



# Practice-Based Teacher Education Benefits Graduate Trainees and Their Students Through Inclusive and Active Teaching Methods

Shelley Stromholt<sup>1</sup> · Benjamin Wiggins<sup>2,3</sup>  · Bailey Von der Mehden<sup>4</sup>

Accepted: 20 September 2023 / Published online: 16 October 2023  
© The Author(s) 2023

## Abstract

The next generations of science, technology, engineering, and math (STEM) workers are being trained in college and university classrooms by a workforce of instructors who learn pedagogical practice largely on the job. While inclusive instructional practices and their impacts are increasingly well-studied, this training is difficult to instill within the professional development that most STEM professors receive before teaching their students. The Science Teaching Experience Program for Upcoming PhDs (STEP-UP) at the University of Washington was built to prepare future professors for inclusive excellence by guiding them through the literature in education research and providing them a space to practice active and inclusive teaching techniques. This study of STEP-UP uses a design-based approach to understand graduate trainee and undergraduate perceptions of the most salient aspects and outcomes of the program. Our study found that trainees used opportunities to practice inclusive teaching methods with a cohort of their peers, and crucially that these methods were evident in trainee-taught courses through multiple lines of evidence. STEP-UP-trained instructors used inclusive teaching strategies that helped students to feel socioemotionally supported. This study offers a model program that fosters inclusion and equity in undergraduate STEM classrooms through improving teaching professional development for graduate students.

**Keywords** Graduate · Inclusive · Pedagogical training · Design research · Practice-based

---

Shelley Stromholt and Benjamin Wiggins contributed equally to this work.

---

Extended author information available on the last page of the article

## Introduction

Post-secondary STEM faculty have the opportunity to educate future generations of scientists and technical experts, but they do so typically without explicit training in educational best practices (Auerbach & Andrews, 2018; Stains et al., 2018). Understanding of supported strategies in STEM teaching has improved considerably in past decades (AAAS, 2009; Bennett et al., 2020; Freeman et al., 2014; Handelsman, 2007; NRC, 2003; Schwartz et al., 2016), but the professional development needed to create truly excellent STEM learning remains inconsistent (Bradforth et al., 2015; Mack & Winter, 2018; Seymour & Hunter, 2019; White et al., 2021). The underlying problems training STEM faculty, which include misaligned rewards structures and undervaluation of teaching and learning, persist (Erdmann et al., 2020; Gess-Newsome et al., 2003). This lack of transformative teaching for professors is a major factor in systematic problems like poor retention and inequitable outcomes especially for students from backgrounds that are traditionally underserved (Banilower et al., 2018; Pfeifer et al., 2023; Sithole et al., 2017; Theobald et al., 2020; Zemenick et al., 2022). Exploration of better professional development during STEM instructional careers is ongoing (Bouwma-Gearhart et al., 2019; Callens et al., 2019; Chen et al., 2021; Du et al., 2019; Erdmann et al., 2020; Stephens et al., 2012; Von Vacano et al., 2022), but will only be maximized by innovations for early-career scientists to develop their practice as inclusive educators prior to the start of their faculty roles (Calabrese Barton & Tan, 2019; Dewsbury, 2017). The program described below is intended to help in these efforts.

