

# Brief Training and Intensive Mentoring Guide Postdoctoral Scholars to Student-Centered Instruction

R. M. Price,<sup>†\*</sup> C. J. Self,<sup>‡</sup> W. C. Young,<sup>§</sup> E. R. Klein,<sup>||</sup> S. Al-Noori,<sup>¶</sup> E. Y. Ma,<sup>#</sup> and A. DeMarais<sup>@</sup>

<sup>†</sup>School of Interdisciplinary Arts and Sciences and <sup>‡</sup>School of Technology, Engineering & Mathematics, University of Washington Bothell, Bothell, WA 98011-8246; <sup>§</sup>Department of Biology, University of Washington, Seattle, WA 98195-1800; <sup>¶</sup>Fred Hutchinson Cancer Research Center, Seattle, WA 98109-1024; <sup>||</sup>Educurious Partners, Seattle, WA 98102; <sup>#</sup>School of Interdisciplinary Arts and Sciences, University of Washington Tacoma, Tacoma, WA 98402-3100; <sup>@</sup>Department of Biology, University of Puget Sound, Tacoma, WA 98416

## ABSTRACT

The Science Teaching Experience Program—Working in Science Education (STEP-WISE) provides teaching experience for postdoctoral scholars holding full-time research appointments. Through a combination of mentorship, deliberate practice, and feedback, the postdocs learn and apply inclusive, evidence-based pedagogies. STEP-WISE is integrated into postdocs' demanding schedules and is sustainable for institutions to run. Here, we assess the effectiveness of STEP-WISE. We used the Classroom Observation Protocol for Undergraduate STEM instruction to quantify instructor and student behaviors in 20 STEP-WISE class sessions from seven courses designed and taught by postdocs in the program. We found that all of the postdocs used student-centered teaching strategies. Also, using a design-based research framework, we studied the program to identify the salient components of its design. Four interconnected key elements contribute to the program's success: 1) two training sessions, 2) a precourse meeting with the mentor, 3) implementation of active-learning strategies with support, and 4) debriefing with the mentor after each class session. STEP-WISE is a replicable model to support postdocs seeking training and experience in evidence-based teaching practices geared to improving undergraduate education and transforming pedagogical practice. We conclude that high-impact teaching can be learned early in a career with streamlined training and intensive mentoring.

## INTRODUCTION

Undergraduate instruction in science, technology, engineering, and mathematics (STEM) needs to improve (American Association for the Advancement of Science [AAAS], 2011; Henderson *et al.*, 2011; National Research Council, 2012; Freeman *et al.*, 2014; Theobald *et al.*, 2020; Harris *et al.*, 2020), and researchers are uncovering evidence for effective and inclusive strategies that address this need for improvement (AAAS, 2011; Tanner, 2013; Freeman *et al.*, 2014; Wilson and Brame, 2018; Theobald *et al.*, 2020). Even as the research progresses, most STEM instructors still do not use student-centered teaching strategies, relying instead on the familiar—and demonstrably less effective (Freeman *et al.*, 2014) and racially biased (Theobald *et al.*, 2020)—format of lecture-only classes (Stains *et al.*, 2018). According to Stains *et al.* (2018), for example, only 18% of more than 2000 STEM university class sessions were student-centered; the rest of the sessions were didactic (55% of the class sessions) or interactive lecture (27%).

A number of barriers interfere with the ability of STEM instructors to adopt more effective teaching (AAAS, 2011, 2019). For example, instructors may not know which strategies to adopt (Henderson *et al.*, 2011) or even how to adopt strategies as they

Gili Marbach-Ad, *Monitoring Editor*

Submitted Mar 31, 2021; Revised Sep 10, 2021; Accepted Sep 13, 2021

CBE Life Sci Educ December 1, 2021 20:ar64

DOI:10.1187/cbe.21-03-0083

\*Address correspondence to: R. M. Price (beccap@uw.edu).

© 2021 R. M. Price *et al.* CBE—Life Sciences Education © 2021 The American Society for Cell Biology. This article is distributed by The American Society for Cell Biology under license from the author(s). It is available to the public under an Attribution–Noncommercial–Share Alike 3.0 Unported Creative Commons License (<http://creativecommons.org/licenses/by-nc-sa/3.0>).

“ASCB®” and “The American Society for Cell Biology®” are registered trademarks of The American Society for Cell Biology.

were intended to be implemented (Offerdahl *et al.*, 2018). Faculty time is often limited, which impedes the ability to foster extended immersion that promotes successful change, such as that provided through professional learning communities (Henderson *et al.*, 2011).

A potential solution to this conundrum is to train early-career scientists, such as postdoctoral scholars, how to teach more effectively (Ebert-May *et al.*, 2015). This solution benefits postdocs as they enter the faculty job market, because faculty hiring criteria include not only more requirements than they have in the past, but also teaching experience that is demonstratively high quality (Fleet *et al.*, 2006; Alberts *et al.*, 2014; Kelsky, 2015). Ebert-May and colleagues developed the FIRST IV program, a highly successful model for training postdocs to teach with evidence-based and student-centered methods, and the postdocs who completed this program continued to leverage these teaching strategies years later as faculty members (Ebert-May *et al.*, 2015; Derting *et al.*, 2016; Emery *et al.*, 2020). Briefly, the postdocs in this program were selected from throughout the United States to attend an intense, 4-day pedagogical training in each of two consecutive summers. Postdocs taught their own undergraduate courses during the intervening school year, consulting with mentors as they did so (Ebert-May *et al.*, 2015).

Another successful approach to providing postdoctoral scholars with skills to teach has been the Institutional Research and Academic Career Development Awards (IRACDA) from the National Institute of General Medicine (National Institute of General Medical Sciences, 2021). Through this granting opportunity, institutions have developed programs that train postdocs in both effective pedagogy and research. Currently, more than 60 institutions—the majority of which are minority serving—have IRACDA programs (National Institute of General Medical Sciences, 2021). These programs are holistic, addressing multiple skills required for success in academic careers, including teaching, and graduates of the programs have been successful in obtaining and succeeding in faculty positions (Rybarczyk *et al.*, 2011, 2016; Eisen and Eaton, 2017). However, the pedagogical training is intensive; in one IRACDA program, for example, postdocs take a semester-long course on effective teaching and then teach a semester-long course (Eisen and Eaton, 2017).

A drawback to the IRACDA and FIRST IV programs is that they are time-intensive for postdoctoral scholars—scientists at a career stage with many conflicting demands on their time. Moreover, FIRST IV necessitates that participants travel, rather than integrating opportunities for teaching-based professional development into postdocs' everyday responsibilities. Another drawback has been, because of their federal support, the exclusion of international postdocs from the training opportunities. Also, both FIRST IV and IRACDA have been expensive—paying for travel and intensive mentoring or multiyear stipends for postdocs.

The Science Teaching Experience Program—Working in Science Education (STEP-WISE) is a practice based, collaborative model for training postdocs in student-centered teaching (Figure 1). This model removes the burden of travel for training and the requirement to teach an entire semester-long course. Postdocs collaborate in teams of three to deliver a 10-week undergraduate seminar course at their home institutions. The program is designed for postdocs in 100% research

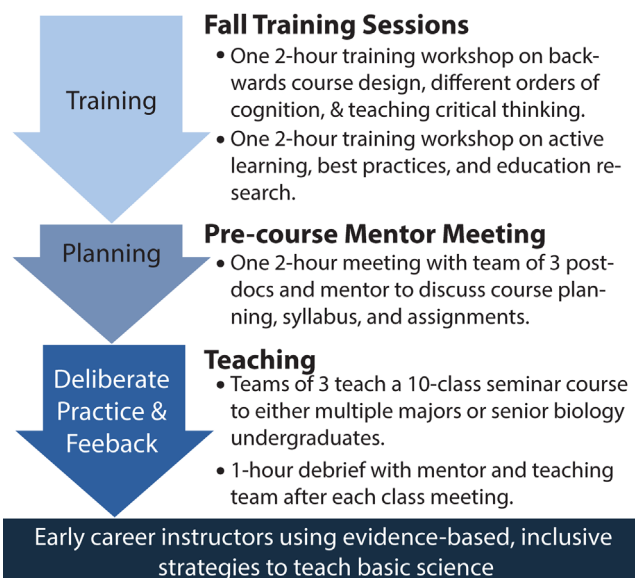


FIGURE 1. Overview of STEP-WISE.

positions, so research productivity can be maintained throughout the apprenticeship. Pedagogical training occurs alongside their research and is addressed for a few hours each week that are distributed over two quarters, so postdocs are practicing the skill that faculty need to teach and conduct research simultaneously. Much of the training occurs during debriefs after each class session, when a teaching mentor facilitates self-reflection that is directed by a rubric about what worked in class and what to improve. STEP-WISE is cost-efficient and sustainable from the perspective of the institution. The cost is built into the curriculum, so long as seminar courses are taught, because the instructor of record for the course is the teaching team's faculty mentor. The mentor advises the team as they prepare the syllabus, attends each class session to observe, and facilitates the debrief.

Here, we provide a rigorous analysis of the success of STEP-WISE in training postdocs to teach with student-centered strategies. To test its effectiveness, we compared the kinds of instruction used by STEP-WISE postdocs to the kinds reported in Stains *et al.*'s (2018) national database of the instructional styles used by STEM faculty. Both this study and Stains *et al.*'s (2018) study used the same instrument, the Classroom Observation Protocol for Undergraduate STEM instruction (COPUS; Smith *et al.*, 2013), to record teaching styles employed in classrooms and the frequency at which they were used. Specifically, we tested the hypothesis that postdocs in STEP-WISE would use student-centered instruction that encouraged collaborative learning among students and that they would do so at a higher rate than the instructors in Stains *et al.*'s (2018) study.

In addition to this quantitative analysis, we used a design-based research framework (Cobb *et al.*, 2003; Design-Based Research Collective, 2003) to collect and analyze data that explore how STEP-WISE was effective at helping beginning instructors use evidence-based and inclusive pedagogies and to identify the salient features of STEP-WISE that promoted postdocs' use of evidence-based and inclusive pedagogies.

## STEP-WISE

Since 2011, STEP-WISE has been providing teaching experiences for postdocs who have 100% research appointments. The STEP-WISE mission is to “engage a diverse pool of postdoctoral scholars at the University of Washington (UW) and affiliate institutions in a closely mentored apprenticeship to learn how to teach scientifically with inclusive, demonstrably effective, student-centered pedagogies” (Science Teaching Experience Program, 2021a). Most postdoctoral scholars at the UW conduct research in biomedical fields, and all the courses that STEP-WISE postdocs teach are tied to biology curricula. Postdocs work in teams of three to develop and teach a special-topics seminar course with no more than 24 undergraduate students enrolled. STEP-WISE expands the curricula offered by the UW Department of Biology, the UW Bothell School of Interdisciplinary Arts and Sciences, the UW Bothell School of Science, Technology, Engineering and Mathematics, and the University of Puget Sound Biology Department. The topics are cutting edge, and have included the neuroscience of pain and addiction, CRISPR/Cas 9, the microbiome, and disease treatments.

