

Building Excellence in Scientific Teaching: How Important Is the Evidence for Evidence-Based Teaching when Training STEM TAs?[†]

Lorelei E. Patrick^{1,2}, Hillary A. Barron^{1*}, Julie C. Brown³, and Sehyo Cotner¹

¹Biology Teaching and Learning Department, University of Minnesota, Minneapolis MN 55455;

²Department of Biological Sciences, Fort Hays State University, Hays, KS 67601;

³School of Teaching and Learning, University of Florida, Gainesville, FL 32611

Evidence-based teaching practices (EBTP)—like inquiry-based learning, inclusive teaching, and active learning (AL)—have been shown to benefit all students, especially women, first-generation, and traditionally minoritized students in science fields. However, little research has focused on how best to train teaching assistants (TAs) to use EBTP or on which components of professional development are most important. We designed and experimentally manipulated a series of presemester workshops on AL, dividing subjects into two groups. The Activity group worked in teams to learn an AL technique with a workshop facilitator. These teams then modeled the activity, with their peers acting as students. In the Evidence group, facilitators modeled the activities with all TAs acting as students. We used a mixed-methods research design (specifically, concurrent triangulation) to interpret pre- and postworkshop and postsemester survey responses. We found that Evidence group participants reported greater knowledge of AL after the workshop than Activity group participants. Activity group participants, on the other hand, found all of the AL techniques more useful than Evidence group participants. These results suggest that actually modeling AL techniques made them more useful to TAs than simply experiencing the same techniques as students—even with the accompanying evidence. This outcome has broad implications for how we provide professional development sessions to TAs and potentially to faculty.

INTRODUCTION

Several recent reports act as a “call to arms” for science, technology, engineering, and mathematics (STEM) educators to implement evidence-based teaching practices (EBTP) into their courses to better prepare future scientists and to create a scientifically literate citizenry (1–6). Despite the exhortations for reform, many STEM faculty members are reluctant to change their teaching practices, for reasons ranging from a perceived or actual lack of support to a concern over sacrificed time for covering content (7, 8). Most of these reform efforts have focused on STEM faculty teaching lecture-based courses; until recently, little work has explored lab courses and the teaching assistants (TAs, some of whom may become faculty members) who often teach them. In the United States, graduate student and undergraduate student TAs teach the majority of under-

graduate science lab courses at most colleges and universities (9, 10) and can strongly influence the academic and career trajectories of the students they teach (11). However, traditional TA programs vary in quality and may result in a lower-quality education for undergraduate students (12). Thus, it should come as no surprise that the TAs in charge of many laboratory sections are similarly ill prepared to use EBTP in their own teaching (13). The result is TAs who are unlikely to implement such teaching practices and ill-equipped to facilitate inquiry in laboratory courses, the most natural venues for students to engage in scientific practices and acquire critical STEM competencies. While we imagine the challenges we address here span the STEM disciplines, our focus is on biology courses and their associated TAs.

Although TA professional development (TAPD) can take many forms, precourse workshops are the most common training format offered to biology TAs (9). However, there is a great deal of variability in the duration and content of these workshops, and often the focus is more on course logistics than on pedagogy (13). The Summer Institutes on Scientific Teaching (14, 15)—with its emphasis on active learning, inclusive teaching, and assessment—has formed the basis for TAPD and workshops at several institutions (16, 17). In some workshops, TAs actively model teaching practices to their peers (16); however, in most PD sessions,

*Corresponding author. Mailing address: 3-154 Molecular and Cellular Biology Building, 420 Washington Ave. SE, Minneapolis, MN 55455. Phone: 218-428-2689. E-mail: fult0050@umn.edu.

Received: 5 October 2020, Accepted: 2 December 2020, Published: 29 January 2021

[†]Supplemental materials available at <http://asmscience.org/jmbe>

©2021 Author(s). Published by the American Society for Microbiology. This is an Open Access article distributed under the terms of the Creative Commons Attribution-Noncommercial-NoDerivatives 4.0 International license (<https://creativecommons.org/licenses/by-nc-nd/4.0/> and <https://creativecommons.org/licenses/by-nc-nd/4.0/legalcode>), which grants the public the nonexclusive right to copy, distribute, or display the published work.

TAs watch and participate in teaching activities as if they were students or observers (17–20). General TA workshops for all disciplines can be effective, increasing TA self-efficacy and positive attitudes towards teaching and lowering communication apprehension (21, 22). Workshops that focus on the teaching techniques (e.g., facilitating inquiry) that TAs will need for a specific course may be more effective for both TAs and their students than generalized workshops (12), but this may not be the case in all situations (18). In fact, both approaches have been reported to be effective. Despite a growing body of work detailing PD for STEM TAs (23 and references above), there is little consensus on the importance of specific training components, or how best to evaluate TAPD (17). Similarly, questions abound regarding how best to train TAs to facilitate inquiry using evidence-based techniques in a sustained manner.

Building Excellence in Scientific Teaching for TAs in biology

A new TAPD program at the University of Minnesota, “Building Excellence in Scientific Teaching” (BEST), has the dual aims of (1) developing TAPD modeled after core features of the Summer Institutes faculty PD program and (2) determining which aspects of TAPD are associated with our desired outcomes: acceptance and implementation of evidence-based teaching in the biology teaching laboratory.

An initial workshop offered in our department in fall 2016 engaged 36 TAs in a 2-day presemester workshop to discuss Scientific Teaching (ST). The workshop was facilitated by six faculty alumni of the Summer Institutes, two of whom are veteran Summer Institutes facilitators. Facilitators and participants discussed some of the recent literature on active learning and modeled in-lab strategies for low-stakes assessment, inclusion, and actively engaging students. Feedback from this effort was overwhelmingly positive: participants expressed gains in confidence with respect to learning about ST and strategies for facilitating inquiry. When asked for suggestions for improvement, input included:

- I would love to see something like this offered on a regular basis for teaching assistants in all CBS [College of Biological Sciences] courses in the format of something like a weekly 1-hour discussion (whether it be a course or simply a group meeting). I think having dedicated time to discuss and dialogue about ST concepts is incredibly valuable for continued growth as a teacher and to get feedback and new ideas on how to implement strategies in my own classrooms.
- I felt like a majority of the training was focused on getting the TAs to believe that group work and research laboratory experiences are a good thing for students. We didn't need convincing.

Workshop facilitators were particularly intrigued by the “didn't need convincing” comment above. Facilitators of the Summer Institutes for faculty have recently detected that faculty participants may be entering the Summer Institutes with a higher degree of “buy-in” to the ST model than earlier participants, so Summer Institutes facilitators have been spending less time on the evidentiary basis of ST (J. Blum, personal communication). These comments led us to question which workshop components are necessary and most effective when training TAs. To date, this question has not been satisfactorily addressed. In particular, we questioned how much TAs accept the value of ST and, furthermore, whether discussion of the empirical evidence would impact TA acceptance and use of EBTP.

To determine which training components were most important, we designed and experimentally manipulated a series of presemester TA workshops. Using convenience samples in a quasi-experimental study (24), we used the findings from our first set of workshops to inform the second set of workshops. In the first set, we tested what we call our “Activity versus Evidence” hypothesis: TAs who actively model ST practices but get no evidence for the effectiveness of those activities will find the workshop to be more valuable than TAs who are presented with evidence for ST but do not actively model ST, thereby experiencing the teaching practices from a student's perspective. Building on our results from our first experiment, we formulated and tested a new hypothesis—the “Cherry-On-Top” hypothesis: If we experimentally manipulate workshop groups such that all TAs actively model teaching practices, but only one group is exposed to evidence, there will be no differences between the experimental groups; however, we expect TAs will appreciate knowing the evidentiary basis for the techniques they are learning. In other words, like the cherry on top of an ice cream sundae, evidence is nice but it isn't the most critical part of the TAPD workshops.

We report here on the primary research question motivating this work: Does presenting TAs with the evidentiary basis of ST lead TAs to value and use ST practices and increase their perceived knowledge of ST more than not presenting this evidence?

Conceptual framework

We contextualize our work in the conceptual framework of constructivism. Constructivism posits that meaning-making and learning occur through social dimensions (25). John Dewey contended that thinking, and therefore knowledge construction, was not merely a spontaneous process but, rather, a product of specific experiences (26). Dewey's “transactional constructivism” pragmatically shifts the focus in learning theory to that of the perceptions and habits of the learner or subject (27). We position the TAs in this study as active participants in their own learning of ST practices and specifically tend to their perceptions of what experiences they value in their learning processes.

Institution and target population

Our institution is a large research-intensive university in the Midwest of the United States serving ~32,000 undergraduate and ~16,000 graduate students. Our department offers ~11 introductory biology lecture and laboratory courses every semester. Our target population is composed of all laboratory and lecture TAs in this department. Approximately 75 to 100 TAs are involved in our departmental teaching duties each semester. Both undergraduate and graduate students serve as TAs. The vast majority of our TAs are responsible for teaching and grading one or two sections of a multisection lab course. This study was approved by our institution's IRB (#STUDY00000941).

PART I. TESTING THE ACTIVITY VS. EVIDENCE HYPOTHESIS

Methods

In order to test the hypothesis that TAs will value modeling activities in a low-stakes environment over being exposed to evidence of the activities' effectiveness, we

experimentally manipulated how activities were presented to TAs in a presemester workshop. Here, we define a low-stakes environment as one in which a mistake made while using a teaching technique is made in front of peers and facilitators, not in front of students, so that TAs can gain experience and feedback without losing face in front of students. Based on feedback at the end of the fall 2017 semester, we invited TAs to participate in an optional workshop on AL teaching techniques in January 2018 (hereafter, the "January Workshops"). We offered two workshop sessions that differed in how material was presented to the TA participants. All TAs teaching introductory biology labs or serving as course assistants in lecture courses were invited, via e-mail in December 2017, to participate in a January workshop. When TAs signed up for a workshop, they were allowed to choose either the morning or the afternoon workshop session or they could indicate whether their availability was flexible. The workshop organizers placed flexible TAs into morning or afternoon sections to distribute TAs as evenly as possible with regard to course, experience (new vs. returning), gender, and level (undergraduate TA vs. graduate TA). Demographic information for the TAs in each session is provided in Table I. Attendees in each session were further

TABLE I.