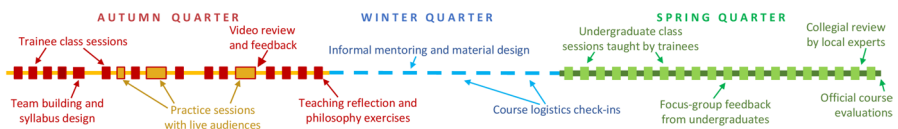
Research on professional development in STEM highlights specific benefits of teaching techniques that facilitate active learning and support students' emotional needs (Beals et al., 2021; Handelsman et al., 2022; Jones & Kahn, 2017; NRC, 2015; Robnett et al., 2018). These complex teaching practices may be best learned in the formative stages of new faculty careers, both for uptake of modern teaching techniques and also for the broadest possible impact on students (Handelsman et al., 2004; Moreira et al., 2019; NRC, 2018). Most STEM teaching faculty build scientific disciplinary habits in graduate school, but few future STEM faculty have opportunities to learn and practice interpersonal teaching techniques (Henderson & Dancy, 2007). A similar lack of professional development around assessment hampers faculty (Brownell & Tanner, 2012), especially as summative assessments drive much of student anxiety and dissatisfaction within STEM education (Drew, 2011; Green et al., 2007; Kuh et al., 2014). Many STEM teaching faculty have only limited exposure to socioemotional teaching practices (Dewsbury, 2017), even as the need for more inclusive teaching in STEM is becoming increasingly apparent (Killpack & Melón, 2016; O'Leary et al., 2020). Exposure to evidence-based instructional practices during graduate training is likely to catalyze ongoing sustainable growth with socioemotional tools that help students to identify with and be resilient in science (Darling-Hammond, 1998; Gasser et al., 2018; Rosenberg et al., 2018; Ruzek et al., 2016; Tharani et al., 2017; Thomas & Zolkoski, 2020; Zeichner, 1987).

Instilling active, inclusive teaching skills is a natural continuation of widespread efforts to improve on early-career capacity in STEM teaching. National

calls for better graduate training programs have echoed this need (Leshner, 2018; Leshner & Scherer, 2019; Love Stowell et al., 2015; Tanner & Allen, 2006), as well as the inherent benefits of pedagogical training for the research field as well (Feldon et al., 2011). Some of these needs may be met by course-based learning about pedagogy (Baumgartner, 2007; Deshler et al., 2015), while other methods include explicit mentorship (Lockwood et al., 2014), community-based practice (Brower et al., 2007; Price et al., 2021), or apprenticeship in K-12 teaching (Ufnar & Shepherd, 2021). Frequently, these efforts are targeted at the level of the graduate teaching assistant (DeChenne et al., 2015; Deshler et al., 2015; Gilmore et al., 2014; Reeves et al., 2016; Rivera, 2018; Schussler et al., 2015; Wan et al., 2020), which is an environment that is well suited for improving TA-led classroom effectiveness and introducing graduate trainees to pedagogical development. In this paper, we focus on practice-based professional development for future instructors (Stroupe et al., 2020) to build on these prior efforts.

The Science Teaching Experience Program for Upcoming PhDs (STEP-UP) at the University of Washington was built to provide a mentored teaching experience for doctoral students in STEM disciplines. STEP-UP is a teaching professional development experience advocated for by research, and was originally proposed and designed by STEM graduate students (Love Stowell et al., 2015). Their goals in creating this training program (in the first cohort of which they participated) were to increase their range of skills that would transfer into research professions, to gain and practice skills that would serve them in mixed research/teaching careers, and to learn more about the teaching methods that would make them more employable in a wider range of industries. During four cohorts, the program featured an in-person autumn course for future professors to (1) learn about active learning strategies and the supporting literature (Bielik et al., 2023; Yik et al., 2022), (2) consider how issues of identity, equity, and justice affect science teaching and learning (Byars-Winston et al., 2015; Flores et al., 2023; Shultz et al., 2022), (3) teach a mock class session for an audience of volunteer undergraduates, and (4) develop their practice and identities as educators through self-reflection and feedback. Over the subsequent two quarters of each cohort, participants then designed and taught an undergraduate course (see Fig. 1). While all autumn trainee courses were taught in-person, trainee-taught courses in the spring of 2020 were taught synchronously online.

Core teaching practices in STEP-UP (Table 1) form the basis for class sessions, during which graduate trainees spend most of their time in direct practice of teaching techniques and methods ranging from microskills (like deliberately pausing after



**Fig. 1** This diagram represents the primary elements of the STEP-UP training program and gives examples (in smaller type) of the activities and practices happening during different quarters of the year-long program. For example, graduate trainees do exercises during class sessions at the end of autumn quarter about their teaching reflections and philosophy