Through their experience in STEP-WISE, 168 postdocs have gained the experience they need to evaluate whether they want a career that involves undergraduate teaching, as well as the experience necessary to be competitive for faculty positions with a significant teaching component. While the program focuses on teaching effectiveness, and therefore faculty positions, a number of graduates have found the skills they learned useful as they transition to careers in the biotech industry (e.g., Nygren, 2018).

The apprenticeship takes place over two academic quarters. In the Autumn, postdocs receive instruction in two 2-hour training sessions for the whole cohort (Figure 1, Training). In the first session, they learn about backward design (Wiggins and McTighe, 2005), and each team identifies a course topic based on their shared research expertise. In the second session (Figure 1), they learn about active-learning strategies (Ma *et al.*, 2021), such as the jigsaw (Aronson and Bridgeman, 1979; Clarke, 1994), gallery walk (Francek, 2006), think-pair-share (Tanner, 2013), unbiased call (Waugh and Andrews, 2020), and quick writes (Tanner, 2013). Before their teaching begins, the postdocs work in their teams of three to design a quarter-long seminar. This planning includes at least one session with their faculty mentor to discuss the course design and improve the syllabus (Figure 1, Planning).

The second quarter of the apprenticeship is the teaching quarter (Figure 1, Deliberate Practice and Feedback), which can be Winter, Spring, or Summer. During this term, the postdocs teach their seminar, with each postdoc leading three classes. Classes meet once per week for 10 weeks in the Winter and Spring quarters, and twice a week for 5 weeks in an accelerated Summer quarter. The 10th and final class session is usually reserved for undergraduates presenting final projects or drafts of final projects. The entire teaching team, including the mentor, attends each class session, after which they debrief as a team for an hour. The mentor notes the inclusive teaching strategies used, including strategies used to promote scientific approaches such as critically analyzing experimental design, interpreting data, and proposing new experiments. In the class sessions and debriefs, the postdocs learn collaboratively from one another's teaching. After the debrief, the mentor shares a written observation of each class session with the teaching team.

Debriefs also address assessment, grading strategies, and time management. Mentors advise the postdocs to limit their class preparation time to 2–4 hours per class session and to limit their grading time to 1–2 hours per week. However, postdocs typically begin their teaching by spending an excess of 10 hours of preparation per class. In the assessment of the STEP-WISE presented here, the postdocs reported that they averaged 8.5 hours of preparation (min. = 2.8, max. = 32,  $N = 18$ ) across each of their three class meetings.

## Scientific Teaching

The design of STEP-WISE is informed by scientific teaching, a pedagogical philosophy that frames teaching as a continual process of hypothesis testing (Handelsman *et al.*, 2004, 2006; Couch *et al.*, 2015). Every teaching decision is based on the hypothesis that students will learn what an instructor intends, and the instructor is constantly collecting data to test that hypothesis for all students in the course. An aspect of inclusivity that is built into STEP-WISE is the expectation that postdocs employ a variety of assessments to measure learning in the entire population of students.

STEP-WISE leadership has ensured that the program stay current with best pedagogical practices, employing many of the criteria specified in a taxonomy of scientific teaching (AAAS, 2011; Couch *et al.*, 2015) and refinements to well-established teaching techniques (e.g., discussions of unbiased call; Waugh and Andrews, 2020; Cooper *et al.*, 2021). Couch *et al.* (2015) describe scientific teaching as occurring in courses in which content is well aligned to course learning goals and in which students apply scientific processes, participate actively in their learning, and engage in higher-level cognition. The initial two 2-hour training sessions introduce postdocs to these aspects of teaching.

The first training session introduces postdocs to backward design (Wiggins and McTighe, 2005; Figure 1), the practice of identifying course-level learning goals, and breaking these down into goals for each class session (Handelsman *et al.*, 2004; Wiggins and McTighe, 2005; Felder and Brent, 2016; Sana *et al.*, 2020). As part of this introduction, we also present Bloom's taxonomy (as described by Crowe *et al.*, 2008; Dirks *et al.*, 2014; Felder and Brent, 2016), so postdocs can begin thinking about different orders of cognition as they write learning goals. We emphasize that analyzing and evaluating data from the primary literature and designing new experiments both occur at higher orders of cognition. By the end of this workshop, members of each teaching team have met each other, identified a topic for their seminar, brainstormed on the final project that they want students to accomplish, and identified a set of potential learning goals. The current version of this workshop (as of Autumn 2020) also encourages postdocs to reflect on the teaching practices that the workshop facilitators used, including learning goals articulated across Bloom's taxonomy, unbiased call, using participants' names, individual thinking and writing time, group work, a structured worksheet, and whole-group discussion.

A major component of scientific teaching is on inclusive teaching to reach the full diversity of students in the classroom (Handelsman *et al.*, 2006). The second training session (Figure 1) emphasizes this component by introducing teaching methods that are both active and inclusive (as in

AAAS, 2011), and it is described in detail, complete with worksheets, elsewhere (Ma *et al.*, 2021). In academic year 2018–2019, when this study took place, the postdocs completed a jigsaw, a technique in which groups of students form different areas of expertise, and then students reshuffle into new groups where each member shares the expertise from their first groups (Aronson and Bridgeman, 1979; Clarke, 1994). The expert groups in this jigsaw studied one of three articles about the efficacy of active-learning strategies: 1) an article describing an activity or homework assignment that helps students understand figures from the primary literature (Round and Campbell, 2013), 2) an article that finds the jigsaw to be a more collaborative group learning exercise than when students complete an entire worksheet as a group (Theobald *et al.*, 2017), and 3) an article that evaluates how well instructors adopt student-centered learning (Ebert-May *et al.*, 2011). In addition to the jigsaw, this training session includes a think–pair–share (Tanner, 2013; Cooper *et al.*, 2021), which ensures that each person thinks individually and pairs up with another student to discuss; gallery walk (Francek, 2006), in which students take turns writing their group’s answers at different stations; quick writes (Tanner, 2013), in which students write down their own ideas; and unbiased call (with the option to pass or not be called; Waugh and Andrews, 2020) about open-ended questions. STEP-WISE is designed to encourage and support postdocs as they implement these inclusive strategies through deliberate, classroom practice followed by reflection after each class session during the teaching quarter.

Approximately 1 month before teaching, the team of postdocs returns to the idea of learning goals, higher orders of cognition, and scientific processes at a precourse meeting with their mentor (Figure 1). This is the point in STEP-WISE at which the postdocs and their mentor begin working closely together and when the postdocs begin to receive feedback about the course plan. The postdocs prepare for this meeting by drafting their syllabus and identifying one or two articles per postdoc from the primary literature that they want to use as the focus for their teaching. The mentors review the learning goals on the syllabus to ensure that they include higher orders of cognition and scientific processes. At this meeting, the mentor also discusses the final project with the postdocs to make sure that it focuses on applying scientific processes, such as students proposing an experiment that would follow up on one of the articles they read.

During teaching (Figure 1), the mentors remind postdocs to focus on their learning goals as they decide which components of the content to expand and contract. The goal is for mentors to remind postdocs to consistently use the data they are collecting (e.g., what students have written in worksheets or on whiteboards, their responses during unbiased call, the questions they ask while instructors circulate) to dictate the pace of class, aligning content to students’ needs.

## MATERIALS AND METHODS

We used data from the 21 STEP-WISE participants (seven courses, each taught by three postdocs) from the 2018–2019 cohort (Table 1) and three faculty mentors (authors R.M.P., S.A.-N., and E.Y.M.). We chose this year because we had funding to hire an external evaluator (author E.R.K.).

**TABLE 1. Demographics of the 2018–2019 STEP-WISE cohort (N = 21)**

Gender	10 men 11 women
Race	15 white 3 Latinx (any race) 3 Asian
U.S. citizens	14 yes
Years since PhD (mean, SD)	2.6 years ± 2.4
Prior teaching experience (median, min.–max.)	
Teaching assistantship	1, 0–3 courses
Guest lectures	0, 0–2 class meetings
Instructor of record	0, 0–6 courses

Two of the classes were aimed at second-year students from a variety of majors (including biology) and the other five were for senior biology majors (Supplementary Material Table 1). Four of the courses were offered at the UW Seattle. The other three courses occurred at the UW Bothell, a master’s regional branch campus that is classified as an Asian American and Native American Pacific Islander minority-serving institution. Classes ranged in size from 10 to 24 undergraduate students (mean = 16, SD = 5).

The postdocs were beginning instructors. They had taught between 0 and 6 classes (median = 0), guest lectured between 0 and two times (median = 0), and been graduate teaching assistants for between 0 and 3 courses (median = 0).

The UW Institutional Review Board determined that the data used in this research were exempt from review under category 1, normal education practices (STUDY00005114).

## Comparing the Teaching of STEP-WISE Postdocs with a National Sample

**Capturing Instructor and Student Behaviors.** While many classroom observation protocols exist (e.g., Eddy *et al.*, 2015; Hora, 2015; Durham *et al.*, 2018), the COPUS (Smith *et al.*, 2013) is becoming a standard instrument for quantifying the amount of time spent on active instruction in classrooms (e.g., Lund *et al.*, 2015; Stains *et al.*, 2018; Zagallo *et al.*, 2019; McNeal *et al.*, 2020; Denaro *et al.*, 2021). Moreover, the large set of COPUS data by Stains *et al.* (2018), as linked to in their supplementary materials, provides a control group that represents how most instructors teach STEM courses in the United States. Because our goal is to measure the strengths of the program, we intentionally chose to evaluate the success of STEP-WISE with respect to overall teaching patterns, rather than comparing this program to others for postdocs. Comparing STEP-WISE to other programs that successfully increase the amount of active learning, such as FIRST IV, would require a different, and particularly sensitive, observation protocol (Hora, 2015).

To analyze STEP-WISE classes with the COPUS, we made audio recordings of each instructor, typically from the second class session taught by each. Using audio recordings as the basis for COPUS data meant that the evaluator could focus her observations on behaviors that were not necessarily captured by the COPUS while she took field notes. The instructor wore the recorder, so we could hear him or her moving around the room and the voices of students changing. We could also hear them



writing on the whiteboard. One instructor was inadvertently not recorded, and therefore, our sample consists of 20 postdoctoral scholars. Class sessions were 2 hours long, but we only coded the first hour of instruction. This decision was made to be conservative: more announcements and other kinds of administration that is instructor-centered, rather than student-centered, tend to occur in the first hour of a class session. In STEP-WISE classes, it is also common to have the second hour of class particularly student-centered, for example, with students working in groups to make predictions, articulate hypotheses, and design experiments to test those hypotheses. The observations conducted in parallel to this study, combined with notes written by mentors, confirm that student-centered instruction continued in the second hour of the classes.