Demographic information, prior teaching experience, and prior training of the teaching assistants participating in the January Workshops as well as course types taught and survey response numbers.

		A Group	E Group
Participants	RSVP	19	19
	Attended	19	19
Gender	Female	13	13
	Male	6	6
Academic level	Undergraduate	16	12
	Graduate	3	7
Teaching experience	New	4	3
	Returning	15	16
Previous training	Yes	6	9
	No	13	10
Course type	Lab (lead TA)	11	15
	Lab (helper)	5	4
	Lecture	3	0
Survey responses	Total: Presemester respondents	19	19
	Total: Postworkshop respondents	19	19
	Total: Postsemester respondents	8	12
	Matched: Presemester to postworkshop only	11	7
	Matched: All time points	8	12

Previous training indicates TAs who participated in previous workshops provided by the department. Lead TAs are responsible for leading one or two sections of a lab course on their own. Helper TAs help students in a lab section but are not responsible for leading the lab. Lecture TAs serve as course assistants for one or more of the lecture courses in our department. Total survey responses indicates the total number of surveys completed at each time point and provides sample sizes for Figures 1 to 3. Matched survey responses indicate the number and time points of surveys completed by the same respondents. TA = teaching assistant.

subdivided into four teaching teams, each consisting of four or five TAs, again attempting to distribute TAs as evenly as possible in terms of course, experience, gender, and level. A total of 38 TAs attended the workshop, 19 in each session. All four authors facilitated both workshops.

The workshop activities are outlined in Table 2. The first hour of both workshops covered the same material in the same manner. The facilitators introduced themselves, discussed the objectives of the workshop, and sorted TAs into their teaching teams. This introduction was followed by an ice-breaker in their teaching teams. The TAs then participated in a figure jigsaw activity that introduced them to the evidence for the overall benefits of active learning (3, 28–30). Table 3 summarizes the focal activities covered in the workshops.

Activity (A) Group

The A Group met for three hours in the morning. After completing the initial activities outlined in the previous section, each Teaching Team was assigned their focal activity and facilitator, and—with the help of their facilitator—learned about the activity and best practices for facilitating the activity. They then practiced how they would model this activity to their peers (the other Teaching Teams). During the next hour, each Teaching Team modeled their focal activity as instructors to their peers, who acted as students in a mock classroom setting. The session closed with, first,

a brief discussion of how the TAs might incorporate these activities into their own classrooms and, finally, a minute paper summarizing what they learned and any areas of confusion or concern. The TAs then left and did not overlap with the TAs in the afternoon session.

Evidence (E) Group

The E Group met for three hours in the afternoon. TAs participated in the initial activities as outlined above. After the jigsaw activity, each of the remaining focal activities was presented to the TAs. The Teaching Teams acted as groups of students participating in the activity and the facilitator acted as the instructor in a mock classroom setting. Each facilitator also presented evidence for the effectiveness of their respective technique and provided facilitation tips. For example, in a discussion of the jigsaw activity, students were introduced to the findings of Doymus *et al.* (31), which demonstrated that students learned material significantly better during a jigsaw activity compared with a control group. The session ended like the previous one, with a brief discussion on how TAs might implement the techniques in their own labs, followed by a final minute paper.

Data collection and analysis

Surveys were emailed to TAs one week prior to the workshops, immediately after the workshops concluded, and

TABLE 2.
Summary of the January 2018 active learning workshop schedule.

Same Introductory Activities			
5 minutes	Facilitator introductions Workshop objectives: • Discuss and experience multiple active learning strategies • Learn how to implement them in your lab sections • Discuss strategies for diversity and inclusion Break into Teaching Teams		
25 minutes	Ice-breaker: pasta tower challenge (https://tinkerlab.com/spaghetti-tower-marshmallow-challenge/)		
30 minutes	Figure jigsaw introducing evidence for benefits of active learning in Teaching Teams (3 [Fig. 2], 28 [Fig. 1], 29 [Figs. 1 and 3], 30 [Fig. 1])		
Activity (A) Group		Evidence (E) Group	
45 minutes	Break Out Groups: each Teaching Team learns one of the following activities assisted by a facilitator: • Easy assessment techniques • Games and simulations • Case studies • Sequence strips	105 minutes	Each facilitator presents the evidence for the effectiveness of the following activities then models the activity with TAs acting as students in their Teaching Teams: • Easy assessment techniques • Games and simulations • Case studies • Sequence strips
60 minutes	Each Teaching Team models their focal activity while their peers act as students		
15 minutes	Wrap up	15 minutes	Wrap up

The latter two hours of the workshops differed between the Activity and Evidence groups, with the Activity group modeling activities as instructors and the Evidence group being exposed to the evidence and experiencing the activities as students.

TA = teaching assistant.

TABLE 3.
Focal activities covered in the January Workshop.

Activity	TA familiarity	References	Description
Jigsaw	Familiar to most TAs	31, 40	Each team receives a different piece of information (in this case, a figure from a journal article). The members of each Teaching Team become “experts” of their piece of information. New teams are then formed, composed of one member from each of the original teams. Participants then teach each other their piece of information.
Easy assessment techniques	Unfamiliar to most TAs	40–43	Student assessment techniques that emphasize broad participation and require little class time or preparation, such as think-pair-share, throat-vote, multiple-hands/multiple-voices, polling methods, minute papers, etc.
Games and simulations	Familiar to some TAs	40, 44–46	These are activities in which biological concepts are illustrated to students through a competitive game, model building, or other similar activity. We demonstrated the “Tournament of Kitchen Utensils” game illustrating natural selection, adapted from https://www.biologycorner.com/worksheets/naturalselection.html .
Case studies	Familiar to some TAs	40, 47	Biological concepts are presented to students as “real-life” scenarios accompanied by questions and problems that students answer using the content knowledge they have acquired from the course. We demonstrated this technique with a case study from the National Center for Case Study Teaching in Science website (http://sciencecases.lib.buffalo.edu/cs/).
Sequence strips	Unfamiliar to all TAs	40, 48	Students are given strips of paper, each containing one step of a complex process (such as meiosis) or procedure (such as loading a DNA visualization gel) and must sort the strips into the correct order.

at the end of the semester. The TAs had 1 or 2 weeks to complete the surveys and received reminder e-mails twice. Those who participated in these optional January Workshops were incentivized for their participation, including pre- and postworkshop surveys, with \$75 Amazon gift cards. Postsemester survey responses were incentivized with \$5 Starbucks gift cards. Specifically, surveys included items asking TAs to self-report their knowledge and perceptions of active learning, inclusive teaching, and facilitating inquiry, as well as how often they implemented these strategies in their own lab sections. Survey items involved both Likert-type responses and open-ended prompts (e.g., Of the training topics you selected as valuable above, which do you think was the MOST valuable? Why?). These questions were discussed and written by several biology education researchers in collaboration with a former TA who had taught within our department, but they were not otherwise validated. The survey questions can be found in Appendix 1.

Our findings were interpreted using the mixed-methods approach of concurrent triangulation (41); specifically, we used survey metrics to gather both quantitative and qualitative data. We report on both sets of data: the quantitative data have value, but, given our small sample size, the qualitative data better contextualize the TA experience.

Quantitative survey responses were downloaded from the Qualtrics survey platform to Excel, then analyzed in R (42) and graphed using the ggplot2 package (43). We used the non-parametric Mann–Whitney–Wilcoxon and Kruskal–Wallis tests to determine whether there were significant differences between groups.

We interpreted TA responses to the open-ended questions using first- and second-cycle qualitative analysis (44).

In the first cycle, open codes were identified and assigned to TA responses, then categorized by two research assistants. In the second cycle, categories were organized into main emerging themes by comparing their similarities and differences.

We triangulated between the two types of data via a series of small-group discussions involving biology education researchers, postdocs, graduate students, and undergraduate students in the lab group. Specifically, we discussed how the emerging themes from the qualitative data can inform any trends detected in the quantitative data. This group-level dialogue informs our discussion, below.

Results

All TAs completed the presemester to postworkshop surveys; fewer TAs responded to the postsemester survey. The number of responses to each survey (sample sizes) is given in Table 1. There was no significant difference in TA perceptions of their knowledge of active learning (AL) between groups prior to the January Workshops (Fig. 1; $p = 0.26$). The TAs reported significantly greater knowledge of AL after the workshop (A group: $p = 0.01$; E group: $p = 0.002$). There was also a significant difference between groups after the workshop, with Evidence group participants reporting significantly greater perceptions of their knowledge of AL than the Activity group participants (Fig. 1; $p = 0.04$). However, while TA perceptions of their knowledge of AL remained greater at the end of the semester than prior to the workshop, the difference between the Activity and Evidence groups disappeared by the end of the semester (Fig. 1; $p = 0.77$).