**Table 1** Core teaching practices

Category	Teaching practices in STEP-UP (starting points for practitioners offered as citations in parentheses)
Teaching Strategies	<ul style="list-style-type: none"> <li>• Framing active learning activities (Bean &amp; Melzer, 2021)</li> <li>• Encouraging inclusive student participation (K. D. Tanner, 2013)</li> <li>• Facilitating difficult conversations (Priftanji et al., 2020)</li> <li>• Asking provoking questions (probing and pressing) (Windschitl et al., 2020)</li> <li>• Using public speaking voice within a classroom to direct attention and foster discussion (Snow &amp; O’connor, 2016)</li> <li>• Throwing attention to speakers gently in a way that warmly gives them time to briefly process the question (Jiménez &amp; Orozco, 2021)</li> <li>• Activating passive lecture material using active learning techniques tied to deliberate practice (Williams &amp; O’Dowd, 2021)</li> </ul>
Supporting Students Socioemotionally	<ul style="list-style-type: none"> <li>• Connecting emotionally with students (Reyes et al., 2012)</li> <li>• Communicating to students that instructors have their perceived best interests at heart (Morrison et al., 2021)</li> <li>• Assessing students using methods likely to build trust (Shepard, 2021)</li> </ul>
Assessment	<ul style="list-style-type: none"> <li>• Assessing students fairly and equitably (Siegel et al., 2008)</li> <li>• Assessing students in ways that challenge higher thinking (G. Wiggins, 1998)</li> </ul>
Design + Preparation	<ul style="list-style-type: none"> <li>• Designing a class starting from learning goals (G. P. Wiggins &amp; McTighe, 1998)</li> <li>• Writing an effective learning goal (Sawyer, 2006)</li> <li>• Creating a syllabus (Addy et al., 2021)</li> <li>• Conceptualizing teaching through six theories of education (Sawyer, 2006)</li> </ul>
Reflection	<ul style="list-style-type: none"> <li>• Understanding research data around active learning (Brame, 2016)</li> <li>• Making meaning of student teaching evaluations (Carless &amp; Winstone, 2020)</li> </ul>

a question) all the way up to macroskills (like designing and carrying out an entire 20-minute active learning module). Scaffolding these practice sessions to build on previously-learned elements of teaching and to become progressively more difficult is important, as are frequent opportunities for metacognition around those practice sessions and how this will translate to authentic teaching (Ford & Yore, 2012; Van Es & Sherin, 2002). For clarity hereafter, we refer to participants in this program as graduate trainees when their own learning is central and instructors when highlighting the perspective of their undergraduate students. As a core design principle of the program, STEP-UP intentionally and frequently engages participants in problems of practice around equity and identity in science education, with the broader goal of developing their skills to address some of the inequities found in science classrooms.

## Theoretical Framework

This study of STEP-UP draws on Clarke and Hollingsworth’s (2002) Interconnected Model of Professional Growth (IMPG) in teaching, which describes the specific mechanisms and the concurrent, nonlinear nature of teacher professional growth in four domains: Personal Domain, External Domain, Domain of Practice, and Domain of Consequence. Clark and Hollingsworth developed this model based on several longitudinal studies on teachers enacting new practices and reflecting on both teaching and

student learning, as they became more central participants in a community of practice (Wenger, 1998). This model accounts for teacher learning as both the development of different kinds of teacher knowledge (Shulman, 1987) and as practice through apprenticeship (Lave & Wenger, 1991). This includes accounting for the *situated* nature of learning by documenting the features of a social setting that afford or constrain the practices that make up learning, as teachers move through cycles of enaction and reflection. This model is useful for examining a practice-based teacher education program like STEP-UP because the IMPG accounts for the complex ecology of learning (Scott et al., 2020) that trainees encounter through the multiple entry points and interrelated processes of the year-long course/teaching sequence. Course features such as role-play, introduction to educational research, and collaborating with other trainees to design and teach multiple times provided multiple entry points for trainees to approach teaching. Because we hope this research will inform the administrators of similar programs in the future, we used this model to focus on the specific programmatic features of STEP-UP that supported graduate trainees. This study focuses on the following activities in each domain of the IMPG:

