyond only preparing to teach. However, the delayed tests reveal a critical distinction between the effects of preparing for and actually teaching—namely, that actually teaching results in a more persistent understanding of the material beyond that which is acquired while preparing to teach.

Finally, the current study included an exploratory post-questionnaire to assess students' invested effort, perceived difficulty, and satisfaction with the learning environment during learning. The results across both experiments indicated no evidence that preparing to teach or explaining to others leads to greater motivation to invest effort during learning, or to a greater sense of satisfaction with the learning environment. This is somewhat surprising, given the differences in learning outcomes found in both experiments. However, the timing at which the post-questionnaire was administered (i.e., at the end of the experiment rather than immediately following learning) may have contributed to this result.

4.2. Theoretical contribution

On the theoretical level, the results are consistent with the predictions in which a learning strategy that tends to promote generative processing during learning (i.e., learning by teaching) helps learners build long-lasting understanding compared to one that tends to mainly promote essential processing (i.e., learning by preparing to teach). Consistent with the cognitive theory of multimedia learning (Mayer, 2005, 2009), learning by teaching is intended to encourage learners to select the most relevant material from the

[2] As in Experiment 1, an ANOVA revealed no significant differences across groups regarding the eight items on the exploratory post-questionnaire. This is again possibly due to asking participants to retroactively report their levels of effect, perceived difficulty, and satisfaction experienced during learning, after they completed the comprehension test.

lesson, organize it into a meaningful representation, and integrate it with their prior knowledge. In line with previous work on the teaching expectancy effect (Annis, 1983; Bargh & Schul, 1980), it was expected that preparing to teach would promote the cognitive processing necessary for selecting the key elements of information from the lesson and organizing these elements as presented in the lesson. In line with previous work on the tutor learning effect (Roscoe & Chi, 2007), it was expected that actually teaching would promote the cognitive processing necessary for integrating the lesson content with information from prior knowledge by elaborating on and reorganizing the representation developed during preparation. The findings of this study provide preliminary support for the idea that preparing to teach and actually teaching might prime different cognitive processes during learning. Specifically, the generative nature of actually teaching appears to promote deeper processing than that which occurs while preparing to teach. The results of Experiment 2 are particularly supportive of this notion, suggesting that the effects of actually teaching are maintained over time, whereas the effects of teaching expectancy becomes less stable over time. Overall, the current study provides preliminary evidence towards developing a theoretical understanding of how learning by teaching works.

4.3. Practical contribution

On the practical level, the current study provides support for the idea that large improvements in student understanding can be achieved through engaging in a basic teaching task. In particular, considerable learning gains were found even when students were only asked to provide a very short (i.e., less than 5 min) video-recorded lecture of the material *as if* they were teaching someone else. Thus, learning was enhanced even without the physical presence of a student to be taught, and therefore, without any of the student–teacher interactions which may or may not be unique to teaching (e.g., providing feedback, receiving feedback, answering questions).

These findings provide important implications for educational practice, by showing that the full benefits or learning by teaching depend on students actually producing prepared explanations of the material. This implies that merely studying *as if* one had to later teach the material may not be as effective for students to achieve long-term learning. In other words, the findings of the current study to some extent validate the efficacy of one central component (i.e., explaining to others) of many popular classroom activities, including small-group discussion (e.g., Webb, 1982), reciprocal teaching (e.g., Palincsar & Brown, 1984), and other cooperative learning environments (e.g., Slavin, 1983). In short, this study supports the idea that learning activities that require students to both prepare and produce explanations for others can result in better long-term understanding of the material.

4.4. Limitations and future directions

One criticism of the current study is that the teaching group performed an additional activity (i.e., presenting a video-recorded lecture) compared to the preparation and control groups, and therefore, simply spending additional time with the material may have been responsible for learning gains. However, participants in the teaching condition did not have the lesson or their notes available to them while presenting their video-recorded lecture. Thus, any additional time that students in the teaching group used to generate their lecture (i.e., up to 5 min) cannot be referred to as extra study time. Rather, it can be more appropriately referred to as teaching time—time spent retrieving and elaborating on the material acquired during preparation in order to present a coherent explanation that could be understood by another person. Therefore, the teaching group did have more time during which they could rehearse the material from the lesson. Future research should investigate whether this finding holds if the preparation and control groups were provided with additional time.