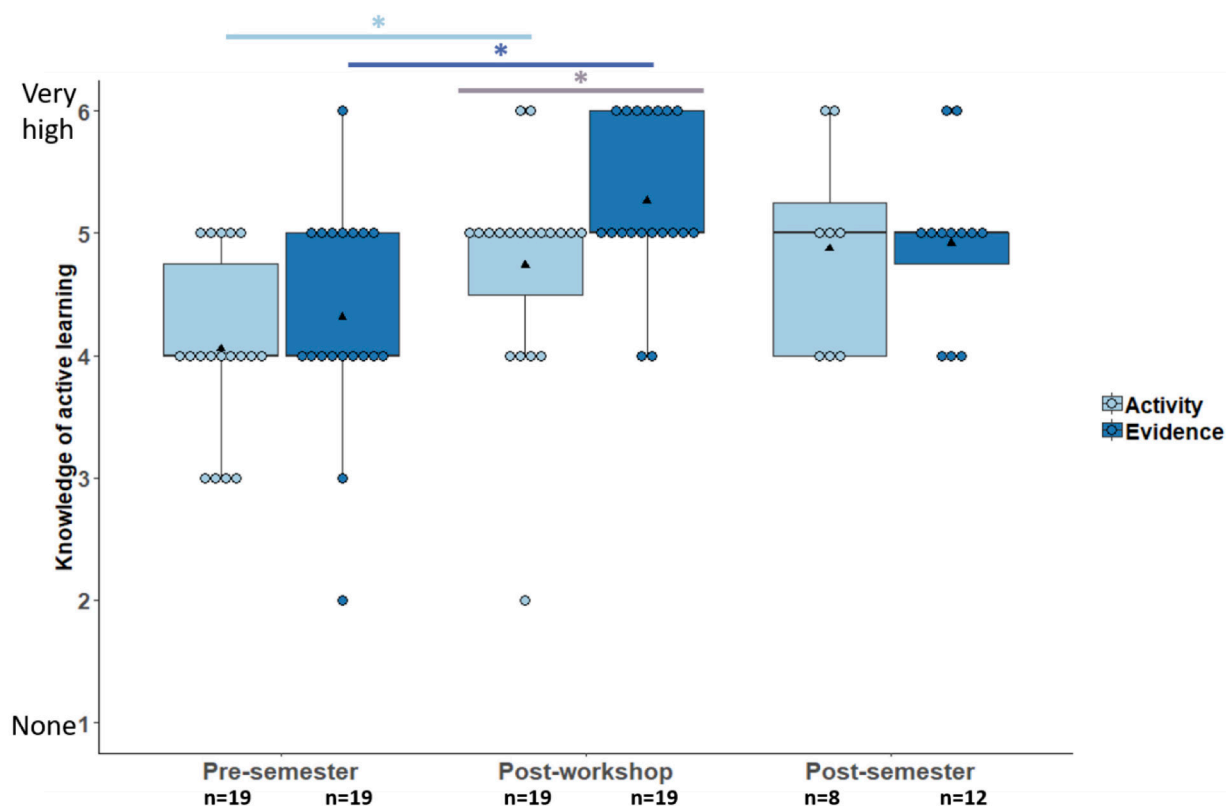


FIGURE 1: TA responses to the question “Please indicate your knowledge of active learning.” Responses ranged from 1 (None) to 6 (Very High). Activity = actively modeled teaching techniques as instructors but were not presented with evidence; Evidence = experienced teaching techniques as students and were presented with evidence. The dark horizontal bar indicates the median, the shaded box indicates the interquartile range, and the whiskers depict 1.5 times the interquartile range. The individual circles indicate individual respondents. The black triangles indicate the mean. * = $p \leq 0.05$.

The Activity group participants found the workshop activities more valuable ($p = 0.053$) after the workshop than did the Evidence group participants (Fig. 2). This difference between groups increased during the semester, with the Activity group finding the activities significantly more valuable than the Evidence group ($p < 0.05$; Fig. 2). This suggests that getting to model an activity to their peers is more valuable than learning the evidentiary basis for the teaching practices and experiencing the practices as students. This finding was the basis for the design of the August workshops (discussed below, in “Part II”).

Easy assessment, which requires little to no materials or prior planning, was the workshop technique both groups of TAs reported using most often, an average of *several times* during the semester (Fig. 3). The remaining techniques were reportedly used *never* to *a few times* during the semester (Fig. 3). Similar trends were observed when TAs were asked to report their most and least valuable workshop topic at the end of the semester: easy assessment techniques were reported the most valuable for the majority of participants from both groups and were never reported as the least valuable. A larger proportion of Activity group participants reported the figure jigsaw—the only evidence for the overall effectiveness of AL that they were exposed to—as the most valuable workshop activity, compared with the

Evidence group. Teaching techniques that required more materials and advanced planning, particularly games and simulations, were the least valuable workshop activities reported for both groups. On postsemester surveys, TAs voiced an interest in future workshops focusing on inclusive teaching (e.g., avoiding implicit bias, involving all students in lab activities) and facilitating inquiry (e.g., guiding students, rather than providing immediate answers; using failure as a teaching tool).

Summary

Both qualitative and quantitative data suggest that it is important for TAs to engage in the activities that are being taught. There were no compelling indications that being presented with the evidence for evidence-based teaching (EBT) is critical—either in terms of knowledge or practice of EBT.

PART II. TESTING THE “CHERRY ON TOP” HYPOTHESIS

Methods

Based on the results from the initial workshop (detailed and summarized in the Results, above), we wanted to fur-

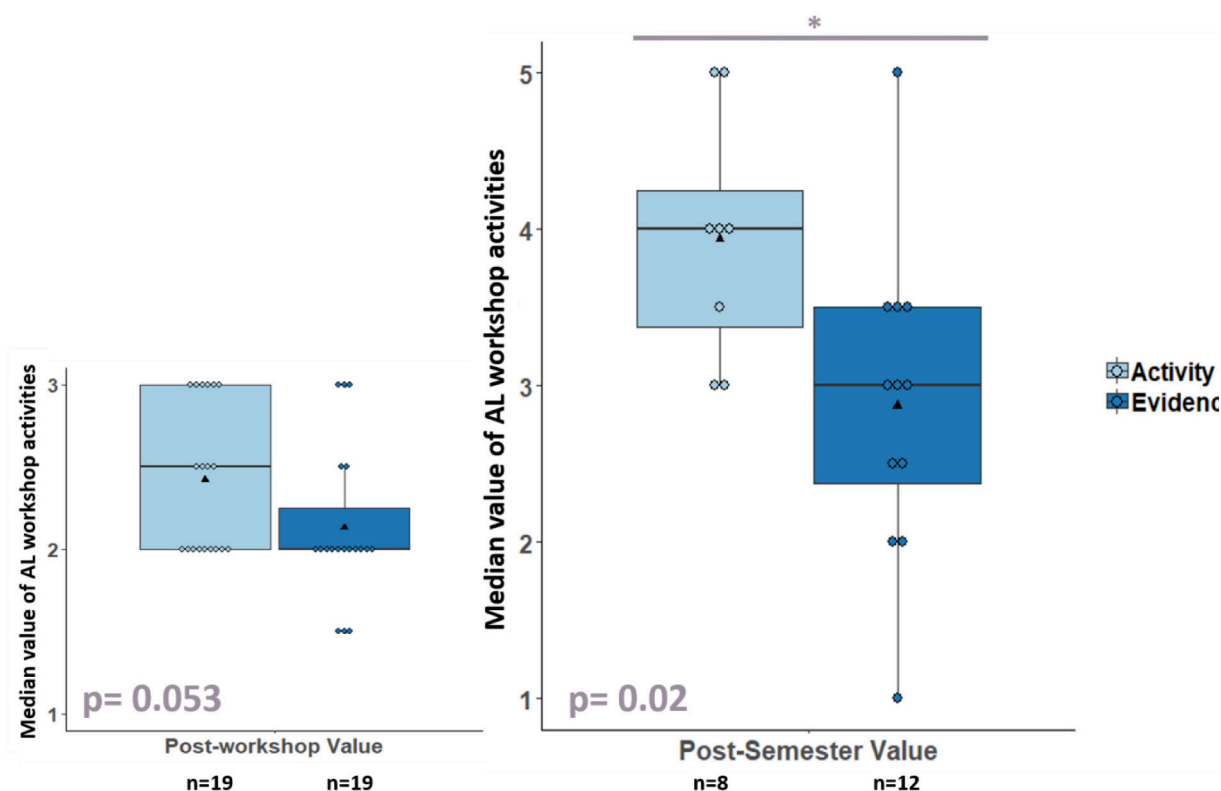


FIGURE 2. Reported value of workshop activities in the first workshop. We took the median of each participant's responses to how much they valued each workshop component to create a metric of the overall value of the workshop for each TA. The postworkshop responses were on a scale of 1 (not at all valuable) to 3 (extremely valuable), whereas the postsemester responses were on a scale of 1 (not at all valuable) to 5 (extremely valuable). Activity = actively modeled teaching techniques as instructors but were not presented with evidence; Evidence = experienced teaching techniques as students and were presented with evidence. The dark horizontal bar indicates the median, the shaded box indicates the interquartile range, and the whiskers depict 1.5 times the interquartile range. The individual circles indicate individual respondents. The black triangles indicate the mean. * = $p \leq 0.05$.

ther disentangle the relative importance of evidence and modeling in TA professional development. In response to feedback from participating TAs in the spring 2018 semester, we focused the experimental workshop topics on inclusive teaching and facilitating inquiry. These half-day, experimental workshops were held in August 2018 (hereafter, the "August Workshops"). Due to the number of TAs attending and the difficulty of securing lab rooms with enough seating, we divided TAs among three rooms, ~20 TAs per room, with three facilitators in each room. Demographic information for the TAs in each group is summarized in Table 4. Teaching assistants were further divided into Teaching Teams of five to seven TAs, for a total of three Teaching Teams per room; each Teaching Team had its own facilitator. Authors LEP, HAB, and SC were the lead facilitators in each room and were assisted by two additional facilitators, in each room, who were knowledgeable about ST. The workshop activities are summarized in Table 5.

The treatments differed in their exposure to introductory material for each of the sub-sessions.

- No Evidence Group: TAs in this group modeled the focal techniques in both sessions and were *not* exposed

to evidence for the effectiveness of the techniques in either session.

- Evidence Group: TAs in this group also modeled the focal techniques and were exposed to evidence demonstrating the effectiveness of the techniques in both the Inclusive Teaching and Facilitating Inquiry sessions.
- A–E Group: TAs in this group got the No Evidence group materials for the Inclusive Teaching session and Evidence group materials for the Facilitating Inquiry session.