Authors R.M.P. and C.J.S. coded the audio recordings with the COPUS. They followed the training protocol outlined in the paper that introduces the COPUS (Smith *et al.*, 2013), first talking through the different behaviors, then reviewing the sample videos the authors of the COPUS recommended for training. Because we did not have video data of the classes we were coding, we began by listening to the training videos without watching them. We then compared our initial audio-only coding of the videos to a second set of coding based on videos, and we confirmed that our coding was consistent between these modalities. R.M.P. and C.J.S. calibrated their scoring with the COPUS by sharing the codes that they assigned for the training videos and discussing any differences. Once they reached a level of comfort with the coding videos that were available online, they moved onto coding STEP-WISE classes. The COPUS documents a set of instructor and student behaviors that occur in 2-minute intervals. R.M.P. and C.J.S. would listen to the audio at the same time but recorded their codes independently. At every 2-minute interval, they would stop and compare what they had coded, resolving disagreements through discussion to reach a consensus. At times they referenced field notes taken by author E.R.K. and observations written by mentors as a way to double-check their coding. They also referenced worksheets that instructors distributed to students, for example, to determine whether instructors asked students to make predictions. R.M.P. is the executive director of STEP-WISE, but C.J.S. has no connection to the program; her independence from STEP-WISE was essential for minimizing bias due to conflict of interest. Notes about how they applied the COPUS are included in Supplementary Table 2.

**Statistical Analysis.** After each class was coded for the instructor and student behaviors present in every 2-minute interval, data were summarized to obtain the percentage of time each behavior occurred during the class. Instructors were the unit of analysis. Class summaries were then analyzed using the COPUS Analyzer (COPUS Analyzer, 2021) to document how each class fell into the clusters described by Stains *et al.* (2018). To compare the behaviors of instructors and students in STEP-WISE and the data collected by Stains *et al.* (2018), we used the Mann-Whitney *U*-test to determine whether each behavior was significantly different between populations. The Mann-Whitney *U*-test is a nonparametric comparison of the medians between two populations, analogous to the parametric *t* test; we considered *p* values <0.05 to indicate a statistically significant difference between medians (Wilcoxon, 1945; Mann and Whitney,

1947). The test produces a *U*-statistic, from which we can calculate a *p* value measuring the probability of obtaining a value at least as extreme as the observed value if there is no difference between the populations. This test is appropriate for samples that differ in population size, and it is conservative in this case, because the power decreases when the sample imbalance is large. The fact that we observe statistical differences—even with such large sample size imbalances—is noteworthy. The *p* values were then adjusted to account for multiple testing using the Holm-Bonferroni method (Holm, 1979). Behaviors were compared between STEP-WISE and the full Stains data set and also between STEP-WISE and cluster 6 from Stains *et al.* (2018), a cluster of student-centered instructors often typified by group work and worksheets. All STEP-WISE classes were also categorized as cluster 6.

Spearman rank correlations were calculated among the STEP-WISE behaviors across all classes to generate a correlation heat map.

These analyses were conducted in R (R Core Team, 2020).

### Identifying the Key Elements of the Program's Design

The results of the COPUS analysis indicated that the STEP-WISE postdocs teach actively and to a great extent compared with even the most active instructors in a national data set (Stains *et al.*, 2018). We adopted a design-based research framework (Cobb *et al.*, 2003; Design-Based Research Collective, 2003) to explore which design elements of STEP-WISE were particularly salient to the postdocs' outcomes. Design-based research is often used iteratively to improve programs through time; here, we conducted the first of these iterative analyses. Our intent was to identify the observed outcomes of STEP-WISE and then to connect them to the designed elements that led to those outcomes. These data are derived from the same seven courses—and the teams of three that taught each—that were included in the COPUS analysis.

**Data Collection.** Aligned to the analytic framework of design-based research (Cobb *et al.*, 2003; Design-Based Research Collective, 2003), data about the programmatic elements of STEP-WISE and the nature of postdocs' collaborations were obtained from field notes (Emerson *et al.*, 2011), as well as analysis of audio recordings of whole-group training sessions.

Field notes were recorded by author and external, independent evaluator E.R.K. between September 2018 and August 2019. Just as our apprenticeship is divided into early training and deliberate practice, our data collection focused on observation of the training and the postdocs' teaching (i.e., pedagogies of enactment; Grossman *et al.*, 2009).

E.R.K. made field observations (Emerson *et al.*, 2011) approximately every third week of class to observe each postdoc once and to capture the trajectory of each team's collective learning. The field observations included the 2-hour class session and the 1-hour debrief. During each class, E.R.K. focused on documenting the main instructor's teaching practice, co-instructors' collaboration, undergraduate students' participation and engagement, and the mentor's mode of support and engagement. During the debriefs, she focused on how mentors structured the hour, the forms of support and feedback offered by the mentor, and how postdocs engaged with their mentor and one another.

We supplemented these data streams with reflections by the participants obtained from 30-minute focus group interviews conducted by E.R.K. with each teaching team (Supplementary Table 3). These focus groups were transcribed to obtain a fine-grained depiction of participants' experiences and allow for coding and analysis. Nineteen of the STEP-WISE participants also completed an online evaluation of the apprenticeship within 1 month of completing their teaching (Supplementary Table 4). To obtain data from the mentor's perspective, E.R.K. conducted semistructured interviews with each mentor (Supplementary Table 3). These interviews revealed how mentors were prepared to support novice teachers, to understand their mentoring experiences during and before the year under study, and to collect feedback on program structures and supports.

Finally, E.R.K. analyzed handouts that the mentors provided at Autumn training sessions, worksheets created by the postdocs for their classes, email messages between postdocs and mentors, and mentors' written observations of each class session.

**Data Analysis.** We analyzed data qualitatively, using each teaching team (postdocs and mentor) as the unit of study. Each team of postdocs formed small, embedded units of analysis, given the collaborative learning that took place during debrief sessions. Our goal was to study the outcomes for this STEP-WISE cohort and identify key elements of the program design.

We used Dedoose v. 8.2.14 to index across data streams, writing additional memos (Miles *et al.*, 2013) that documented noteworthy aspects about each artifact and making connections across artifacts. Rather than conducting an ethnographic analysis of only our field notes, we used open coding (Patton, 2014) across all types of qualitative data to identify and explore the data, and we used structural coding to analyze the effect of particular goals of the program. The coded data were then used to relate these findings to the designed structures and the mentors' reflections, allowing us to build arguments ("higher-level claims" in Strauss and Corbin, 1990) about why certain aspects of the program's design were identified as key elements that impacted participants' outcomes. Accordingly, the *Results* section provides descriptions of the key program design elements in relation to the experiences and outcomes for the postdocs.

Most of the qualitative data we collected did not lend itself to quantitative analysis, because of the nature of open-ended and semistructured prompts (Patton, 2014). However, the evaluation survey that the postdocs completed included multiple-choice and rating questions (Supplementary Table 4). The results from these questions were analyzed with basic descriptive statistics in Microsoft Excel.

## RESULTS

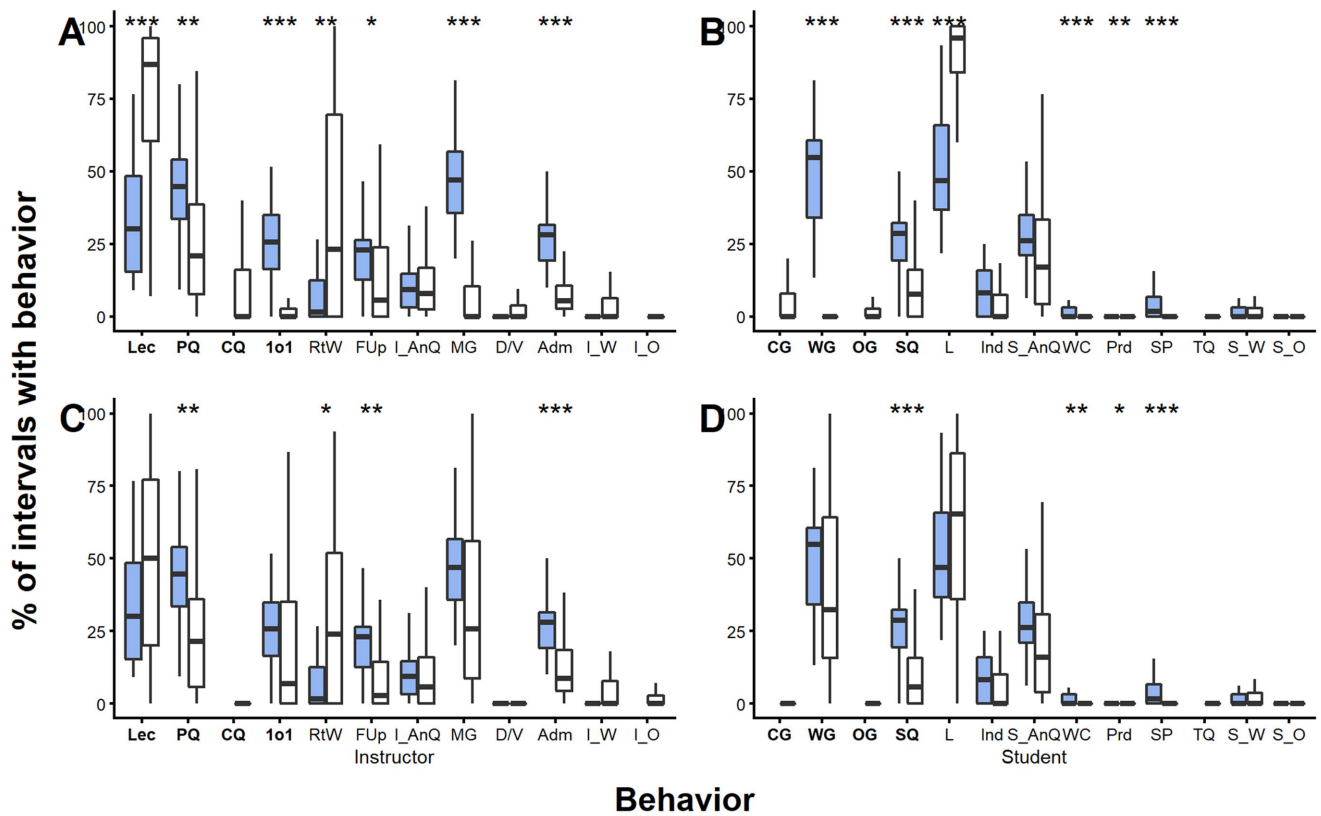
Our comparison of STEP-WISE teaching to a national sample, supplemented by the programmatic evaluation conducted with a design-based research framework, indicate that STEP-WISE is meaningful and successful. We first present the results of our analysis quantifying the amount of active learning observed and the relationships among instructor and student behaviors. We then present the four key elements in the design of STEP-WISE (Table 2). The key elements are arranged in the order during which they appear in the program. Quotations are from STEP-WISE postdocs, except when indicated. Sections about the key elements also include results from the evaluation survey.