- **External Domain:** Teachers encounter new information or practice, such as seeing new practices modeled and participating in in-service meetings. In STEP-UP, this included immersion and reflection on inclusive, active teaching strategies in the trainee course (autumn). The External Domain influences the Domain of Practice and the Personal Domain, which in turn may influence the Domain of Consequence.
- **Domain of Practice:** Teachers try out new practices through professional experimentation. In STEP-UP, graduate trainees engaged in this domain both through a variety of role playing and practice in the trainee course (autumn) and again when they taught undergraduate courses (spring).
- **Personal Domain:** Teachers develop changes in knowledge, beliefs, and attitudes based on their professional experimentation with new information or practices. In STEP-UP, this includes trainees' perceptions of these changes during and after the program, about their own teaching practice, their identities as teachers, and about student outcomes and student learning.
- **Domain of Consequence (Salient Outcomes):** Professional experimentation with new information or practices, as well as related teacher changes in beliefs and attitudes, result in undergraduate or classroom outcomes. In our study of STEP-UP, we focused especially on how trainees' professional experimentation with skills they developed in STEP-UP led to undergraduate perceptions of emotional well-being and *self-efficacy* (also described as *confidence* in data collection methods using common language) in surveys as learners (Lawson et al., 2007; Trujillo & Tanner, 2014; Zee & Koomen, 2016). Ideally, programmatic features of STEP-UP that support observable trainee outcomes will be identified and described.

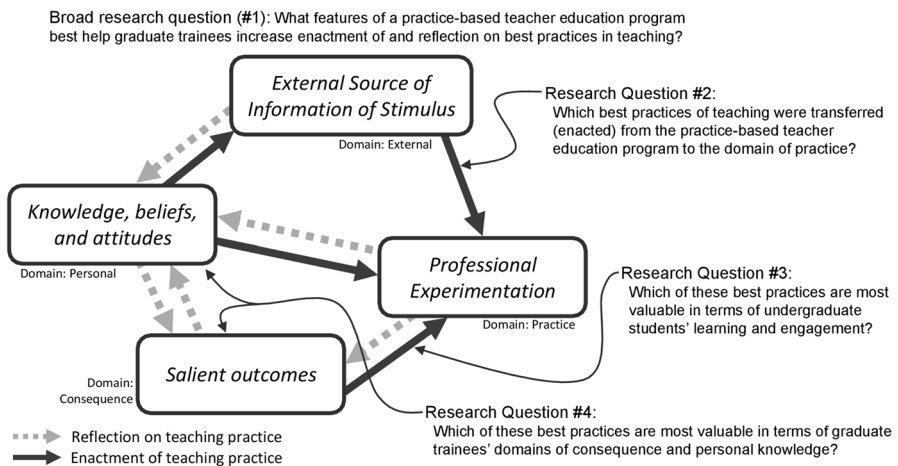
## Current Study

In this study of STEP-UP, we use the IMPG to map the experiences of individuals and groups of graduate trainees in each domain, as they participate in professional learning, apply what they have learned through teaching, reflect on their experience, and plan for future iterations. These cycles of *practice, enactment with students, and reflection* throughout the program are also in line with the approaches of Design-Based Research (DBR), discussed further in the next section.

## Research Questions

This study draws on the IMPG (Clarke & Hollingsworth, 2002) for teachers (see Fig. 2) to address the following questions:

1. What features of a practice-based teacher education program best help graduate trainees increase enactment of and reflection on best practices?
2. What teaching practices were transferred from the practice-based teacher education program to the domain of practice (i.e., enactment)?
3. Which of these practices are most valuable in terms of undergraduate students' learning and engagement?
4. Which of these practices are most valuable in terms of graduate trainees' domains of consequence and personal knowledge?