In the remainder of the workshop, *all* participants actively modeled the same focal techniques, which were covered in the latter portion of the sessions. In the Inclusive Teaching session, we spent time with the TAs co-constructing group norms for the session such that we could collectively uphold a supportive and inclusive learning space. The group norms activity was also described to the TAs as a mechanism they could use with their students to create a positive lab environment. Then, TAs worked in their Teaching Teams with their facilitator to develop a strategy to handle an inclusive teaching scenario. Each Teaching Team then role-played their solution for the other groups

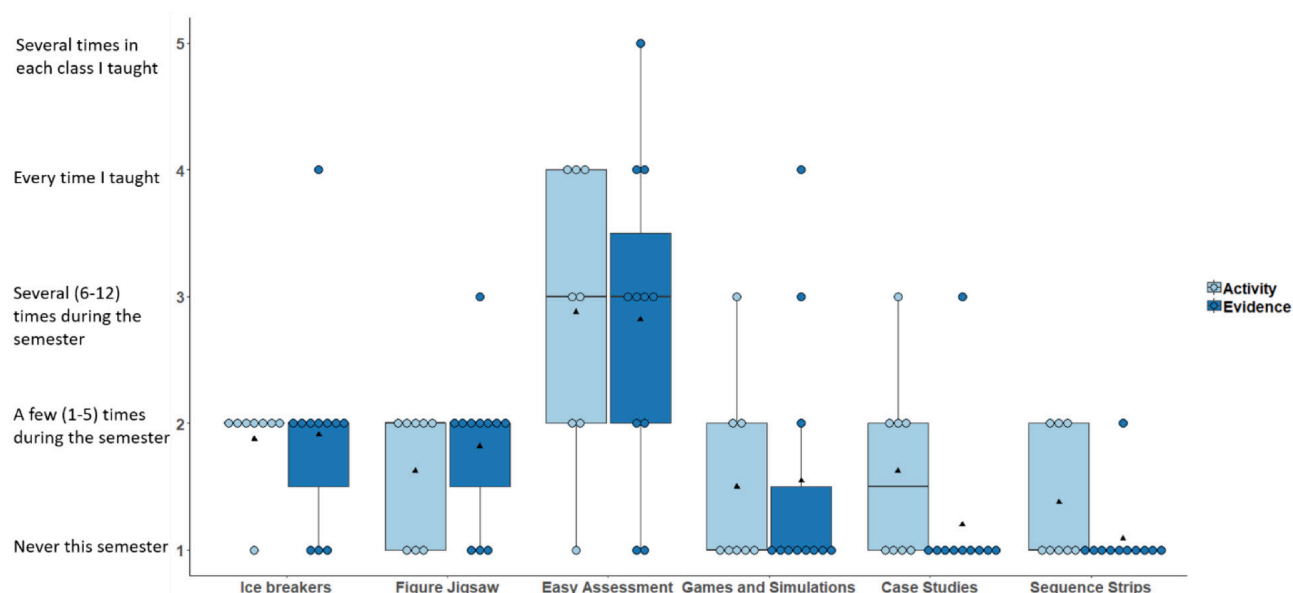


FIGURE 3. Self-reported frequency of use of workshop activities learned in the first workshop during the semester: Activity = actively modeled teaching techniques as instructors but were not presented with evidence, $n=8$; Evidence = experienced teaching techniques as students and were presented with evidence, $n=12$. The dark horizontal bar indicates the median, the shaded box indicates the interquartile range, and the whiskers depict 1.5 times the interquartile range. The individual circles indicate individual respondents. The black triangles indicate the mean.

or discussed it with them and facilitated a discussion of other strategies that could also have been effective. In the Facilitating Inquiry session, each Teaching Team learned a different strategy to facilitate inquiry with their facilitator; Teaching Teams then modeled these strategies with their peers acting as students in a mock lab setting. All sessions wrapped up with a minute paper in which the TAs wrote about what they found effective and what they wanted to learn more about. (All workshop materials will be included in a forthcoming publication.)

Data collection and analysis

Data collection and analysis largely mirrors that of Part I (discussed above); however, in Part II, TAs were incentivized to complete the surveys with \$5 to \$10 Starbucks gift cards. Due to low survey response rates (summarized in Table 4), we combined A–E response data for the Inclusive Teaching session with the No Evidence group and response data for the Facilitating Inquiry session with the Evidence group.

In addition to the Likert-scale questions, the open-ended question “How important is it for us to present evidence for the effectiveness of scientific teaching practices (including inclusive teaching, active learning, and assessment) during the workshop? Explain your answer” was included in the postworkshop and postsemester survey. We interpreted TA responses to the open-ended questions using first- and second-cycle qualitative analysis (44). In the first cycle, two research assistants developed open codes that reflected the main ideas of the TA responses. Open coding allowed the research assistants to compare and contrast the TA responses, and thus categorize the responses accordingly.

Second-cycle qualitative analysis allowed us to look deeply at the content of the categories, compare similarities and differences, and refine them into the emergent categories we reported in Table 6.

Triangulation discussions, similar to those discussed above, allowed us to better interpret the survey data; specifically, qualitative findings added nuance to the quantitative results, as discussed below.

Results

Survey response rates were very low (Table 4); \$5 to \$10 coffee cards were not large enough incentives to encourage completion. Very few TAs completed all surveys, and only slightly more completed two surveys, meaning that few of the survey responses could be matched. Very few TAs had attended the January Workshops. Because of these small sample sizes, we do not feel comfortable reporting that there are statistical differences between groups or time points for the second experiment. However, we feel that the trends are clear—if a bit noisy—so these are what we report below.

There were no differences in TA perceptions of their knowledge of Inclusive Teaching and Facilitating Inquiry between the No evidence and Evidence groups prior to the August Workshops in fall 2018 (Fig. 4, presemester). After attending the workshop, TAs in both groups reported that their perceptions of their knowledge of Inclusive Teaching and Facilitating Inquiry increased substantially, and these remained higher at the end of the semester than prior to the start of the semester (Fig. 4, postworkshop and postsemester). Both groups also found the workshop topics valuable, with no discernable differences between groups (Fig. 5).

TABLE 4.

Demographic information, prior teaching experience, and prior training, of the teaching assistants participating in the August Workshops as well as course types taught and survey response numbers.

		A Group	A/E Group	E Group
Participants	RSVP	24	22	22
	Attended	20	16	17
Gender	F	12	12	11
	M	8	4	6
Academic level	Undergraduate	16	12	12
	Graduate	4	4	5
Experience	New	15	11	13
	Returning	5	5	4
Previous training	Yes	4	6	7
	No	16	10	10
Attended January workshop	No	17	15	14
	A	1	0	2
	E	2	1	1
Course type	Lab (lead TA)	14	15	14
	Lab (helper)	5	0	2
	Lecture	1	1	1
Survey responses	Total: Presemester respondents	10	8	12
	Total: Postworkshop respondents	6	7	12
	Total: Postsemester respondents	5	5	11
	Matched: All time points	3	2	7
	Matched: Presemester to postworkshop	2	2	3
	Matched: Presemester to postsemester	0	1	2
	Matched: Postworkshop to postsemester	0	2	1
	Unmatched: One time point only	8	4	2

Previous training indicates TAs who participated in previous workshops provided by the department. Attended January workshop indicates which January workshop section, if any, TAs attended. Lead TAs are responsible for leading one or two sections of a lab course on their own. Helper TAs help students in a lab section but are not responsible for leading the lab. Lecture TAs serve as course assistants for one or more of the lecture courses in our department. Total survey responses indicates the total number of surveys completed at each time point and provide sample sizes for figures 4 to 8. Matched survey responses indicate the number and time points of surveys completed by the same respondents. Unmatched indicates the number of survey respondents who only completed a single survey.

TA = teaching assistant.

Teaching assistants reported using the Inclusive Teaching techniques *several times* during the semester, on average, and there did not appear to be substantial differences in frequency of use between the No evidence and Evidence groups (Fig. 6A). Those in the No evidence group reported using the Facilitating Inquiry techniques *a few times* on average, while the TAs in the Evidence group reported using the Facilitating Inquiry techniques more frequently, *several times* during the semester on average (Fig. 6B). There was little consensus on which activities were most and least useful between groups and time points; this lack of agreement leads us to conclude that each of the workshop activities we offered was useful to at least *some* of the participating TAs.

We also asked TAs how important evidence was in convincing them to use ST practices. On average, TAs reported that being presented with evidence of the effectiveness of various teaching practices was very important postworkshop and postsemester (Fig. 7). The overwhelming majority of TAs found the evidence *moderately*, *very*, or *extremely important*. Responses from the No evidence and Evidence groups were not discernably different from each other (Fig. 7). Next, TAs could explain their choices—from “not at all important” to “extremely important.” These open responses were then coded, and two coders came to consensus on five categories for the codes. Of these categories, four grouped naturally into the “values the evidence”