### STEP-WISE Instructors Use Student-Centered Strategies at a High Rate

Seven of the 12 instructor behaviors were significantly different between the STEP-WISE and Stains *et al.* (2018) data (Figure 2A, asterisks, and Supplementary Table 5), and two of the instructor behaviors were absent from STEP-WISE courses (clicker questions, CQ; other instructor behaviors, I\_O). STEP-WISE postdocs spent significantly less time lecturing (Stains *et al.*, 2018; Figure 2A and B and Supplementary Table 5), and the instructors spent significantly more time on administration (setting up activities, Adm,  $p < 0.0001$ ), following up after activities ( $p = 0.02$ ), working with students one-on-one (1o1,  $p < 0.0001$ ), and guiding group work (MG,  $p < 0.0001$ ; Figure 2 and Supplementary Table 5). Many of the instructors in Stains *et al.*'s (2018) sample used clicker questions (CQ), but STEP-WISE instructors did not. Instead, they posed open-ended questions (PQ,  $p < 0.001$ ). Finally, significantly more instructors in the Stains *et al.*'s (2018) study wrote on the board in real time (RtW,  $p < 0.01$ ). Collectively, these results indicate that

**TABLE 2. Key design elements of STEP-WISE**

Key elements	Features
1. Autumn training sessions	<ul style="list-style-type: none"> <li>• Begin designing a course through backward design (session 1).</li> <li>• Experience active-learning strategies as postdocs learn about them (session 2).</li> <li>• Introduce postdocs to primary literature supporting different active-learning strategies (session 2).</li> </ul>
2. Precourse meeting with mentor	<ul style="list-style-type: none"> <li>• Hold a structured meeting between the teaching team and the faculty mentor approximately 1 month before the start of the course</li> <li>• Refine learning goals for the course and class meetings.</li> <li>• Begin scaffolding a final project that will unfold throughout the term.</li> <li>• Arrange digital collaboration.</li> </ul>
3. Implementing active learning with support	<ul style="list-style-type: none"> <li>• Implement inclusive teaching strategies, such as think-pair-share, jigsaw, gallery walk, and unbiased call.</li> <li>• Collaborate across instructional team members in class.</li> <li>• In-class collaboration among instructors.</li> <li>• Write and use worksheets that structure class sessions.</li> </ul>
4. Debriefs	<ul style="list-style-type: none"> <li>• Reflect on the class just taught, centering the experiences of the lead instructor.</li> <li>• Elicit additional reflection from co-instructors.</li> <li>• Build on postdocs' reflections by praising instruction, making suggestions for improvement, comparing with examples of success and failures from mentor's own teaching, and modeling language for classroom use.</li> </ul>

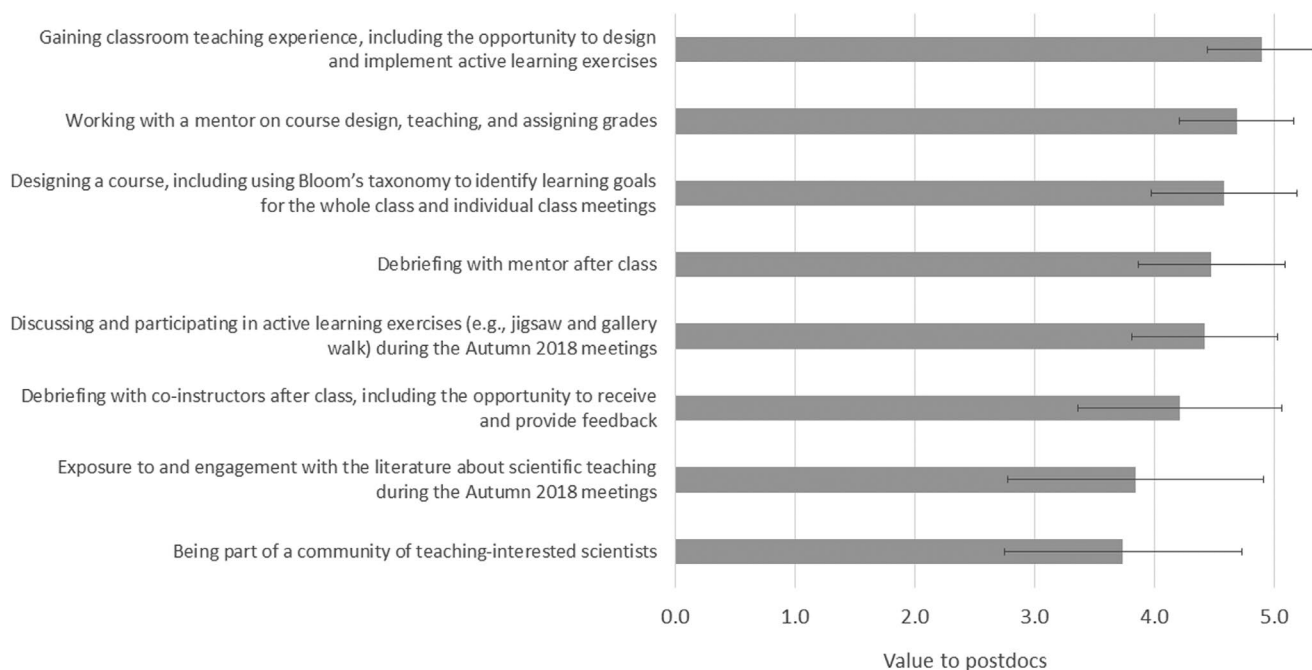


**FIGURE 2.** Summary of the COPUS results demonstrating student-centered teaching in STEP-WISE. Box plots of the percentage of 2-minute intervals with each behavior (x-axes) in the STEP-WISE courses (blue,  $N = 20$ ) and the courses coded by Stains *et al.* (2018) (white,  $N = 2008$ ) showing each behavior in the COPUS for (A) instructors and (B) students, and in the cluster 6 subsample of Stains *et al.* (2018) ( $N = 157$ ) that is most similar to the STEP-WISE data for (C) instructors and (D) students. Horizontal lines are medians, boxes include 50% of the data, whiskers extend to the farthest observation no farther than 1.5 times the interquartile range from the median. A single interval can include multiple behaviors, so more than one behavior may be present 100% of the time. \*Adjusted  $p < 0.05$ ; \*\*adjusted  $p < 0.01$ ; \*\*\*adjusted  $p < 0.0001$ . Labels for the first four variables for instructors and the first four variables for students are in boldface to indicate that they were used in the cluster analysis by Stains *et al.* (2018) to identify whether instruction in a class session was classified as didactic lecture, interactive lecture, or student-centered. Instructor behaviors: 1o1 Discussion with one or a few individuals, Adm Administration, CQ Clicker question, D/V demonstration or video, FUp Follow-up/feedback, L\_AnQ Answering student questions, L\_O Other, L\_W Waiting, Lec Lecturing, PQ Posing question, RtW Real-time writing, MG Moving through the room and guiding student work. Student behaviors: CG Clicker question groups, Ind Individual work, L Listening, OG Other group activity, Prd Making a prediction, S\_AnQ Student answering question, SQ Student asks question, S\_O other, S\_W students waiting, SP Student presentation, TQ Test or quiz, WC Whole class discussion, WG Working in groups. Variables are defined in more detail in Supplementary Table 2.

STEP-WISE postdocs’ teaching practices were more student-centered than those in the national database.

Student behavior also differed between the STEP-WISE classes and those studied by Stains *et al.* (2018). Six of the 13 student behaviors were significantly different between the STEP-WISE and Stains *et al.* (2018) data (Figure 2B, asterisks, and Supplementary Table 5), and another three were absent from STEP-WISE courses (clicker question groups, CG; other assigned group, OG; and test or quiz, TQ; Supplementary Table 5). The students in STEP-WISE classes spent significantly more time working in groups (WG,  $p < 0.0001$ ) and significantly less time listening to lecture (L,  $p < 0.0001$ ). Students in the STEP-WISE classes also spent significantly more time asking questions (SQ,  $p < 0.0001$ ) and presenting their ideas to the class (SP,  $p < 0.0001$ ). However, there was no significant difference in the amount of time students worked independently (Ind,  $p = 0.18$ ).

We also compared a subset of 157 classes in Stains *et al.* (2018)—their cluster 6—that were most similar to STEP-WISE seminars (Figure 2C and D). All of the courses depicted in Figure 2C and D are classified as student-centered by Stains *et al.*’s (2018) criteria. Compared with Stains *et al.* (2018) cluster 6 classes, more time in STEP-WISE classes is spent with both instructors posing (PQ,  $p = 0.0005$ ) and students asking (SQ,  $p < 0.0001$ ) questions. This result is apparent from the fact that four of the instructor behaviors and four student behaviors differed significantly between the STEP-WISE and the similarly taught courses in Stains *et al.* (2018) (Figure 2C and D, asterisks). The time spent following up on student ideas also corresponds to more time that instructors spend following up on student ideas (FUp,  $p = 0.0003$ ). Other behaviors trend toward being more active in STEP-WISE classes, although the results are not significant after applying



**FIGURE 3.** Mean value ( $\pm$ SD) of different program components as rated on the STEP-WISE exit survey ( $N = 19/21$  responses). Scale is 1–5, where 5 is most valuable.

the Holm-Bonferroni correction for multiple comparisons (Holm, 1979). For example, instructors in STEP-WISE classes tend to lecture less (Lec,  $p = 0.41$ ), and, correspondingly, students in STEP-WISE classes tend to spend more time working either on their own (Ind,  $p = 0.663$ ) or in groups (WG,  $p = 0.61$ ). Students in the STEP-WISE course trend toward spending less time listening (L,  $p = 0.61$ ), an activity that is often considered to be passive rather than active.

### Behaviors within STEP-WISE

Some instructor behaviors in STEP-WISE are highly correlated with one another and with specific student behaviors (Supplementary Figure 1 and Supplementary Tables 6 and 7). For example, in 63% of the intervals in which instructors lectured (Lec), students were answering questions (S\_AnQ); 83% of the intervals that had lecture, also included instructors posing questions (PQ). Thus, the very nature of the lecture was student-centered. Also, students asked questions in different contexts, including when they worked in small groups (WG, 68% of intervals in which students asked questions; Supplementary Table 8), one-on-one conversations (1o1, 55% of intervals), and while they were listening to lecture (L, 42% of intervals). Some behaviors that characterized STEP-WISE lecture included instructors posing questions (PQ), students listening (L), and students answering questions aloud to the whole class (S\_AnQ). Students completing group work structured by worksheets (WG) is highly correlated to instructors moving and guiding throughout the class (MG) and one-on-one/small-group discussion (1o1). Lecture (Lec) and group work (WG) were negatively correlated with each other (Supplementary Figure 1 and Supplementary Tables 6 and 7).

### Key Elements of STEP-WISE

Four elements of the STEP-WISE design were found to be particularly salient for influencing postdocs' adoption of student-centered teaching. We present these in the same order that they appear in the program sequence (Figure 1). Note that the order in which we present the key elements does not reflect the weight of their importance. In the *Discussion*, we consider how these elements interact and support each other, and it is the synergy among the elements that leads to the program's success.