**Fig. 2** Research questions from this study are mapped onto a close adaptation of Clark & Hollingsworth's (2002) Interconnected Model of Professional Growth (IMPG)

## Methods

To design the program and address our research questions, we took a design-based research (DBR) approach (Anderson & Shattuck, 2012; Barab & Squire, 2004; Scott et al., 2020; The Design-Based Research Collective, 2003). Design-based research situates research in real-world educational settings, focuses on the design and testing of an intervention, and uses mixed methods (Scott et al., 2020). Multiple iterations of a design-enact-study cycle are conducted and the findings in each cycle inform further enactment. This approach is appropriate for this study because the purposes of DBR are to inform practice, contribute to theory, and develop products (Bannan-Ritland, 2003; Barab & Squire, 2004). While DBR studies generally do not yield measurable effect sizes, they provide “rich descriptions of the contexts in which the studies occurred, the challenges of implementation, the development processes involved in creating and administering the interventions, and the design principles that emerged” (Anderson & Shattuck, 2012). DBR studies often include both quantitative and qualitative approaches, which provide opportunities for both graduate trainees and undergraduate participants to share their experiences and give feedback on the program activities. In particular, observations and interviews/focus groups have the potential to provide high internal validity (LeCompte et al., 1993). DBR aligns with this study’s goals to contribute to program development and to the broader goal of supporting teaching professional development experiences for advanced doctoral students in STEM (Connolly et al., 2018).

Throughout the analysis, we used a DBR approach to better identify design principles and instructional approaches that can be generalized to other programs to support mentored teaching experiences for early career STEM faculty. A DBR approach allows us to contribute to the broader field and further specify our theoretical model by providing practical principles for the design of learning environments in the service of specific learning outcomes. DBR allows for a collaborative, mixed methods approach to the analysis in which expertise from the learning sciences can contribute to an interdisciplinary perspective that provides a “rich picture of how the instructional tools and their implementation influence student learning” (Scott et al., 2020).

The first year of data collection was a pilot of the program and focused primarily on iterative refinement of tools. In this study, the tools described, analysis, and results are primarily from data collected in 2021–2022. One important exception is the inclusion of the short Undergraduate Exit Ticket survey; data for this survey from both 2020–2021 and 2021–2022 academic years are reported here. All exit tickets were collected using online surveys, so the online nature of spring 2020 courses did not impact the method of collection of this data.

In addition to quantitative survey data, we also collected qualitative data from surveys, observations, and focus groups to provide rich accounts of graduate trainees’ and undergraduates’ experiences to address our research questions. Below, we describe these methods in more detail by research question, including a summary of data collected in Table 2.

**Table 2** Data collected

	Autumn: preparation course	Spring: trainee teaching	Spring: trainee reflections
2020	<ul style="list-style-type: none"> <li>• <i>Autumn Course Post-survey</i> (6 trainees)</li> </ul>	<ul style="list-style-type: none"> <li>• <i>Course Observations</i> (6 trainees)</li> <li>• <i>Undergraduate Exit Ticket</i> (127 responses)</li> </ul>	<ul style="list-style-type: none"> <li>• <i>Spring Teaching Survey</i> (6 trainees)</li> <li>• <i>End-of-Year Focus Group</i> (6 trainees)</li> </ul>
2022	<ul style="list-style-type: none"> <li>• <i>Autumn Course Post-survey</i> (13 trainees)</li> </ul>	<ul style="list-style-type: none"> <li>• <i>Course Observations</i> (13 trainees)</li> <li>• <i>Undergraduate Exit Ticket</i> (644 responses)</li> <li>• <i>Alumni Interviews</i> (3 2020 trainee alumni)</li> </ul>	<ul style="list-style-type: none"> <li>• <i>Spring Teaching Survey</i> (13 trainees)</li> <li>• <i>End-of-Year Focus Group</i> (12 trainees)</li> </ul>



## Participants

Participants were late-stage STEM PhD students enrolled in a large, public university in the Northwest region of the USA. We collected data from 19 graduate students (trainees) enrolled in the program (6 in 2019–2020; 13 in 2021–2022) and 355 undergraduate students (primarily seniors in STEM majors) enrolled in the trainees' courses. Graduate trainees were in their 4th year of a STEM Ph.D. program on average, and self-identified as 37% students of color, 53% women, 26% neurodivergent, and 26% as LGBTQ+. Participation was voluntary and mechanisms for protection of research participants were controlled by the university's institutional review board (under study #UWIRB00006242). The experience equates to training on the order of roughly 30% FTE (full-time employee) for three quarters and is free for participants.