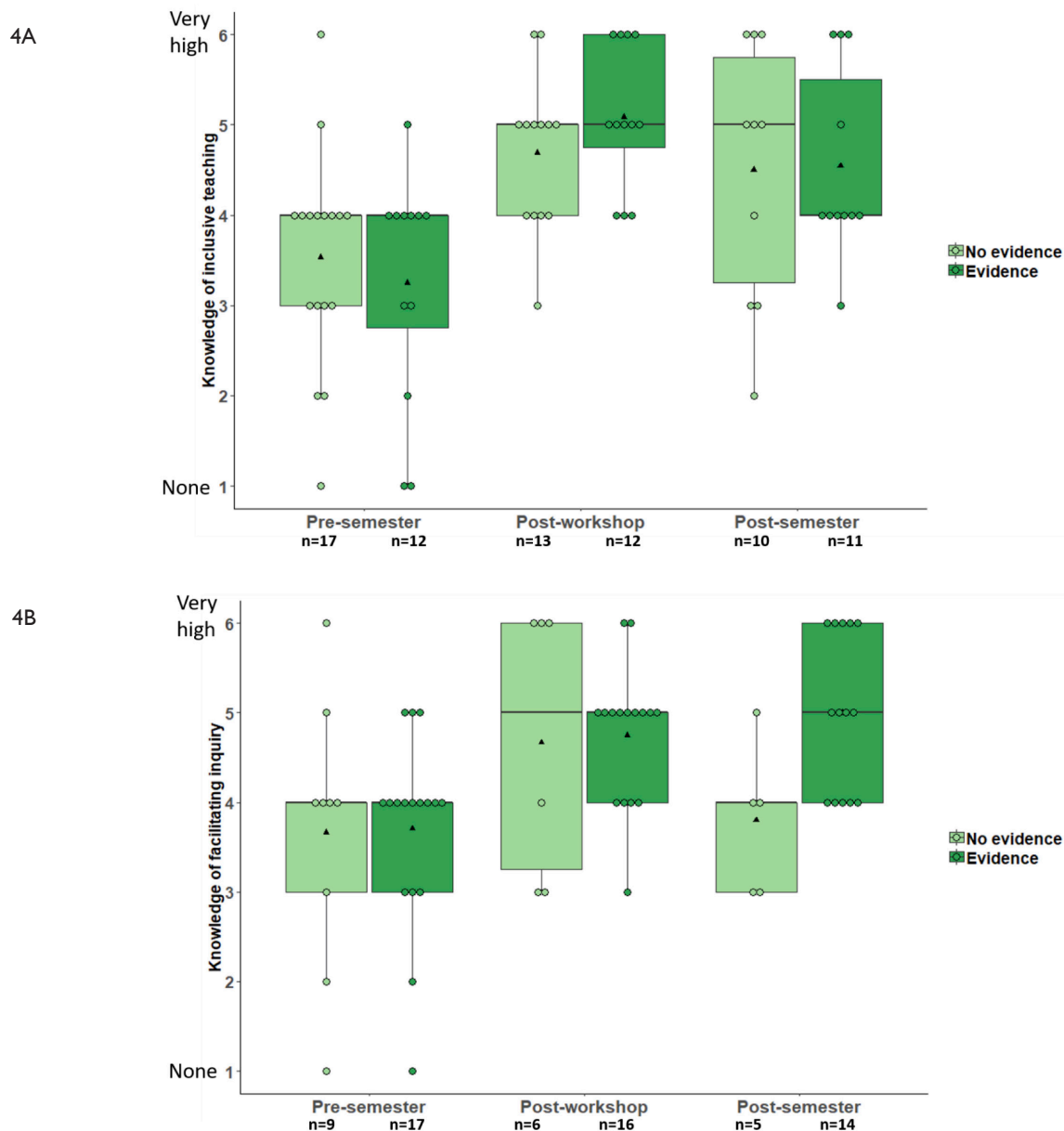


FIGURE 4. TA knowledge of inclusive teaching (A) and facilitating inquiry (B). Responses ranged from 1=None to 6=Very High. The dark horizontal bar indicates the median, the shaded box indicates the interquartile range, and the whiskers depict 1.5 times the interquartile range. The individual circles indicate individual respondents. The black triangles indicate the mean.

emergent theme, and one, “not useful,” clearly belonged to a separate, “doesn’t value the evidence,” theme. The five categories are described below.

Helps with implementation. Codes in this category made a connection between the evidence they were presented with and the practical application of the teaching techniques under investigation. For example: “Because in many fields, there are always approaches that are great in theory and work in ideal conditions, but in practice, they are difficult to implement. I believe teaching is similar.” Ten TA responses included this code.

Improves understanding. These codes specifically addressed the *how* or *why* of different teaching techniques

and stated that the evidence helped them understand the pedagogy. For example: “Because students don’t always like these methods but if we have a concrete understanding of why it works then we can explain it to them too.” Twenty-three TA responses included this code.

Increases motivation. These codes attributed some of their motivation for implementing ST to its evidentiary basis. For example: “It is motivating to see that what we will be doing will have a strong impact on each student’s learning.” Ten TA responses included this code.

Cherry on top. These codes referred to learning about the evidence as a nice, but not necessary, addition to the workshop. For example: “I think it drives the point

TABLE 5.
Summary of the August 2018 Inclusive Teaching and Facilitating Inquiry workshop activities.

Inclusive Teaching			
	No Evidence	A–E	Evidence
20 minutes	<ul style="list-style-type: none">• Introductions• Objectives: Develop and apply strategies for creating inclusive classrooms• Session norms• Introduction to inclusive teaching and hidden identities	<ul style="list-style-type: none">• Introductions• Objectives: Develop and apply strategies for creating inclusive classrooms• Session norms• Introduction to inclusive teaching and hidden identities	<ul style="list-style-type: none">• Introductions• Objectives: Develop and apply strategies for creating inclusive classrooms• Session norms• Evidence for why inclusive teaching is important• Introduction to inclusive teaching and hidden identities
30 minutes	Break-Out Groups: each Teaching Team brainstorms a solution to a different scenario assisted by a facilitator		
30 minutes	Each Teaching Team role-plays their solution to a given scenario followed by discussion of alternative strategies		
5 minutes	Wrap up		
10 minutes	Break		
Facilitating Inquiry			
	No Evidence	A–E	Evidence
20 minutes	<ul style="list-style-type: none">• Objectives:<ul style="list-style-type: none">– Learn & model strategies for facilitating inquiry– Develop a toolkit of teaching strategies for facilitating inquiry• Introduction to inquiry• Examples of inquiry labs taught by TAs	<ul style="list-style-type: none">• Objectives:<ul style="list-style-type: none">– Learn & model strategies for facilitating inquiry– Develop a toolkit of teaching strategies for facilitating inquiry• Introduction to inquiry• Examples of inquiry labs taught by TAs• Evidence showing the effectiveness of inquiry-based teaching and learning	<ul style="list-style-type: none">• Objectives:<ul style="list-style-type: none">– Learn & model strategies for facilitating inquiry– Develop a toolkit of teaching strategies for facilitating inquiry• Introduction to inquiry• Examples of inquiry labs taught by TAs• Evidence showing the effectiveness of inquiry-based teaching and learning
30 minutes	Break-Out Groups: each Teaching Team learns one of the following inquiry strategies assisted by a facilitator: <ul style="list-style-type: none">• Brainstorming techniques• Questioning techniques• Techniques to keep students on track and deal with failure		
30 minutes	Each Teaching Team models their focal activity while their peers act as students		
5 minutes	Wrap up		

All participants actively modeled the workshop activities. Participants in the *No Evidence* Group were not shown evidence for the effectiveness of the focal teaching activities. The *Evidence* Group was presented with evidence. The *A–E* group was exposed to evidence only in the Facilitating Inquiry session

home, but hopefully it's a concept we can happily all get behind." Eight TA responses included this code.

Not useful. These codes clearly communicated that the evidence was not a valuable part of the workshop. For example: "It seems that you are trying to make us believe it is great." Five TA responses included this code.

Additional comments are included in Table 6.

Summary

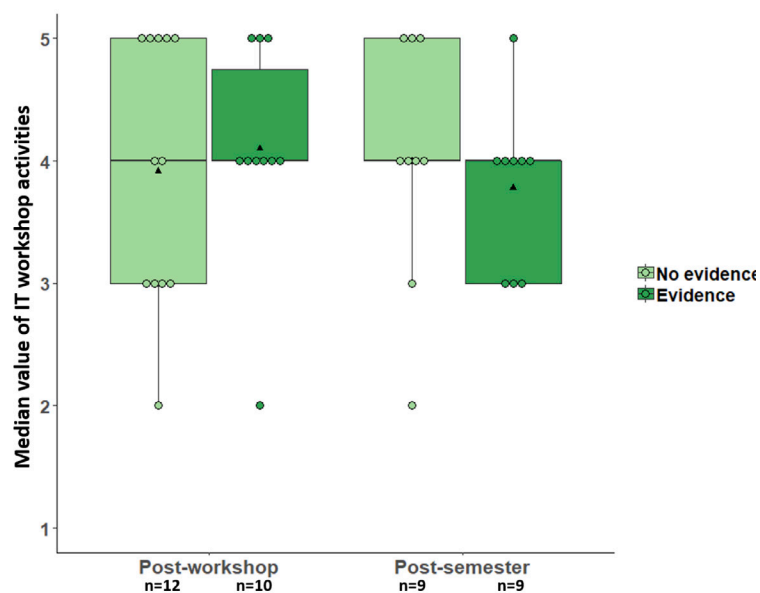
Most TAs value learning about the evidence for EBTP, and cite a variety of reasons for this evaluation.

DISCUSSION

Collectively, our findings indicate that actively mod-

eling teaching practices can be an important component of TAPD workshops, while evidence for the effectiveness of the teaching practices is also valued by the majority of the TAs (Fig. 8). We found that TAs who were able to model ST techniques as instructors (Activity Group) in a low-stakes workshop environment found the workshop topics more valuable and used the techniques more often than TAs who were given evidence for the effectiveness of the same teaching practices but only experienced the techniques as if they were students (Evidence Group). There was no difference between groups when all TAs were able to actively model teaching practices but only some TAs were shown evidence; taken alone, this finding supports our hypothesis that evidence functions as the "cherry-on-top" of TAPD workshops: it's nice to have but not crucial. However, on open-ended responses, most TAs expressed more than just "cherry-on-top" valuation for evidence. They spoke