**Key Element 1: Autumn Training Sessions.** The two Autumn training sessions are critical to the success of STEP-WISE. During the first session, postdocs identified a contemporary topic in the basic biological sciences, as exemplified by the titles of their courses, which included: “Why Your Microbiome Matters,” “DNA and Data: What Genetic Testing Can—and Can't—Tell Us,” and “Harnessing the Immune System to Fight Disease” (Supplementary Table 1). In addition, postdocs learned about backward design (Wiggins and McTighe, 2005). Conversations about setting course goals led to a discussion about clearly articulating learning goals that span different orders of cognition (Anderson and Krathwohl, 2000; Crowe *et al.*, 2008). Postdocs reported that engaging in course design was one of the most valuable components of the program (Figure 3).

At the training sessions, postdocs began learning about student-centered, active teaching by experiencing it from the students' perspective (Figure 3): STEP-WISE mentors “basically ran an active classroom in teaching us how to do it. It was really helpful in how can I engage these students a lot more than I realized I needed to.”

In addition to learning about course design and specific teaching practices, these sessions also provide an opportunity



for postdocs to engage with the primary literature supporting different active-learning strategies. For example, postdocs learned about the jigsaw (Aronson and Bridgeman, 1979; Clarke, 1994) by participating in and reading primary research about it (Theobald *et al.*, 2017). One postdoc reflected that:

being able to see how we did ... a jigsaw, discussing some of the papers and how we actually learned about some of the active learning techniques or methods by doing them in class is pretty creative. So, I think those are the things that benefited me the most ... super useful.

Postdocs also engaged with literature on additional research-based practices, including ways to support students' understanding of figures and tables (Round and Campbell, 2013). In their exit survey, participants rated that exposure to active-learning strategies and educational research were highly valued (4.4 and 3.8 respectively; 5 = most valuable; Figure 3).

**Key Element 2: Precourse Meeting with Mentor.** Another key element of STEP-WISE is the structured meeting that the teaching team had with their faculty mentor approximately 1 month before starting their course. This meeting began to establish what would become close and valuable relationships within the team and between the team and mentor (Figure 3). Typically, the meeting began by discussing drafts of the learning goals for the course, as well as for the first few sessions of the seminar. Although the postdocs learned about writing learning goals at a variety of cognitive levels (Anderson and Krathwohl, 2000; Crowe *et al.*, 2008), their first drafts were typically only written at low levels. The mentor worked with the postdocs to revise some of the learning goals to include higher-level work. This discussion led to conversations about how to pare down the content of the course, focusing on depth of learning over breadth, as well as how to integrate across each postdoc's section. Another major component of the meeting was thinking about the final project, especially considering how to distribute aspects of that project throughout the term. The session concluded with talking about logistics, such as visiting the classroom ahead of time, building out the course website on the university's learning management system, practicing saying students' names, and deciding when and where to meet before heading to the classroom for the first class session.

This meeting is also where the mentor ensured that the team had a way to communicate and share files electronically, such as through Google Drive, Dropbox, and Slack. Given time constraints, these means for collaboration were helpful to postdocs, one of whom reported: "We've mostly just discussed things digitally, and it's worked really well."

Postdocs indicated that, while this meeting was essential to the program's success, it did not always feel sufficient: "that [meeting with the mentor] was helpful but I still think it was, like, when we jumped into the first ... class it was like, 'Oh wow. We're really in this.'" Some mentors had a second meeting with their teams before the start of the quarter.

**Key Element 3: Implementing Active Learning with Support.** Postdocs indicated that the most valuable feature of STEP-WISE was the opportunity to gain teaching experience

(Figure 3). For some postdocs, STEP-WISE offered their first teaching experience (Table 1). When postdocs were teaching, they used active-learning strategies often and consistently, minimizing the amount of lecture they included and relying on the strategies emphasized in their Autumn training: think-pair-share, jigsaw, gallery walk, figure facts, and unbiased call.

During class sessions, co-instructors sometimes supported the main teaching postdoc through activities such as writing prompts on the board, helping organize students into jigsaw groups, or, most often, circulating while students worked in small groups. However, co-instructors' roles varied and depended on factors such as their comfort level with the content, clarity of the student task, and interpersonal dynamics within the teaching group.

All of the teams supported their teaching practice with in-class worksheets that served multiple purposes. The worksheets empowered novice teachers to develop their practice and structure their teaching. Postdocs used the worksheets to state learning goals, guide class discussions, structure students' engagement with the content and each other, and, at times, to gradually increase the complexity of tasks.

The worksheets also provided postdocs with opportunities for formative assessment. During in-class small-group activities, postdocs circulated to gauge students' understanding, elicit responses to worksheet prompts, and provide guidance as needed. Additionally, the worksheets were collected at the end of class and typically graded for completion. Some teaching teams also prompted metacognition by including on the worksheet prompts for one thing students learned and one question they still had about the material.

STEP-WISE postdocs routinely engaged students in undertaking authentic scientific practices. Students read and analyzed figures from scientific papers from the primary literature, often using techniques such as the gallery walk and the jigsaw to make sense of what they were studying. Final projects required students to analyze and interpret data, design an experiment, and/or write a proposal for a new research project. The projects were engaging and communicated complex scientific ideas through media such as scientific posters, informational flyers, and podcasts.

Postdocs varied in their success at implementing active-learning strategies. In addition to the success described above, common oversights included not checking in with small groups while circulating; omitting the last component of a gallery walk in which students summarize, synthesize, and share the responses upon returning to their first station; being unorganized about sorting and resorting students into groups for their jigsaw; not being truly random when students were called; and missing the opportunity to build on students' responses after they share; or not repeating students' responses to the rest of the class to ensure that everyone could hear.

**Key Element 4: Debriefs.** The after-class debriefs served as the primary site of learning, the main feature of STEP-WISE, and an element that postdocs found to be "detailed and incredibly valuable" (Figure 3): the "debriefs with our mentor were critically important. This is where most of our learning happened." Debriefs structured self-reflection, team

reflection, and suggestions for improvement from the mentor. Both positive and negative aspects of teaching were discussed.

Postdocs reported that they used mentors' feedback directly in their subsequent practice: "I knew the top three things that I needed to work on. I prioritized that as I was planning my next class." The debriefs also served as a space for collective learning: "The reason that I didn't make some of the same mistakes that were made before was because I recognized it coming up with another person and so I knew how to improve." Collegiality within the teaching teams may have helped foster this collective learning:

[I was] working with a team where we all talked in a very nice and friendly and constructive way about the things that worked and didn't work, I think it helped me improve, like, be critical, realize the things that needed more work on. But, in a really nice, positive, encouraging way and also kind of learning from what [the other postdocs] needed [to] work on as well. Thinking how we could help each other learn, I think that was a process that was also very insightful.

Across all three mentors, debrief sessions roughly followed the same structure. The mentors often did not have expertise in the content of the course and focused instead on pedagogy. They began with open-ended reflection from the main instructor of the class session. The mentor then elicited reflection from the co-instructors, the depth and complexity of which varied among teams. Finally, the mentor built on the postdocs' reflection to provide feedback and advice, sometimes praising instruction or comparing challenges to examples from his or her own teaching:

[The mentor] drew examples from her own teaching experience and preparing for teaching multiple times. Doing so ... shows her expertise, but in a way that highlights how she has learned how to teach, given that most of her examples are when something went awry or that she learned over the years (like stripping down her slides over multiple times teaching same course). (E.R.K. memo of February 3, 2019)

Mentors built on postdocs' reflection, praising the postdocs and providing critical feedback about the positive and negative pedagogical aspects of class, such as timing, logistics and implementation of active-learning strategies; the volume level while speaking to the class or repeating student questions; whether instructors engaged with students while circulating among small groups; how the worksheet and multimedia resources worked; and how postdocs supported students through primary literature.

Debriefs often concluded with the team discussing their plan for the next class session, ways to coordinate grading across instructors, and strategies for scaffolding students' work toward the final project.

## DISCUSSION

Our results indicate that STEP-WISE postdocs quickly adopted scientific teaching despite a lack of prior teaching experience. By their second class session, all postdocs in the 2018–2019 STEP-WISE cohort used student-centered instruction in the first hour of the 2-hour class session, as defined by the COPUS (Smith

*et al.*, 2013). These successes can be tied to the intentional, key elements of STEP-WISE (Table 2) that impacted and informed participants' experiences and that we identified using the guiding framework of design-based research (Cobb *et al.*, 2003; Design-Based Research Collective, 2003). Although the *Results* section presented each key element as a separate component in the design of the apprenticeship, in actuality, the key elements interact with each other to make the program successful.

These findings can inform the design of other programs that provide mentored teaching experiences for scientists. Collectively, these results suggest that high-impact teaching can be learned early in a career with minimal training through an institutionally sustainable program. In the STEP-WISE model, the time cost is distributed among mentors more than postdocs, so the postdocs can teach with minimal interruption to their research. The financial cost of the mentoring is folded into a standard science department's curriculum.

### STEP-WISE Teaching Is Student-Centered and Inclusive

The teaching methods that STEP-WISE postdocs used corresponded to students' engagement in active and inclusive ways of learning. A plethora of evidence demonstrates that classes that prioritize lecture as a mode of learning are far less effective than ones that are student-centered (Deslauriers *et al.*, 2011; Freeman *et al.*, 2014) and that student-centered approaches elevate all students' mastery of content while decreasing disparity among demographic groups (Eddy and Hogan, 2014; Harris *et al.*, 2020; Theobald *et al.*, 2020). Consequently, a number of studies are documenting ways in which instructors use class time (Deslauriers *et al.*, 2011; Eddy and Hogan, 2014; Stains *et al.*, 2018). Strategies that promote inclusive learning include working independently (Figure 2A, Ind), for example, students writing their own ideas and thus engaging with the material by thinking through questions posed before discussion with peers or the whole class (WC, Tanner, 2013). Students can also talk through ideas in small groups (Figure 2B, working in groups, WG; student question, SQ) and among the whole class (Figure 2B, whole-class discussion, WC; student question, SQ; Smith *et al.*, 2013; Eddy *et al.*, 2015; Hora, 2015; Durham *et al.*, 2018). These student practices are orchestrated by instructor behaviors, such as moving through the classroom and guiding discussion (Figure 2A, instructor poses a question, PQ; instructor and student(s) are engaged in a one-on-one or one-on-few discussion, 1o1; instructor is moving and guiding, MG). This type of immediate feedback for students promotes learning (Figure 2A, activity follow-up, FUp; instructor answering a question, AnQ; instructor and student(s) are engaged in a one-on-one or one-on-few discussion, 1o1; Smith *et al.*, 2013; Eddy *et al.*, 2015; Durham *et al.*, 2018).

The interplay among instructor and student behaviors is captured in the correlation matrix (Supplementary Figure 1 and Supplementary Tables 6 and 7). Correlations between instructor behaviors and student behaviors offer some insight into when lecture occurs in a STEP-WISE class. Lecture, when structured appropriately, can be part of a suite of student-centered, effective teaching strategies (Hora, 2015). For example, high correlation exists between when STEP-WISE instructors lecture and when students answer questions (S\_Anq). This pattern suggests that the instructors' lectures are interactive and are to some degree student directed.