## Data Collection

**Surveys** To examine features of practice-based teacher education (RQ1), we distributed surveys at the end of the quarter (Supplemental A) to collect responses from graduate trainees about their experiences, including an *Autumn Course Post-survey* and a *Spring Teaching Survey* (Table 3). This survey focused on trainee perceptions of confidence in a set of core teaching practices, based on the teaching development course design. In 2021, we used a retrospective pretest survey model that asked respondents to rate their perceptions at the time of the survey (post) and then to recall their perceptions before the program started (Gouldthorpe & Israel, 2013). This model is particularly useful for addressing response-shift bias, in which respondents shift their frame of reference used to answer pretest and posttest questions, or when they might over- or underestimate their ratings in the pretest based on limited knowledge (Gouldthorpe & Israel, 2013). Since the graduate trainees' prior knowledge of teaching practices was unknown before their autumn course in teaching methods, using a retrospective survey after trainees' experience with the focal teaching practices (post-autumn-course, pre-spring) was the most appropriate method. Additional posttest questions prompted graduate trainees to reflect on their expectations for the course, intent to implement teaching practices, and suggestions for additions/revisions to the course. Finally, we asked trainees to identify which course features they believed best supported their confidence and then to elaborate on their responses in open-ended survey questions.

**Course Observations** To examine the development of teaching practices from professional development through mentored teaching experiences (RQ2), we conducted *Course Observations* of graduate trainees' teaching. Graduate trainees recorded their teaching in at least one class session (~35–65 minutes), which occurred either in-person or online (6 online observation videos in 2020; 13 in-person observation videos in 2022). We created an observation protocol (sample questions in Supplemental A) based on the observable teaching practices and then developed content logs for each video. To develop content logs, we conducted repeated viewings of the

**Table 3** Analysis of data collection tools by research question

Research question	Interconnected model domain	Conceptual focus	Data collection tool	Analysis
Research Question #1: What features of a practice-based teacher education program best help graduate trainees increase enactment of and reflection on best practices?	All domains Focus on Personal Domain: Knowledge, Beliefs, Attitude (self-efficacy)	Graduate trainees learn, implement, and reflect on inclusive, active teaching strategies in a practice-based teacher education program to create valuable learning experiences for undergraduate students. Measures of trainee perceptions of self-efficacy in a set of core teaching practices, based on the course design features.	<i>Autumn Course Post-survey</i> <i>Spring Teaching Survey</i>	Likert-scale items: Descriptive statistics Open-ended items: Content analysis (constant comparative method) and analytical induction techniques
Research Question #2: What teaching practices were transferred from the practice-based teacher education program to the domain of practice (i.e., enactment)?	External Domain and Domain of Practice: Professional Expertise Personal Domain: Knowledge, Beliefs, Attitude (self-efficacy)	Graduate trainees take skills they learned in a practice-based teacher education program and we observe that happening in an undergraduate course.	<i>Course Observations</i>	Content logging
Research Question #3: Which of these practices are most valuable in terms of undergraduate students' learning and engagement?	Domain of Consequence: Salient Outcomes	Trainee experiences with successful teaching strategies; reflection on the program and preparation for the future (e.g., job market).	<i>End-of-Year Focus Group</i>	Content analysis (constant comparative method) and analytical induction techniques focused on the core teaching practices
Research Question #4: Which of these practices are most valuable to graduate trainees?	Domain of practice: Professional experimentation and Personal domain: knowledge, skills, attitude	Undergraduate student perceptions of course equity, inclusivity, opportunities for active participation, instructor impact on student outcomes. Alumni reflections on how their trainee experience contributed to their job search and current enactment in teaching roles.	<i>Undergraduate Exit Ticket</i> <i>Alumni Interview</i>	Open-ended items: Content analysis (constant comparative method) and analytical induction techniques Content analysis (constant comparative method) and analytical induction techniques

videos to document enactments of graduate trainees' observable teaching practices according to the project's theoretical framework and course design (Derry et al., 2010). This strategy allowed for triangulation of the *Course Observations* data with survey, focus group, and exit ticket data (Denzin, 1978).