5A



5B

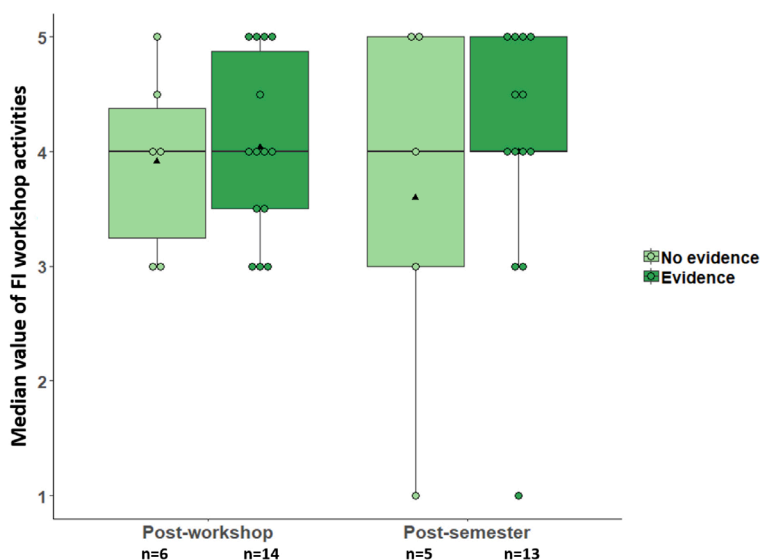


FIGURE 5. Reported value of workshop activities for the inclusive teaching (A) and facilitating inquiry (B) sessions of the second workshop. We took the median of each participant's responses to how much they valued each workshop component to create a metric of the overall value of the workshop for each TA. The responses were on a scale of 1 (not at all valuable) to 5 (extremely valuable). The dark horizontal bar indicates the median, the shaded box indicates the interquartile range, and the whiskers depict 1.5 times the interquartile range. The individual circles indicate individual respondents. The black triangles indicate the mean.

about the value of this evidentiary basis on their students' motivation to learn and understanding of course content, as well as on their own (the TAs') willingness to implement certain AL strategies in lab sections. For that reason, we propose a modified hypothesis for how to view the evidence in structuring TAPD: specifically, evidence functions as the chocolate (or strawberry, or caramel) sauce on the ice-cream sundae—without the sauce, you don't have a sundae, but it's not the main part. In this metaphor, having TAs model the supported activities is the ice cream—the foundation of the sundae. But the evidence is still important.

Implications

We recognize that there is a vast body of research working to understand how professional development influences science educators—people who have chosen teaching careers—in the primary and secondary school settings (for example, see 45). However, there has been less work on how professional development impacts science educators who have not necessarily chosen teaching as a career—namely TAs. Therefore, we reaffirm this work in the postsecondary space as novel and impactful.

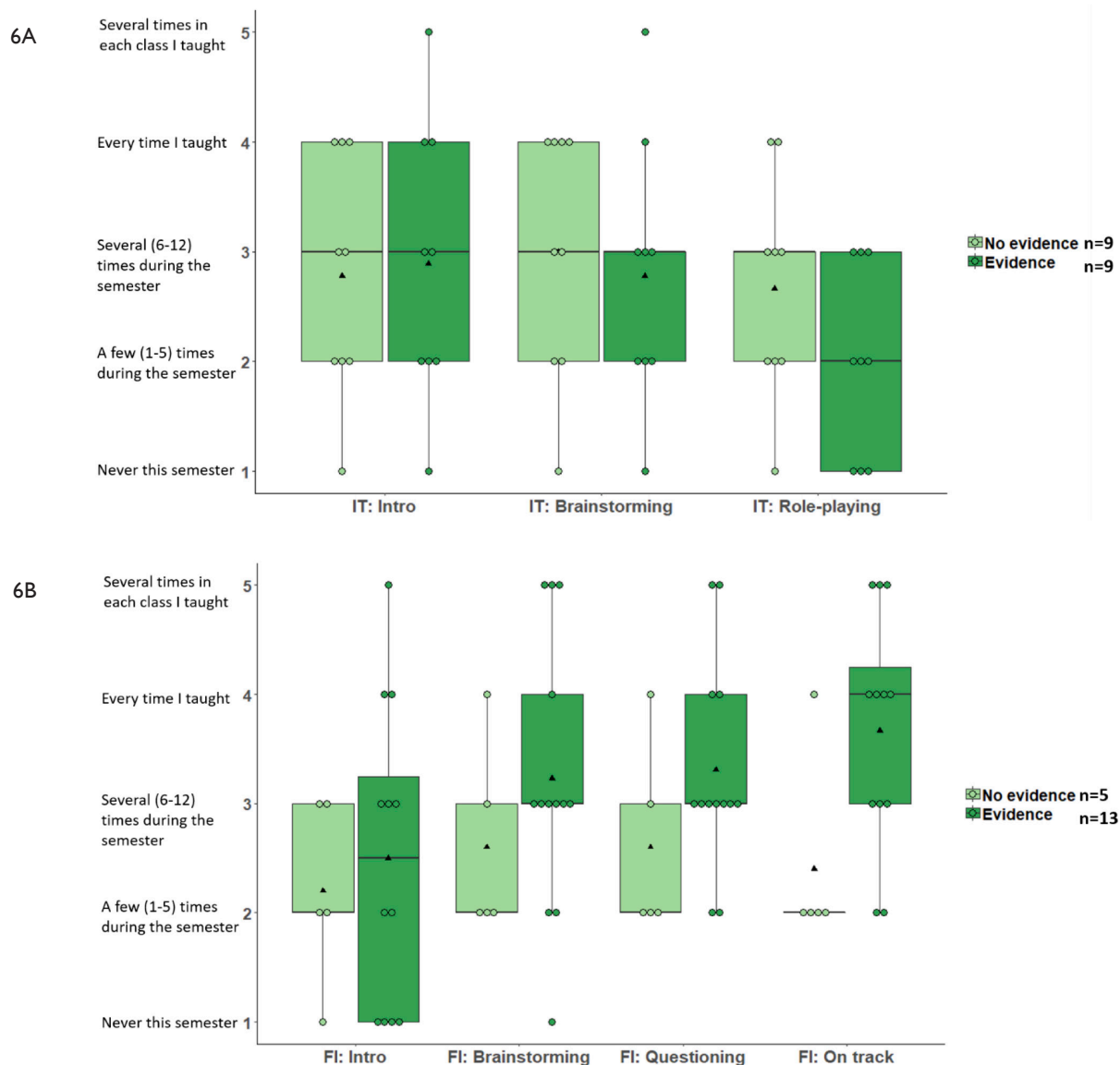


FIGURE 6. Self-reported frequency of use of workshop activities for the inclusive teaching (A) and facilitating inquiry (B) sessions of the second workshop. The dark horizontal bar indicates the median, the shaded box indicates the interquartile range, and the whiskers depict 1.5 times the interquartile range. The individual circles indicate individual respondents. The black triangles indicate the mean

We are aware of only one other study that has experimentally manipulated TAPD workshop conditions. Hughes and Ellefson (12) designed, implemented, and assessed two separate workshops for TAs teaching an inquiry-based introductory biology lab course. They found that TAs who attended workshops about inquiry-based pedagogy were more effective in terms of self-reported and student-based metrics than TAs who attended workshops on general teaching best practices. Our research extends this scant literature base by providing evidence about the perceived value of our TAPD for TAs' instruction, and according to specific experiences.

Some previous work has explored the important components of TAPD workshops. Based on the literature,

Wheeler *et al.* (18, 46) asserted that presenting evidence for EBTP, observing the lab coordinator/facilitator or experienced TA demonstrate teaching practices, and learning theory and pedagogy are important components of TAPD. However, these researchers were unable to demonstrate that their presemester workshop and weekly in-prep meeting TAPD positively influenced TA confidence, self-efficacy, or teaching beliefs (18). Our findings suggest that Wheeler's proposed TAPD components may not be crucial, at least for presemester workshops.

Some TAPD programs involve facilitators modeling teaching practices, with TAs acting as students (e.g., 17), while others have TAs practice modeling as instructors, but only after everyone has experienced the technique as stu-

TABLE 6.
Sample TA responses on the importance of evidence.

Response	No. of Responses (n=32)	Example Comment
Helps with implementation	10	Because I don't have any experience in this type of teaching, and it seems like it takes a lot of time and that you are not actually passing along any information, so to see that these tools work makes it more likely for me to put them into practice.
Improves understanding	23	It helped me to understand how effective these strategies can be in classroom settings
Increases motivation	10	As a student, I am hesitant to fully engage in active learning because I do well in traditional classroom environments, but seeing the evidence helps convince me that active learning is worth being a part of.
Cherry on top	8	Evidence is always good. I didn't select the most extreme option only because I feel that discussing these topics without evidence is still useful, but evidence just makes it that much more convincing!
Not useful	5	I feel I already knew all of this information coming into the workshop.

The survey question was "How important is it for us to present evidence for the effectiveness of scientific teaching practices (including inclusive teaching, active learning, and assessment) during the workshop?" and respondents were asked to explain their answer.

dents (47), so TAs were not learning new teaching practices from their peers, as in the present study. In both of these previous studies, the TAPD occurred during the weekly prep meetings, and in both cases, the effect on TAs was mixed, meaning that not all teaching practices were implemented by TAs at levels deemed acceptable by the researchers. For example, Becker *et al.* (47) found that TA implementation of

the focal teaching practices was highly variable and did not lead to differences in student learning outcomes, which led them to suggest that neither modeling-based nor evidence-based approaches to TAPD were sufficient to convince TAs to use EBTP and that prior experiences as students greatly influence TA adoption of such practices. However, their results suggest to us that not all TAs were comfortable