### A Sustainable, Beneficial, and Adoptable Program for Early-Career Researchers

We suggest that STEP-WISE is a readily adoptable model for institutions to train postdoctoral scholars in evidence-based instruction of the sort promoted by scientific teaching (Handelsman *et al.*, 2004, 2006; Couch *et al.*, 2015). Its sustainability is part of what makes it powerful. STEP-WISE is a model that is financially stable, because the costs of the program are incorporated into the university's standard curriculum. The training and intensive mentoring are accounted for, because the mentors are instructors of record, and from the university's perspective, the postdocs are guest lecturers.

While STEP-WISE relies on postdocs volunteering their time, the program is sustainable in terms of workload for postdocs in full-time research positions. It is structured to be offered in conjunction with a standard research load. The two 2-hour Autumn training sessions are held in the evenings, and the mentor session is scheduled at a time when the teaching team can meet. Teaching is done either in the morning or the afternoon, so postdocs devote at most half of a day to teaching. The planning and design of the course is shared by the entire team of three, although each postdoc is primarily responsible for teaching in only three class sessions, which minimizes the time they spend preparing for class. Moreover, the program does not charge any registration, tuition, or travel, all of which are common in training programs.

Postdoctoral trainees are exploring their career options (National Academy of Sciences, 2014; Gibbs *et al.*, 2014), and many are deciding whether to pursue careers that are primarily research, primarily teaching, or a combination (Sauermann and Roach, 2012). STEP-WISE functions as a way for postdocs to determine whether they enjoy teaching and to explore how teaching impacts other aspects of their professional interests. Training in pedagogy supports teaching expertise, as well as increasing trainees' ability to conduct (Feldon *et al.*, 2011) and communicate (Shortlidge and Eddy, 2018) research. Additionally, postdocs who have no or little previous teaching experience are able to begin building a portfolio of their teaching using artifacts they developed and employed during their STEP-WISE course that they can use in job applications.

Other cost-effective programs exist that support postdocs who want to learn to teach with evidence-based and inclusive practices. For example, the Postdoc Academy offered through the Center for the Integration of Research, Teaching and Learning (CIRTL Network, 2020a) has massive open online courses (MOOCs) that promote the professional development of postdocs. Accompanying the courses are guides for facilitating small-group, live discussions about the content. Other MOOCs are available under topics such as evidence-based (CIRTL Network, 2020b) and inclusive teaching (EdX, 2021). A challenge with MOOCs, however, is their very low completion rate (Pursel *et al.*, 2016), combined with the lack of the key elements around personal feedback that we have identified as critical to the success of STEP-WISE.

The existence of programs similar to STEP-WISE suggests that this model might be straightforward to replicate. For example, the University of Oregon Science Literacy Program also applied scientific teaching (Handelsman *et al.*, 2004, 2006; Couch *et al.*, 2015) to support the professional development of graduate teaching assistants and undergraduate learning assis-

tants (Vandegrift *et al.*, 2020). Major similarities to STEP-WISE are the emphasis on mentoring by experienced faculty paired with learning to teach by teaching. That program was supported by an external grant and so may be harder than STEP-WISE to sustain. As another example, Emory University has the On Recent Discoveries by Emory Researchers program, which, like STEP-WISE, encourages teams of postdoctoral scholars to participate in training and then collaborate to design a course that shares teaching responsibilities (Sales *et al.*, 2007; Hue *et al.*, 2010), although a detailed overview and evaluation have, to our knowledge, not yet been published (Sales *et al.*, 2007; Hue *et al.*, 2010).

### Interactions among the Key Elements

All of the key design elements (Table 2) of STEP-WISE interact, and the postdocs' excellent teaching (Figure 2) results from the synergy among these four program components. The debriefs within the instructional team—key element 4—succeed because of the momentum built in key elements 1–3. The debriefs rely on the foundational work from the Autumn training (key element 1), the first meeting with the mentor that establishes a precedent for constructive feedback aligned to the Autumn trainings (key element 2), and the real-time implementation and support in the classroom from the co-instructors and the mentor (key element 3).

We asked how STEP-WISE was effective at supporting novice science instructors' use of evidence-based and inclusive pedagogies. We found that the jigsaw is one of the effective and powerful inclusive practices (Theobald *et al.*, 2017) introduced in the Autumn training (key element 1) and referenced when mentors met with their teams to prepare for class (key element 2). Postdocs found this technique especially helpful for teaching the primary literature and when implemented with the aid of a worksheet that presented the structure of the activity (key element 3). Teaching teams discussed how to improve the implementation of the jigsaw and the kinds of questions asked in the worksheet during the debriefs (key element 4).

STEP-WISE training (Ma *et al.*, 2021) emphasized think-pair-share (Tanner, 2013; Cooper *et al.*, 2021), gallery walk (Francek, 2006), unbiased call (Waugh and Andrews, 2020), and quick writes (Tanner, 2013), and the instructors used all of these strategies (key element 1). STEP-WISE postdocs are still at the beginning of their teaching careers by the end of the apprenticeship—after all, some have only taught three class meetings in their entire careers. This finding, combined with general knowledge that it is challenging for any instructor to adopt a new technique with fidelity (Offerdahl *et al.*, 2018), makes it understandable that postdocs' implementation of these techniques could improve. In subsequent iterations of STEP-WISE, we are actively addressing these areas in our debriefs (key element 4).

We explored how the design of STEP-WISE promotes the postdocs' use of evidence-based and inclusive pedagogies. The support that postdocs received while they are implementing active learning (key element 3) appears critical to the success of the apprenticeship. We attribute this result to two major factors. First, the mentor engages the postdocs in a cycle of assessment and feedback (key element 4), an element that Vandegrift *et al.* (2020) also report as incredibly helpful when early-career scientists learn teaching (graduate students and undergraduates, in their case). This process requires deep reflection based on



what students have demonstrated about their learning (Henderson *et al.*, 2011). In STEP-WISE, the postdocs are using the data they collect from in-class activities to determine what students learned in each class session (Handelsman *et al.*, 2006), then they reflect on those data with their mentor during the debriefs and strategize on ways to improve instruction (key element 4). Reflection at this magnitude often works best in community (Henderson *et al.*, 2011), such as the STEP-WISE teaching teams (key elements 2–4). Thus, the second factor that emerges from this design is a tight community of collaborators, both within the teaching team (key elements 2–4) and across the broader cohort (key element 1). Finally, the third major factor leading to the success of STEP-WISE is the fact that training is sustained over time, throughout the length of the apprenticeship. The learning community begins with the Autumn training sessions (key element 1) and extends through the teaching quarter (key elements 2–4).

### Limitations and Future Work

This evaluation of the teaching methods that STEP-WISE postdocs apply relies on the use of the COPUS (Smith *et al.*, 2013), an instrument that, perhaps because of its popularity, has been subject to criticism (Lund *et al.*, 2015; Hora, 2015; McConnell *et al.*, 2021). Critics of the instrument find that it is too coarse (Hora, 2015), despite the fact that it documents 25 behaviors (Lund *et al.*, 2015), and that it does not distinguish among forms of assessment that instructors use (McConnell *et al.*, 2021). For simplicity's sake the COPUS assumes that the behaviors exhibited are implemented as intended, for example, that the group work students take part in is appropriately structured (Theobald *et al.*, 2017; Wilson *et al.*, 2018). We tested this assumption qualitatively with observations (Emerson *et al.*, 2011). Our qualitative data confirmed that group work was structured to increase inclusion. The COPUS also does not assess the cognitive level at which students are working, with the possible exception of coding when students make a prediction. Our qualitative analysis of the questions that instructors posed, as presented on worksheets the postdocs designed for students to complete, included analysis and interpretation of figures and aspects of experimental design, all of which represent high orders of cognition (Anderson and Krathwohl, 2000; Crowe *et al.*, 2008).

In addition to the limitations of the COPUS, our analysis is limited by the use of Stains *et al.*'s (2018) data as a control. This control group was appropriate for us to establish that STEP-WISE postdocs use active learning that is at least on par with the student-centered instructors in that data set. However, future analyses could compare the performance of STEP-WISE postdocs with that of instructors who have received other kinds of professional development, including graduates of the FIRST IV (Ebert-May *et al.*, 2015; Derting *et al.*, 2016; Emery *et al.*, 2020) and IRACDA programs (Rybarczyk *et al.*, 2016; Eisen and Eaton, 2017; (National Institute of General Medical Sciences, 2021).

Seminar courses such as those that STEP-WISE postdocs teach are free from what has been called the “tyranny of content” (Petersen *et al.*, 2020), which can limit the adoption of evidence-based and inclusive teaching practices in other courses. The postdocs have the freedom to limit the content they are teaching to dive deeply into data analyses and elicit the

critical-thinking skills necessary to understand and build from a single research paper. This approach contrasts with teaching core courses, such as introductory courses in biology, chemistry, and physics, in which breadth tends to be prioritized over depth. Evidence-based teaching techniques can take more time than traditional lecture, although that trade-off is justified by the fact that students retain more knowledge and develop more cognitive skills (AAAS, 2011). Moreover, evidence-based teaching strategies continue to improve along with ways to ensure that the necessary content is covered (AAAS, 2011; Petersen *et al.*, 2020).

The short training in STEP-WISE is both a benefit and a drawback. We have presented evidence that this approach is effective at supporting postdocs who use a few, targeted strategies for active and inclusive teaching. Other strategies, such as those detailed by Tanner (2013), however, are not addressed. In fact, in earlier iterations of STEP-WISE, we found anecdotally that introducing too many teaching strategies in the trainings could overwhelm the postdocs. This approach means that curriculum-based strategies, such as classroom-based undergraduate research experiences (Bangera and Brownell, 2014; CUREnet, 2021) and perhaps even case studies (Bonney, 2015) are outside the scope of the program.

Currently, STEP-WISE has a related program for graduate students called STEP-UP (Science Teaching Experience Program for Upcoming PhDs, 2021b) that is being evaluated. As we will discuss in the future, the program differs in structure from STEP-WISE, because the needs of and opportunities for postdocs and graduate students differ. Preliminary data support the conclusion that STEP-UP is also successful at supporting trainees in adopting scientific teaching early in their academic careers, though possibly for different reasons (Klein *et al.*, 2019).

STEP-WISE courses differ from typical university classes, in that student enrollment is capped at 24—and it is the small class sizes that probably explain why STEP-WISE instructors did not use clicker questions (Figure 2, CQ). For those postdocs who become professors who teach, many of their courses may be much larger. Nonetheless, the evidence-based strategies that the postdocs have learned can be applied to large classes (Deslauriers *et al.*, 2011; Eddy and Hogan, 2014), although it may take additional practice to apply them in these contexts. Future studies can explore how graduates of STEP-WISE use the strategies they learned in teaching large classes.

While our results indicate that STEP-WISE is successful over the short term, more data are necessary to determine how this model of teacher development works over the long term. The FIRST IV model led to sustained, high-quality teaching (Emery *et al.*, 2020), but comparable data do not exist for STEP-WISE postdocs. An exciting direction of future research would be to assess whether STEP-WISE graduates also maintain this teaching excellence. Furthermore, while we have recorded the instructors enacting inclusive and evidence-based teaching practices, we do not have corresponding data about student learning or about how students feel and behave about their learning.