**End-of-year Focus Group** The End-of-year Focus Group examined graduate trainees' overall experience in STEP-UP (6 in 2020; 12 in 2022). In 2022, we conducted two 75-minute focus groups, with six graduate trainees in each session. Graduate trainees reflected on their aggregated survey data, completed a short journaling activity, discussed their experiences with successful teaching strategies, and reflected broadly on the program and their preparation for the future (e.g., job market).

**Undergraduate Exit Ticket** To examine the value of teaching practices for undergraduate learning and engagement (RQ3), we distributed the *Undergraduate Exit Ticket* (a short survey; sample items in Supplemental A) to those enrolled in STEP-UP trainees' courses in spring 2020 (Likert-scale items) and 2022 (open-ended items). We asked students to complete exit tickets after three class sessions. The prompts in the exit tickets asked students to describe how equitable and inclusive their class was that day, their perceived opportunities for active participation, and how these compared to their typical STEM classes. In 2022, students were also prompted to describe anything their instructor may have done to impact their confidence or well-being as learners. Students earned participation credit for their responses, but could opt for their responses to not be included in the research project. The results presented here include descriptive statistics for both quantitative data from 2020 and coded qualitative data from 2022.

**Alumni Interview** To examine the value of teaching practices in terms of value to early career STEM faculty (RQ4), we used the *Alumni Interview* with three alumni from previous training years as they transitioned into teaching faculty roles. Alumni who had already graduated and continued into teaching careers were recruited for interviews from the prior cohort. Each *Alumni Interview* lasted approximately 30–45 minutes and focused on their current positions/institutions, their experiences on the job market and in a new position, how the training program may have contributed to their job search and current role, and how they are currently applying what they learned with STEP-UP. We recorded each interview, created transcripts, and then analyzed each transcript using the qualitative methods described below.

## Analysis

To analyze data from the tools described above, we used both quantitative and qualitative analysis methods. For Likert-scale items in the surveys and exit tickets, we used descriptive statistics to describe changes in graduate trainees' perceptions. Because of the small sample size for each dataset, we did not conduct tests to determine statistically significant differences, but instead triangulated the quantitative results with qualitative data to strengthen our claims. We report the percentage

of “positive” responses—the sum of the top two positive responses on a five-point scale, typically Agree + Strongly Agree.

To analyze qualitative data in open-ended survey items, focus groups, and exit tickets, we conducted qualitative coding using a combination of constant comparative method (Glaser & Strauss, 1968) and analytic induction (Robinson, 1951). This combination is typical of many approaches to coding because it allows for the testing of hypotheses while allowing for unanticipated outcomes (Merriam & Tisdell, 2015). In this study, the core teaching practices formed a framework for our inquiry as phenomena of interest (analytic induction) and we used the constant comparative method of analysis (both within and between trainee responses) to generate categories that described how graduate trainees and undergraduate students experienced those phenomena. For each dataset, we identified codes based on the teaching practices we expected to arise in each dataset, then iteratively applied and refined both the codes and hypotheses as we analyzed each dataset. The constant comparative analysis also allowed for triangulation between parts of a complex ecology of learning (Scott et al., 2020), in which we compared or cross-checked between datasets (e.g., surveys and observations) and between researchers in order to improve internal validity (Denzin, 1978, 2012).

For focus group and interview data, the research team also employed coding consistency checks by member checking preliminary findings with both the program director and the program participants themselves at different points during the analysis. By involving people who have direct experience and concerns with the program and the data, member checking served to enhance the credibility of the findings (Thomas, 2006).

Finally, informed by work on how inclusivity and emotional well-being are linked to student learning of science (Gannon, 2018; Lee et al., 2021; Moriña, 2017; Ruzek et al., 2016; Yorke et al., 2021), we focused our analysis of the student exit tickets in similar directions. In the analysis of the 2020 undergraduate exit tickets, we focused on how undergraduate students identified courses as “more inclusive compared to other courses.” We included this subset of the 2020 data in our analysis for this study because the wording in these survey items remained appropriate to the latest iteration of data analysis. In the analysis of the 2022 dataset, we first coded a sample of the responses using the qualitative methods described above (constant comparative method and analytic induction) to identify trends in open-ended responses. The analytical focus was on student perceptions of how instructors impacted students’ emotional well-being and self-efficacy as learners (Bandura et al., 1999; Connolly et al., 2018; Gasser et al., 2018; McMullen et al., 2015; Seymour & Hunter, 2019).