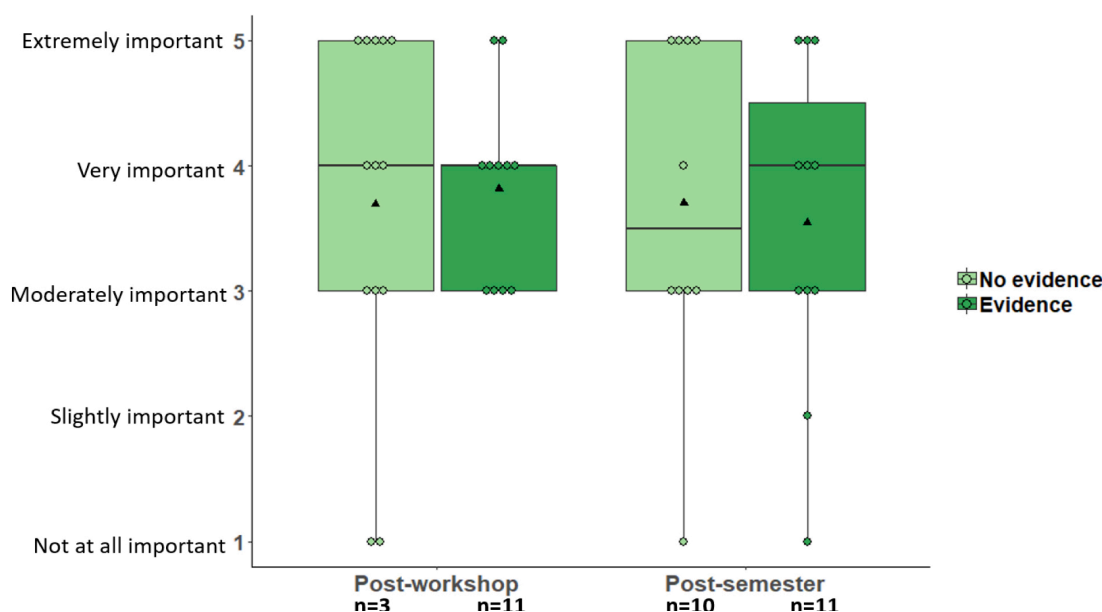


FIGURE 7. TA responses to the question: "How important is it for us to present evidence for the effectiveness of scientific teaching practices (including inclusive teaching, active learning, and assessment) during the workshop?" The dark horizontal bar indicates the median, the shaded box indicates the interquartile range, and the whiskers depict 1.5 times the interquartile range. The individual circles indicate individual respondents. The black triangles indicate the mean.

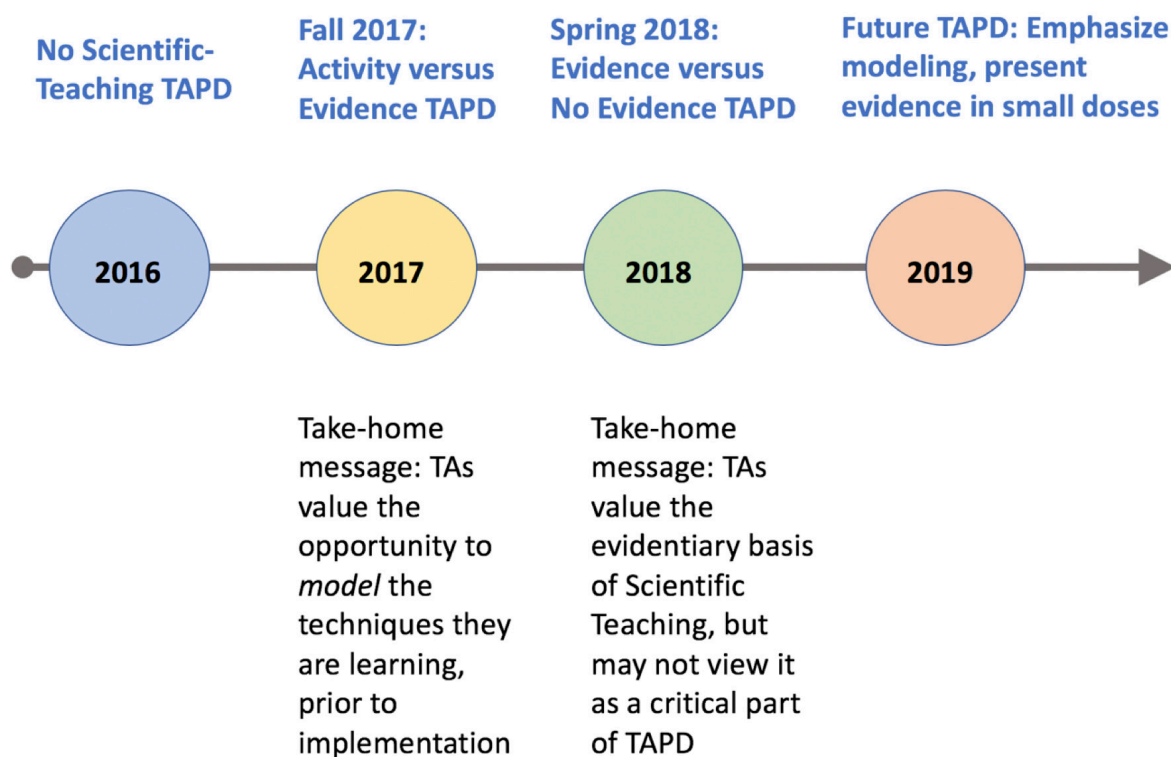


FIGURE 8. A summary of TAPD workshops, and associated take-home messages, in the development of the Building Excellence in Scientific Teaching (BEST) TA training program. TA = teaching assistant; TAPD = TA professional development.

with all focal teaching practices and they shouldn't all be expected to implement them at similar levels. Our findings were similar in that our TAs did not value or report using all techniques at similar levels.

Few workshops have TAs teach their peers EBTP. For example, over the course of four years, Roden *et al.* (16) implemented and assessed a 2-day, presemester ST workshop. Teaching assistants learned a ST practice in small groups, then, using the jigsaw method, shared their focal technique with their peers. They also practiced teaching using a variety of ST practices, received feedback on their teaching, and were able to try again. Teaching assistants reported that they appreciated the opportunity to practice these teaching techniques in a low-stakes environment. Furthermore, similar to our results, TAs reported that their knowledge and use of these teaching practices increased after the workshop.

Taken together, our results and those from the previous studies suggest that key components of presemester TAPD workshops include 1) covering teaching practices that are relevant and useful to the courses for which the TAs are responsible and 2) creating a low-stakes environment for TAs to learn and model teaching practices from their peers. Although, in our study, facilitators did help the TAs learn the group's focal technique/scenario before modeling, most of the TAs were first exposed to the technique from an instructional perspective when their peers demonstrated it. It seems like this distinction shouldn't make a big dif-

ference (i.e., experiencing the technique as students with peers leading vs. with a facilitator leading) but we did see a difference, with Activity group TAs valuing the workshop more than their Evidence group counterparts.

Recommendations

Our future workshops will continue to emphasize modeling of ST—in the low-stakes, non-threatening environment of the TAPD workshops. Some evidence for the effectiveness of the teaching techniques will be presented, but it will not be the focus of initial TAPD. Rather, the evidentiary basis will form more of the discussion for follow-up workshops for experienced TAs who have expressed a deeper interest in pedagogy. This plan also has its roots in other work (Barron *et al.*, in revision) suggesting that novice TAs are not immediately concerned about pedagogy, but that, rather, this interest develops, in *some* TAs, with teaching experience and opportunities for reflection.

Our recommendations for TAPD, based on the experiences detailed above, are the following:

Workshop participants should actively model teaching strategies. The primary differences we observed were between the Activity group and the Evidence group in the spring 2018 presemester workshop, with Activity group participants valuing the workshop more than their Evidence group counterparts. Thus, workshop facilitators should emphasize opportunities for TAs to model what

they are learning, prior to implementation in their own lab sections. We also found that TAs appreciated knowing the evidentiary basis for the techniques being modeled but didn't need the evidence to overshadow the practical aspects of the workshops.

Select teaching practices that are fast, require minimal preparation, and are useful in any class.

Of the techniques presented, discussed, and modeled, TAs reported easy assessment techniques to be the most valuable and the most used. Further, because TA duties can vary by course, instructor, and student population, giving TAs tools that are applicable in multiple contexts (e.g., think-pair-share instead of case studies) is likely to be most useful.

Ask TAs what they need. Our TAPD is a work in progress, largely informed by TA input. In some instances, the TAPD we initially developed was not ideally suited to the participants. Multi-pronged assessment, via surveys, focus groups, and in-lab observations (not all of which are detailed here), has provided several avenues of communication between facilitators and TAs. In response to TA feedback, we've instituted more time for modeling, scaled back on presenting the evidentiary basis for ST, and provided options for continuous, postworkshop, TAPD.

Limitations

Our study does have several limitations worth mentioning. First, our survey response rates—and therefore sample sizes—were low, particularly for the second workshop held in fall 2018, so our results should be interpreted with caution. While we could require TAs to attend the second workshop, we could not require them to complete the surveys, and \$5 coffee cards were not large enough incentives to encourage participation. We had much greater survey completion success when we increased the incentive to \$10 to \$15 and when we were able to incentivize participation in the optional workshop we held in spring 2018. We also acknowledge the limitations of self-reported, survey-based data; a clearer picture could be painted by combining observational data, student reporting, and TA self-reflections. While funding constraints may be prohibitive to increasing financial incentives, in future studies we will consider additional motivations to increase participation. For example, TAs who participate in continuous professional development could list the PD as a “Fellowship in Scientific Teaching” on their resumes or curriculum vitae. Marketing training programs in this way could increase participation and/or response rates.

Second, many of our TAs are undergraduate students, and TAPD for undergraduate TAs has not been, to our knowledge, explored; thus, we were not working with an established framework for TAPD in this population. In addition, undergraduate students at our institution have already been exposed to and participated in most of the teaching practices covered in our workshops in the classes they took

in our department. Therefore, these TAs may have already “bought in” to the effectiveness of ST practices and only needed to practice in a low-stakes environment before implementing these practices in the classes they teach. Recent work has shown that many biology graduate TAs have been exposed to and have bought into ST principles (48), so this may be less of a concern than it would have been only a few years ago. Notwithstanding these recent findings, TAs at institutions where ST is less prevalent may need to be convinced that these teaching practices are useful and may therefore find evidence much more important.

Third, in an effort to make these professional development opportunities available to as many TAs as possible, we recruited TAs who served as lecture course assistants or merely assisted other TAs in the lab; these other TAs were not in charge of their own lab sections. However, we developed much of the workshop material for TAs who were in charge of their own lab sections. Thus, there could have been some mismatch between workshop content and TA responsibilities, resulting in a few TAs finding some of the discussion inapplicable.