### CONCLUSION

STEP-WISE is a sustainable and efficient form of professional development for postdoctoral scholars. During the two quarters of training, postdocs adopt inclusive and evidence-based



teaching methods in their classes, while they each lead instruction in three class sessions. Moreover, postdocs do this teaching while maintaining 100% research positions, and they do not need to travel to participate. This apprenticeship is time-efficient for postdocs, and as a trade-off, the mentors do more work than in other models of professional development. From the perspective of a university and/or school of medicine, the training is sustainable, because mentors are compensated for their time by being the instructors of record for the seminar courses. Moreover, international postdocs can participate, because this sustainable model does not rely on federal funding. As such, STEP-WISE is an easily replicable model for other institutions committed to improving both postdoctoral and undergraduate education, transforming the field for the future.

## ACKNOWLEDGMENTS

We thank Bruce Burgett, William M. Mahoney, John T. Slattery, and Robert Stacey; past, present, and future STEP-WISE postdocs and mentors; students in STEP-WISE courses; and colleagues at UW Seattle, UW Bothell, and University of Puget Sound. STEP-WISE emerged from a program begun by Barbara T. Wakimoto and Merrill B. Hille. STEP-WISE is supported by the UW School of Medicine, the UW College of Arts and Sciences, UW Biology Department, UW Bothell School of Interdisciplinary Arts and Sciences, the UW Bothell School of Science, Technology, Engineering and Mathematics, and the University of Puget Sound. STEP-UP—a complementary program mentioned in the *Discussion*—is based on work supported by the National Science Foundation Innovations in Graduate Education grant no. 1855841.

## REFERENCES

- Alberts, B., Kirschner, M. W., Tilghman, S., & Varmus, H. (2014). Rescuing US biomedical research from its systemic flaws. *Proceedings of the National Academy of Sciences USA*, *111*, 5773–5777. <https://doi.org/10.1073/pnas.1404402111>
- American Association for the Advancement of Science (AAAS). (2011). *Vision and change in undergraduate biology education: A call to action*. Washington, DC. Retrieved October 7, 2021, from <http://visionandchange.org>
- AAAS. (2019). *Levers for change: An assessment of progress on changing STEM instruction* (p. 200). Washington, DC.
- Anderson, L. W., & Krathwohl, D. R. (2000). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives* (Complete ed.). New York, NY: Pearson.
- Aronson, E., & Bridgeman, D. (1979). Jigsaw groups and the desegregated classroom: In pursuit of common goals. *Personality and Social Psychology Bulletin*, *5*(4), 438–446.
- Bangera, G., & Brownell, S. E. (2014). Course-based undergraduate research experiences can make scientific research more inclusive. *CBE—Life Sciences Education*, *13*, 602–606. <https://doi.org/10.1187/cbe.14-06-0099>
- Bonney, K. M. (2015). Case study teaching method improves student performance and perceptions of learning gains. *Journal of Microbiology & Biology Education*, *16*, 21–28.
- CIRTL Network. (2020a). *CIRTL-aligned grants: Postdoc Academy*. Retrieved March 22, 2021, from [www.cirtl.net/sections/44/pages/postdoc-academy](http://www.cirtl.net/sections/44/pages/postdoc-academy)
- CIRTL Network. (2020b). *Course list*. Retrieved March 22, 2021, from [www.cirtl.net/courses/354](http://www.cirtl.net/courses/354)
- Clarke, J. (1994). Pieces of the puzzle: The jigsaw method. In Sharan, S. (Ed.), *Handbook of cooperative learning methods* (pp. 34–50). Westport, CT: Greenwood Press.
- Cobb, P., Confrey, J., diSessa, A., Lehrer, R., & Schauble, L. (2003). Design experiments in educational research. *Educational Researcher*, *32*(1), 9–13. <https://doi.org/10.3102/0013189X032001009>
- Cooper, K. M., Schinske, J. N., & Tanner, K. D. (2021). Reconsidering the share of a think-pair-share: Emerging limitations, alternatives, and opportunities for research. *CBE—Life Sciences Education*, *20*(1), fe1. <https://doi.org/10.1187/cbe.20-08-0200>
- COPUS Analyzer. (2021). *Introduction page*. Retrieved November 6, 2020, from [www.copusprofiles.org](http://www.copusprofiles.org)
- Couch, B. A., Brown, T. L., Schelpat, T. J., Graham, M. J., & Knight, J. K. (2015). Scientific teaching: Defining a taxonomy of observable practices. *CBE—Life Sciences Education*, *14*(1), ar9. <https://doi.org/10.1187/cbe.14-01-0002>
- Crowe, A., Dirks, C., & Wenderoth, M. P. (2008). Biology in Bloom: Implementing Bloom's taxonomy to enhance student learning in biology. *CBE—Life Sciences Education*, *7*(4), 368–381. <https://doi.org/10.1187/cbe.08-05-0024>
- CUREnet. (2021). *Home page*. Retrieved May 28, 2021, from <https://serc.carleton.edu/curennet/index.html>
- Denaro, K., Sato, B., Harlow, A., Aebersold, A., & Verma, M. (2021). Comparison of cluster analysis methodologies for characterization of Classroom Observation Protocol for Undergraduate STEM (COPUS) data. *CBE—Life Sciences Education*, *20*(1), ar3. <https://doi.org/10.1187/cbe.20-04-0077>
- Derting, T. L., Ebert-May, D., Henkel, T. P., Maher, J. M., Arnold, B., & Passmore, H. A. (2016). Assessing faculty professional development in STEM higher education: Sustainability of outcomes. *Science Advances*, *2*(3), e1501422. <https://doi.org/10.1126/sciadv.1501422>
- Design-Based Research Collective. (2003). Design-based research: An emerging paradigm for educational inquiry. *Educational Researcher*, *32*, 5–8.
- Deslauriers, L., Schelew, E., & Wieman, C. (2011). Improved learning in a large-enrollment physics class. *Science*, *332*(6031), 862–864. <https://doi.org/10.1126/science.1201783>
- Dirks, C., Wenderoth, M. P., & Withers, M. (2014). *Assessment in the college science classroom*. New York, NY: W.H. Freeman Scientific Teaching.
- Durham, M. F., Knight, J. K., Bremers, E. K., DeFreece, J. D., Paine, A. R., & Couch, B. A. (2018). Student, instructor, and observer agreement regarding frequencies of scientific teaching practices using the Measurement Instrument for Scientific Teaching-Observable (MISTO). *International Journal of STEM Education*, *5*, 31. <https://doi.org/10.1186/s40594-018-0128>
- Ebert-May, D., Derting, T. L., Henkel, T. P., Middlemis Maher, J., Momsen, J. L., Arnold, B., & Passmore, H. A. (2015). Breaking the cycle: Future faculty begin teaching with learner-centered strategies after professional development. *CBE—Life Sciences Education*, *14*(2), ar22. <https://doi.org/10.1187/cbe.14-12-0222>
- Ebert-May, D., Derting, T. L., Hodder, J., Momsen, J. L., Long, T. M., & Jardeleza, S. E. (2011). What we say is not what we do: Effective evaluation of faculty development programs. *BioScience*, *61*(7), 550–558. <https://doi.org/10.1525/bio.2011.61.7.9>
- Eddy, S. L., Converse, M., & Wenderoth, M. P. (2015). PORTAAL: A classroom observation tool assessing evidence-based teaching practices for active learning in large science, technology, engineering, and mathematics classes. *CBE—Life Sciences Education*, *14*(2), ar23. <https://doi.org/10.1187/cbe.14-06-0095>
- Eddy, S. L., & Hogan, K. A. (2014). Getting under the hood: How and for whom does increasing course structure work? *CBE—Life Sciences Education*, *13*(3), 453–468. <https://doi.org/10.1187/cbe.14-03-0050>
- EdX. (2021). *Inclusive teaching: Supporting all students in the college classroom*. Retrieved March 22, 2021, from [www.edx.org/course/inclusive-teaching-supporting-all-students-in-the](http://www.edx.org/course/inclusive-teaching-supporting-all-students-in-the)
- Eisen, A., & Eaton, D. C. (2017). A model for postdoctoral education that promotes minority and majority success in the biomedical sciences. *CBE—Life Sciences Education*, *16*(4), ar65. <https://doi.org/10.1187/cbe.17-03-0051>
- Emerson, R. M., Fretz, R. I., & Shaw, L. L. (2011). *Writing ethnographic fieldnotes* (2nd ed.). Chicago: University of Chicago Press.
- Emery, N. C., Maher, J. M., & Ebert-May, D. (2020). Early-career faculty practice learner-centered teaching up to 9 years after postdoctoral professional development. *Science Advances*, *6*(25), eaba2091. <https://doi.org/10.1126/sciadv.aba2091>
- Felder, R. M., & Brent, R. (2016). *Teaching and learning STEM: A practical guide*. San Francisco, CA: Jossey-Bass.