Another potential limitation of the current study is the nature of the Doppler Effect lesson used. The lesson used was relatively simple and brief (i.e., approximately 600 words). Therefore, further work is needed to test the effects of learning by teaching and teaching expectancy for more complex learning materials. Importantly, the Doppler lesson did consist of meaningful content involving a cause-and-effect explanation of a scientific system, extending past research investigating the relationship between teaching expectancy and actually teaching (Annis, 1983). Along those lines, future research should explore the applicability of these findings to other academic domains, such as explaining how to perform a mathematical procedure to others.

The external validity of the teaching expectancy condition is also a potential limitation that should be considered when interpreting the results of the current study. In other words, participants completed the comprehension test likely under the impression that they will still teach the material after testing. Although this was done to avoid influencing their performance by telling them they actually were not teaching the material, it also means that the preparation group may not have completed the comprehension test under typical conditions that may occur in the classroom.

Finally, work is needed to determine the degree to which learning by teaching *depends* on teaching expectancy and preparing to teach. For example, it may be that the strong effects found for learning by teaching in this study are due to participants also preparing with the expectation of later teaching. It is likely that learning by teaching would not be as effective if students studied the material without the expectation of teaching before being asked to teach. In other words, the strong effects of learning by teaching may be mediated by how students prepare prior to teaching. This possibility could provide important implications for educational theory and practice, and should be explored in subsequent studies.

Acknowledgment

This research was supported by Grant N000140810018 from the Office of Naval Research.

References

Allen, V., & Feldman, R. (1973). Learning through tutoring: Low-achieving children as tutors. *Journal of Experimental Education, 42*(1), 1–5.
Anderson, L. W., Krathwohl, D. R., Airasian, P. W., Cruikshank, K. A., Mayer, R. E., Pintrich, P., et al. (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives*. New York: Longman.
Annis, L. F. (1983). The processes and effects of peer tutoring. *Human Learning: Journal of Practical Research and Applications, 2*(1), 39–47.
Bargh, J. A., & Schul, Y. (1980). On the cognitive effects 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*, 755–765.
Chi, M., Siler, S., Jeong, H., Yamauchi, T., & Hausmann, R. (2001). Learning from human tutoring. *Cognitive Science, 25*, 471–533.
Cloward, R. (1967). Studies in tutoring. *Journal of Experimental Education, 36*(1), 14–25.
Cohen, J. (1986). Theoretical considerations of peer tutoring. *Psychology in the Schools, 23*, 175–186.
Cohen, P., Kulik, J., & Kulik, 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 scientific texts. *The Journal of the Learning Sciences, 6*(4), 347–365.
Cook, S., Scruggs, T., Mastropieri, M., & Casto, G. (1986). Handicapped students as tutors. *Journal of Special Education, 19*(4), 483–492.
Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J., & Willingham, D. T. (2013). Improving students' learning with effective learning techniques: Promising

directions for cognitive and education psychology. *Psychological Science in the Public Interest, 14*, 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.

Fantuzzo, J., King, J., & Heller, L. (1989). Effects of reciprocal peer tutoring on mathematics and school adjustment: A componential analysis. *Journal of Educational Psychology, 81*(2), 173–177.

Galbraith, J., & Winterbottom, M. (2011). Peer tutoring: What's in it for the tutor? *Educational Studies, 37*(3), 321–332.

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

Ginsburg-Block, M., & Fantuzzo, J. (1997). Reciprocal peer tutoring: An analysis of "teacher" and "student" interactions as a function of training and experience. *School Psychology Quarterly, 12*(2), 134–149.

Graesser, A., Person, N., & Magliano, J. (1995). Collaborative dialogue patterns in naturalistic one-to-one tutoring. *Applied Cognitive Psychology, 9*, 495–522.

Greenwood, C., Delquadri, J., & Hall, R. (1989). Longitudinal effects of classwide peer tutoring. *Journal of Educational Psychology, 81*(3), 371–383.

Griffin, M. M., & Griffin, B. W. (1998). An investigation of the effects of reciprocal peer-tutoring on achievement, test anxiety, and academic self-efficacy. *Contemporary Educational Psychology, 23*, 298–391.