Future directions

Future research should focus on the scalability and transferability of the workshops: will similar TAPD workshops work in other STEM fields, for example with chemistry or physics TAs? Will math TAs value the evidence for ST, but not necessarily see it as critical to their professional development? And will novice TAs at other institutions value the ability to model teaching techniques with their peers? Will TAs at institutions implementing few EBTP desire more evidence for ST? Are our findings transferable to faculty PD? Are there actionable differences in the perceptions of graduate and undergraduate TAs? Future work will address these questions.

SUPPLEMENTAL MATERIALS

Appendix I: Survey instrument

ACKNOWLEDGMENTS

The authors thank the fall 2018 facilitators for all of their help: Petra Kranzfelder, Jenna Hicks, Seth Thompson, Deena Wassenberg, Charles Willis, and Kristina Prescott. Zoe Koth also helped organize the fall 2018 workshop. Sadie Hebert provided invaluable assistance in administering the surveys. We also thank the TAs who participated in the workshops; without them, we would have no data. This project was funded by National Science Foundation award #1712033 (S. Cotner, PI). The authors declare that they have no conflicts of interest.

REFERENCES

1. American Association for the Advancement of Science. 2011. Vision and change in undergraduate biology education: a call to action. 2010. [Online.] http://www.visionandchange.org/VC_report.pdf.
2. Brownell SE, Kloser MJ, Fukami T, Shavelson R. 2012. Undergraduate biology lab courses: comparing the impact of traditionally based “cookbook” and authentic research-based courses on student lab experiences. *J Coll Sci Teach* 41:36–45.
3. Freeman S, Eddy SL, McDonough M, Smith MK, Okoroafor N, Jordt H, Wenderoth MP. 2014. Active learning increases student performance in science, engineering, and mathematics. *Proc Natl Acad Sci* 111:8410–8415.
4. Lopatto D. 2007. Undergraduate research experiences support science career decisions and active learning. *CBE Life Sci Educ* 6:297–306.
5. Luckie DB, Aubry JR, Marengo BJ, Rivkin AM, Foos LA, Maleszewski JJ. 2012. Less teaching, more learning: 10-yr study supports increasing student learning through less coverage and more inquiry. *Adv Physiol Educ* 36:325–335.
6. Feinstein NW, Allen S, Jenkins E. 2013. Outside the pipeline: reimagining science education for nonscientists. *Science* 340:314–317.
7. Miller CJ, Metz MJ. 2014. A comparison of professional-level faculty and student perceptions of active learning: its current use, effectiveness, and barriers. *Adv Physiol Educ* 38:246–252.
8. Patrick LE, Howell LA, Wischusen EW. 2016. Perceptions of active learning between faculty and undergraduates: differing views among departments. *J STEM Educ Innov Res* 17:55–63.
9. Schussler EE, Read Q, Marbach-Ad G, Miller K, Ferzli M. 2015. Preparing biology graduate teaching assistants for their roles as instructors: an assessment of institutional approaches. *CBE Life Sci Educ* 14(3):ar31.
10. Sundberg MD, Armstrong JE, Wischusen EW. 2005. A reappraisal of the status of introductory biology laboratory education in U.S. colleges & universities. *Am Biol Teach* 67:525–529.
11. Bettinger EP, Long BT, Taylor ES. 2016. When inputs are outputs: the case of graduate student instructors. *Econ Educ Rev* 52:63–76.
12. Hughes PW, Ellefson MR. 2013. Inquiry-based training improves teaching effectiveness of biology teaching assistants. *PLOS One* 8:e78540.
13. Roehrig GH, Luft JA, Kurdziel JP, Turner JA. 2003. Graduate teaching assistants and inquiry-based instruction: implications for graduate teaching assistant training. *J Chem Educ* 80:1206.
14. Pfund C, Miller S, Brenner K, Bruns P, Chang A, Ebert-May D, Fagen AP, Gentile J, Gossens S, Khan IM, Labov JB, Pribbenow CM, Susman M, Tong L, Wright R, Yuan RT, Wood WB, Handelsman J. 2009. Summer institute to improve university science teaching. *Science* 324:470–471.
15. Gregg CS, Ales JD, Pomarico SM, Wischusen EW, Siebenaller JF. 2013. Scientific teaching targeting faculty from diverse institutions. *CBE Life Sci Educ* 12:383–393.
16. Roden JA, Jakob S, Roehrig C, Brenner TJ. 2018. Preparing graduate student teaching assistants in the sciences: an intensive workshop focused on active learning. *Biochem Mol Biol Educ* 46(4):318–326.
17. Wyse SA, Long TM, Ebert-May D. 2014. Teaching assistant professional development in biology: designed for and driven by multidimensional data. *CBE Life Sci Educ* 13:212–223.
18. Wheeler LB, Maeng JL, Chiu JL, Bell RL. 2016. Do teaching assistants matter? Investigating relationships between teaching assistants and student outcomes in undergraduate science laboratory classes. *J Res Sci Teach* 54(4):463–492.
19. Pavelich MJ, Streveler RA. 2004. An active learning, student-centered approach to training graduate teaching assistants, pp. FIE-1. In *IEEE, 34th Annual Frontiers in Education, 2004. FIE 2004*, Vol. 2.
20. Marbach-Ad G, Schaefer KL, Kumi BC, Friedman LA, Thompson KV, Doyle MP. 2012. Development and evaluation of a prep course for chemistry graduate teaching assistants at a research university. *J Chem Educ* 89:865–872.
21. Boman JS. 2013. Graduate student teaching development: evaluating the effectiveness of training in relation to graduate student characteristics. *Can J Higher Educ* 43:100–114.
22. DeChenne SE, Kozioł N, Needham M, Enochs L. 2015. Modeling sources of teaching self-efficacy for science, technology, engineering, and mathematics graduate teaching assistants. *CBE Life Sci Educ* 14:ar32.
23. Gardner GE, Jones MG. 2011. Pedagogical preparation of the science graduate teaching assistant: challenges and implications. *Sci Educ* 20:31.
24. Campbell DT, Stanley J. 1963. Experimental and quasi-experimental designs for research 8–12.
25. Vygotsky LS. 1978. *Mind in society: the development of higher psychological processes*. Harvard University Press, Cambridge, MA.
26. Dewey J. 1933. *How we think: a restatement of the relation of reflective thinking to the educative process*. DC Heath.
27. Vanderstraeten R. 2002. Dewey’s transactional constructivism. *J Philos Educ* 36:233–246.
28. Eddy SL, Hogan KA. 2014. Getting under the hood: how and for whom does increasing course structure work? *CBE Life Sci Educ* 13:453–468.
29. Armbruster P, Patel M, Johnson E, Weiss M. 2009. Active learning and student-centered pedagogy improve student attitudes and performance in introductory biology. *CBE Life Sci Educ* 8:203–213.
30. Udovic D, Morris D, Dickman A, Postlethwait J, Wetherwax P. 2002. Workshop biology: demonstrating the effectiveness of active learning in an introductory biology course. *AIBS Bull* 52:272–281.
31. Doymus K. 2008. Teaching chemical equilibrium with the jigsaw technique. *Res Sci Educ* 38:249–260.
32. Handelsman J, Miller S, Pfund C. 2007. *Scientific teaching*. Macmillan, New York.
33. Angelo TA, Cross K, Recknue CA. 1993. *A handbook for college teachers*. Jossey-Bass Publishers, San-Francisco.
34. Brazeal KR, Brown TL, Couch BA. 2016. Characterizing student perceptions of and buy-in toward common formative assessment techniques. *CBE Life Sci Educ* 15:ar73.

35. Tanner KD. 2013. Structure matters: twenty-one teaching strategies to promote student engagement and cultivate classroom equity. *CBE Life Sci Educ* 12:322–331.
36. McDonald KK, Gnagy SR. 2015. Lights, camera, acting transport! Using role-play to teach membrane transport. CourseSource. doi: 10.24918/cs.2015.12.
37. Sadler TD, Romine WL, Menon D, Ferdig RE, Annetta L. 2015. Learning biology through innovative curricula: a comparison of game- and nongame-based approaches. *Sci Educ* 99:696–720.
38. Spiegel CN, Alves GG, Cardona T da S, Melim LM, Luz MR, Araújo-Jorge TC, Henriques-Pons A. 2008. Discovering the cell: an educational game about cell and molecular biology. *J Biol Educ* 43:27–36.
39. Bonney KM. 2015. Case study teaching method improves student performance and perceptions of learning gains. *J Microbiol Biol Educ* 16:21–28.
40. Alozie N, Grueber D, Dereski M. 2012. Promoting 21st-century skills in the science classroom by adapting cookbook lab activities: the case of DNA extraction of wheat germ. *Am Biol Teach* 74:485.
41. Warfa ARM. 2016. Mixed-methods design in biology education research: approach and uses. *CBE Life Sci Educ* 15:rm5.
42. R Development Core Team. 2013. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
43. Wickham H. 2016. *ggplot2: elegant graphics for data analysis*. Springer-Verlag, New York.
44. Saldaña J. 2015. *The coding manual for qualitative researchers*. Sage.
45. Loucks-Horsley S, Stiles KE, Mundry S, Love N, Hewson PW. 2009. *Designing professional development for teachers of science and mathematics*. Corwin Press.
46. Wheeler LB, Clark CP, Grisham CM. 2017. Transforming a traditional laboratory into an inquiry-based course: importance of training TAs when redesigning a curriculum. *J Chem Educ* 94:1019–1026.
47. Becker EA, Easlson EJ, Potter SC, Guzman-Alvarez A, Spear JM, Facciotti MT, Igo MM, Singer M, Pagliarulo C. 2017. The effects of practice-based training on graduate teaching assistants' classroom practices. *CBE Life Sci Educ* 16:ar58.
48. Goodwin EC, Cao JN, Fletcher M, Flaiban JL, Shortlidge EE, Stains M. 2018. Catching the wave: are biology graduate students on board with evidence-based teaching? *CBE Life Sci Educ* 17:ar43.