- Feldon, D. F., Peugh, J., Timmerman, B. E., Maher, M. A., Hurst, M., Strickland, D., ... & Stiegelmeier, C. (2011). Graduate students' teaching experiences improve their methodological research skills. *Science*, *333*(6045), 1037–1039. <https://doi.org/10.1126/science.1204109>
- Fleet, C. M., Rosser, M. F. N., Zufall, R. A., Pratt, M. C., Feldman, T. S., & Lemons, P. P. (2006). Hiring criteria in biology departments of academic institutions. *BioScience*, *56*(5), 430–436. [https://doi.org/10.1641/0006-3568\(2006\)056\[0430:HCIBDO\]2.0.CO;2](https://doi.org/10.1641/0006-3568(2006)056[0430:HCIBDO]2.0.CO;2)
- Francek, M. (2006). Promoting discussion in the science classroom using gallery walks. *Journal of College Science Teaching*, *36*(1), 27–31.
- Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., & Wenderoth, M. P. (2014). Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the National Academy of Sciences USA*, *111*, 8410–8415. <https://doi.org/10.1073/pnas.1319030111>
- Gibbs, K. D. Jr., McGready, J., Bennett, J. C., & Griffin, K. (2014). Biomedical science Ph.D. career interest patterns by race/ethnicity and gender. *PLoS ONE*, *9*(12), e114736. <https://doi.org/10.1371/journal.pone.0114736>
- Grossman, P., Hammerness, K., & McDonald, M. (2009). Redefining teaching, re-imagining teacher education. *Teachers and Teaching*, *15*(2), 273–289. <https://doi.org/10.1080/13540600902875340>
- Handelsman, J., Chang, A., Ebert-May, D., Beichner, R., Bruns, P., Chang, A., ... & Wood, W. B. (2004). Scientific teaching. *Science*, *304*, 521–522.
- Handelsman, J., Miller, S., & Pfund, C. (2006). *Scientific teaching*. New York, NY: Freeman.
- Harris, R. B., Mack, M. R., Bryant, J., Theobald, E. J., & Freeman, S. (2020). Reducing achievement gaps in undergraduate general chemistry could lift underrepresented students into a "hyperpersistent zone." *Science Advances*, *6*(24), eaaz5687. <https://doi.org/10.1126/sciadv.aaz5687>
- Henderson, C., Beach, A., & Finkelstein, N. (2011). Facilitating change in undergraduate STEM instructional practices: An analytic review of the literature. *Journal of Research in Science Teaching*, *48*(8), 952–984. <https://doi.org/10.1002/tea.20439>
- Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, *6*(2), 65–70.
- Hora, M. T. (2015). Toward a descriptive science of teaching: How the TDOP illuminates the multidimensional nature of active learning in postsecondary classrooms. *Science Education*, *99*(5), 783–818. <https://doi.org/10.1002/sce.21175>
- Hue, G., Sales, J., Comeau, D., Lynn, D. G., & Eisen, A. (2010). The American science pipeline: Sustaining innovation in a time of economic crisis. *CBE—Life Sciences Education*, *9*(4), 431–434. <https://doi.org/10.1187/cbe.09-12-0091>
- Kelsky, K. (2015). *The professor is in: The essential guide to turning your Ph.D. into a job*. New York: NY: Three Rivers Press.
- Klein, E. R., Price, R. M., Gardner, R., & Wiggins, B. (2019). The Science Teaching Experience Program for Upcoming PhDs (STEP-UP): Designing for effective & inclusive instructors. BioTAPS Virtual Conference held on November 19, 2019.
- Lund, T. J., Pilarz, M., Velasco, J. B., Chakraverty, D., Rosploch, K., Undersander, M., & Stains, M. (2015). The best of both worlds: Building on the COPUS and RTOP observation protocols to easily and reliably measure various levels of reformed instructional practice. *CBE—Life Sciences Education*, *14*(2), ar18. <https://doi.org/10.1187/cbe.14-10-0168>
- Ma, E. Y., Freisem, K., Al-Noori, S., Klein, E. R., & Price, R. M. (2021). An active learning workshop to teach active learning strategies. *CourseSource*. <https://doi.org/10.24918/cs.2021.36>
- Mann, H. B., & Whitney, D. R. (1947). On a test of whether one of two random variables is stochastically larger than the other. *Annals of Mathematical Statistics*, *18*, 50–60.
- McConnell, M., Boyer, J., Montplaisir, L. M., Arneson, J. B., Harding, R. L. S., Farlow, B., & Offerdahl, E. G. (2021). Interpret with caution: COPUS instructional styles may not differ in terms of practices that support student learning. *CBE—Life Sciences Education*, *20*(2), ar26. <https://doi.org/10.1187/cbe.20-09-0218>
- McNeal, K. S., Zhong, M., Soltis, N. A., Doukopoulos, L., Johnson, E. T., Courtney, S., ... & Porch, M. (2020). Biosensors show promise as a measure of student engagement in a large introductory biology course. *CBE—Life Sciences Education*, *19*(4), ar50. <https://doi.org/10.1187/cbe.19-08-0158>
- Miles, M. B., Huberman, A. M., & Saldana, J. (2013). Tactics for generating meaning. In *Qualitative data analysis: A methods sourcebook* (3rd ed., pp. 277–293). Thousand Oaks, CA: Sage.
- National Academy of Sciences. (2014). *The postdoctoral experience revisited*. Washington, DC: National Academies Press.
- National Institute of General Medical Sciences. (2021). *Institutional Research and Academic Career Development Awards (IRACDA; K12)*. Retrieved March 22, 2021, from [www.nigms.nih.gov/training/careerdev/pages/twdinstres.aspx](http://www.nigms.nih.gov/training/careerdev/pages/twdinstres.aspx)
- National Research Council. (2012). *Discipline-based education research: Understanding and improving learning in undergraduate science and engineering*. Washington, DC. <https://doi.org/10.17226/13362>
- Nygren, P. (2018). *Every piece counts: Skills I gained from STEP helped my job search in unexpected ways*. Retrieved October 7, 2021, from [http://depts.washington.edu/stepuw/every\\_piece\\_counts](http://depts.washington.edu/stepuw/every_piece_counts)
- Offerdahl, E. G., McConnell, M., & Boyer, J. (2018). Can I have your recipe? Using a fidelity of implementation (FOI) framework to identify the key ingredients of formative assessment for learning. *CBE—Life Sciences Education*, *17*(4), es16. <https://doi.org/10.1187/cbe.18-02-0029>
- Patton, M. Q. (2014). *Qualitative research & evaluation methods: Integrating theory and practice* (4th ed.). Los Angeles, CA: Sage.
- Petersen, C. I., Baepler, P., Beitz, A., Ching, P., Gorman, K. S., Neudauer, C. L., ... & Wingert, D. (2020). The tyranny of content: "Content coverage" as a barrier to evidence-based teaching approaches and ways to overcome it. *CBE—Life Sciences Education*, *19*(2), ar17. <https://doi.org/10.1187/cbe.19-04-0079>
- Pursel, B. K., Zhang, L., Jablolkow, K. W., Choi, G. W., & Velegol, D. (2016). Understanding MOOC students: Motivations and behaviours indicative of MOOC completion: MOOC student motivations and behaviors. *Journal of Computer Assisted Learning*, *32*(3), 202–217. <https://doi.org/10.1111/jcal.12131>
- R Core Team. (2020). *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing. Retrieved February 15, 2021, from [www.R-project.org](http://www.R-project.org)
- Round, J. E., & Campbell, A. M. (2013). Figure facts: Encouraging undergraduates to take a data-centered approach to reading primary literature. *CBE—Life Sciences Education*, *12*(1), 39–46. <https://doi.org/10.1187/cbe.11-07-0057>
- Rybarczyk, B. J., Lerea, L., Lund, P. K., Whittington, D., & Dykstra, L. (2011). Postdoctoral training aligned with the academic professoriate. *BioScience*, *61*, 699–705. <https://doi.org/10.1525/bio.2011.61.9.8>
- Rybarczyk, B. J., Lerea, L., Whittington, D., & Dykstra, L. (2016). Analysis of postdoctoral training outcomes that broaden participation in science careers. *CBE—Life Sciences Education*, *15*(3), ar33. <https://doi.org/10.1187/cbe.16-01-0032>
- Sales, J., Comeau, D., Liddle, K., Perrone, L., Palmer, K., & Lynn, D. (2007). An interdisciplinary, undergraduate science course taught by graduate and postdoctoral teacher-scholars. *Journal of College Science Teaching*, *34*(4), 24–30.
- Sana, F., Forrin, N. D., Sharma, M., Dubljevic, T., Ho, P., Jalil, E., & Kim, J. A. (2020). Optimizing the efficacy of learning objectives through pretests. *CBE—Life Sciences Education*, *19*(3), ar43. <https://doi.org/10.1187/cbe.19-11-0257>
- Sauermann, H., & Roach, M. (2012). Science PhD career preferences: Levels, changes, and advisor encouragement. *PLoS ONE*, *7*, e36307. <https://doi.org/10.1371/journal.pone.0036307>
- Science Teaching Experience Program. (2021a). *Home page*. Retrieved March 22, 2021, from <http://depts.washington.edu/stepuw>
- Science Teaching Experience Program. (2021b). *Science Teaching Experience Program for Upcoming PhDs*. Retrieved October 7, 2021, from <https://depts.washington.edu/stepuw/home/step-up>
- Shortlidge, E. E., & Eddy, S. L. (2018). The trade-off between graduate student research and teaching: A myth? *PLoS ONE*, *13*(6), e0199576. <https://doi.org/10.1371/journal.pone.0199576>
- Smith, M. K., Jones, F. H. M., Gilbert, S. L., & Wieman, C. E. (2013). The Classroom Observation Protocol for Undergraduate STEM (COPUS): A new instrument to characterize university STEM classroom practices. *CBE—Life Sciences Education*, *12*(4), 618–627. <https://doi.org/10.1187/cbe.13-08-0154>

- Stains, M., Harshman, J., Barker, M. K., Chasteen, S. V., Cole, R., DeChenne-Peters, S. E. Jr., ... & Young, A. M. (2018). Anatomy of STEM teaching in North American universities. *Science*, *359*(6383), 1468–1470. <https://doi.org/10.1126/science.aap8892>
- Strauss, A., & Corbin, J. (1990). Memos and diagrams. In *Basics of qualitative research: Grounded theory procedures and techniques* (pp. 197–223). Los Angeles, CA: Sage.
- Tanner, K. D. (2013). Structure matters: Twenty-one teaching strategies to promote student engagement and cultivate classroom equity. *CBE—Life Sciences Education*, *12*(3), 322–331. <https://doi.org/10.1187/cbe.13-06-0115>
- Theobald, E. J., Eddy, S. L., Grunspan, D. Z., Wiggins, B. L., & Crowe, A. J. (2017). Student perception of group dynamics predicts individual performance: Comfort and equity matter. *PLoS ONE*, *12*(7), 1–16. <https://doi.org/10.1371/journal.pone.0181336>
- Theobald, E. J., Hill, M. J., Tran, E., Agrawal, S., Arroyo, E. N., Behling, S., ... & Freeman, S. (2020). Active learning narrows achievement gaps for underrepresented students in undergraduate science, technology, engineering, and math. *Proceedings of the National Academy of Sciences, USA*, *117*(12), 6476–6483. <https://doi.org/10.1073/pnas.1916903117>
- Vandegrift, E. V. H., Barber, N. C., Vitale, A., & Ward, T. (2020). Supporting science graduate teaching assistants and undergraduate learning assistants' teaching professional development. *Transformative Dialogues: Teaching and Learning Journal*, *14*, 60–86.
- Waugh, A. H., & Andrews, T. C. (2020). Diving into the details: Constructing a framework of random call components. *CBE—Life Sciences Education*, *19*(2), ar14. <https://doi.org/10.1187/cbe.19-07-0130>
- Wiggins, G., & McTighe, J. (2005). *Understanding by design* (2nd ed.). Alexandria, VA: Association for Supervision and Curriculum Development.
- Wilcoxon, F. (1945). Individual comparisons by Ranking Methods. *Biometrics Bulletin*, *1*, 80–83.
- Wilson, K. J., & Brame, C. J. (2018). Helping practitioners and researchers identify and use education research literature. *CBE—Life Sciences Education*, *17*(1), fe3. <https://doi.org/10.1187/cbe.17-12-0256>
- Wilson, K. J., Brickman, P., & Brame, C. J. (2018). Group work. *CBE—Life Sciences Education*, *17*(1), fe1. <https://doi.org/10.1187/cbe.17-12-0258>
- Zagallo, P., McCourt, J., Idsardi, R., Smith, M. K., Urban-Lurain, M., Andrews, T. C., ... & Lemons, P. P. (2019). Through the eyes of faculty: Using personas as a tool for learner-centered professional development. *CBE—Life Sciences Education*, *18*(4), ar62. <https://doi.org/10.1187/cbe.19-06-0114>