Kafai, Y., & Harel, I. (1991). Learning through design and teaching: Exploring social and collaborative aspects of constructionism. In I. Harel & S. Papert (Eds.), *Constructionism*. Norwood, NJ: Ablex.

King, A., Staffieri, A., & Adelgais, A. (1998). Mutual peer tutoring: Effects of structuring tutorial interaction to scaffold peer learning. *Journal of Educational Psychology, 90*(1), 134–152.

Mastropieri, M., Spencer, V., Scruggs, T., & Talbott, E. (2000). Students with disabilities as tutors: An updated research synthesis. In T. E. Scruggs & M. A. Mastropieri (Eds.), *Educational interventions: Advances in learning and behavioral disabilities* (pp. 247–279). Stamford, CT: JAI.

Mathes, P., & Fuchs, L. (1994). The efficacy of peer tutoring in reading strategies for students with mild disabilities: A best-evidence synthesis. *School Psychology Review, 23*(1), 59–80.

Mayer, R. E. (2008). *Learning and instruction* (2nd ed). Upper Saddle River, NJ: Pearson.

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

Mayer, R. E. (2011). *Applying the science of learning*. Upper Saddle River, NJ: Pearson Merrill Prentice Hall.

Mayer, R. E. (2005). Cognitive theory of multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 31–48). New York, NY: Cambridge University Press.

Mayer, R. E., & Wittrock, M. C. (2006). Problem solving. In P. Alexander, P. Winne, & G. Phye (Eds.), *Handbook of educational psychology* (pp. 287–303). Mahwah. NJ: Erlbaum.

Mayer, R. E., & Wittrock, M. C. (1996). Problem solving and transfer. In D. Berliner & R. Calfee (Eds.), *Handbook of educational psychology* (pp. 45–61). New York: Macmillan.

Morgan, R., & Toy, T. (1970). Learning by teaching: A student-to student compensatory tutoring program in a rural school system and its relevance to the educational cooperative. *Psychological Record, 20*, 159–169.

Palincsar, A. S., & Brown, A. L. (1984). Reciprocal teaching of comprehension-fostering and comprehension-monitoring activities. *Cognition and Instruction, 1*, 117–175.

Ploetznerl, R., Dillenbourg, P., Praier, M., & Traum, D. (1999). Learning by explaining to oneself and to others. In P. Dillenbourg (Ed.), *Collaborative-learning: Cognitive and computational approaches* (pp. 103–121). Oxford: Elsevier.

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

Robinson, D., Schoefield, J., & Steers-Wentzell, K. (2005). Peer and cross-age tutoring in math: Outcomes and their design implications. *Educational Psychology Review, 17*(4), 327–362.

Roediger, H. L., III, & Karpicke, J. D. (2006). Test-enhanced learning: Taking memory tests improves long-term retention. *Psychological Science, 17*, 249–255.

Rohrbeck, C., Ginsburg-Block, M., Fantuzzo, J., & Miller, T. (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). Understanding tutor learning: Knowledge-building and knowledge-telling in peer tutors' explanations and questions. *Review of Educational Research, 77*(4), 534–574.

Roscoe, R. D., & Chi, M. T. H. (2008). Tutoring learning: The role of explaining and responding to questions. *Instructional Science, 36*, 321–350.

Rosenthal, R. (1991). *Meta-analytic procedures for social research*. Beverly Hills, CA: Sage.

Slavin, R. E. (1983). *Cooperative learning*. New York: Longman.

Sweller, J., Ayres, P., & Kalyuga, S. (2011). *Cognitive load theory*. New York: Springer.

Sweller, J. (2005). Implications of cognitive load theory for multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 19–30). New York: Cambridge University Press.

Topping, K., & Ehly, S. (2001). Peer assisted learning: A framework for consultation. *Journal of Educational and Psychological Consultation, 12*(2), 113–132.

Webb, N. M. (1982). Peer interaction and learning in cooperative small groups. *Journal of Educational Psychology, 74*(5), 642–655.

Weinstein, C. E., & Mayer, R. E. (1985). The teaching of learning strategies. In M. C. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed., pp. 315–327). New York: Macmillan.

Wittrock, M. C. (1989). Generative processes of comprehension. *Educational Psychologist, 24*, 345–376.