### Positionality of the authors

We include the positionality of the authors to situate ourselves within the context of this study. Qualitative coding and interpretation are subjective and researcher experiences may influence coding although every effort to be objective is attempted. One co-first author (S.S.) is an external research and evaluation consultant with an M.Ed. and PhD in science education. She has extensive experience in both design-based

research (DBR) and program evaluation. She co-designed the data collection tools and conducted the bulk of the data analysis. The other co-first author (B.W.) is an affiliate faculty member at the institution studied and a professor at a local community college. He has a MS in molecular science and a PhD in education. He was the director of the STEP-UP program and research PI although he did not himself collect or analyze data in order to minimize bias. The third author (B.V.) is a Ph.D. student focused on discipline-based education research (DBER) while situated in a traditional biology department. She has an undergraduate and master's degree in biology. Her master's degree research was also focused in DBER. She helped to interpret results and helped with manuscript preparation.

## Results

*Research Question #1:* What features of a practice-based teacher education program best help graduate trainees increase enactment of and reflection on best practices?

### Teaching Self-efficacy

Analysis of graduate trainees' survey data demonstrated that STEP-UP contributed to their self-efficacy in all of the focal teaching practices over the course of the program. The greatest gains in self-efficacy overall were as follows (Table 4): activate passive lecture material using active learning techniques (86% increase pre to post), create a syllabus (86% increase), assess students in ways that challenge higher-level thinking (79% increase), design a class session starting from learning goals (79% increase), and strategies for calling on students gently (79% increase).

Graduate trainees reported that their confidence in teaching practices continued to increase during their spring teaching experiences (Table 5). All trainees reported a continued increase in their confidence to frame active learning strategies while teaching, which they attributed to the heavy focus in the autumn course. This makes sense in light of prior work focusing on learning communities, such as in Polizzi et al. (2021) and Zhou et al. (2023). Other frequent self-efficacy gains (83% of trainees) during teaching were in asking provoking questions, assessing students using methods likely to build trust, creating a syllabus, designing a class session starting from learning goals, and writing an effective learning goal. Graduate trainees reported that the experience of teaching multiple times helped them gain self-efficacy, and, importantly, comfort in their teaching (Table 5). This allowed them opportunities to try new things, reflect on the process, and pivot as needed. Fewer trainees reported increases in confidence to productively pull meaning out of student-written teaching evaluations (33% of trainees) (likely due to the fact that they had not received these at the time of the survey) and connecting emotionally with students (42% of trainees). Several graduate trainees reported that they faced challenges in connecting with students for a variety of reasons including COVID-related absences, masking, and teaching only a section of a course. Trainees noted that

opportunities to interact with students on a 1:1 basis made a difference in how they were able to develop skills in this area as well.

Managing a classroom of students allowed me to experiment with different strategies and receive real-time feedback about student understanding and engagement. If one approach was not working, I could pivot and try another and see tangible changes in student responses. I also learned from watching my co-instructors' approach to teaching the same group of students. –*Spring Survey*

### Key Course Features

Graduate trainees identified a number of features unique to the autumn course that supported their confidence in teaching practices in the surveys. The most valuable features were opportunities to practice microskills, role play activities in which graduate trainees tried out scenarios as both instructors and students, collaborating with their colleagues (being part of a cohort of learners), and discussing educational research. For most graduate trainees, these activities represented unique opportunities in their graduate career to practice key teaching skills and iteratively design course syllabi, activities, and assessments. More information can be found in Supplemental B.

The teaching activities we did in the fall were very helpful for developing the teaching strategies listed [here]. My co-teacher and I also did a practice lesson before the course started and that helped us improve our framing of active learning activities. –*Spring Survey*

*Research Question #2: What teaching practices were transferred from the practice-based teacher education program to the domain of practice (i.e., enactment)?*

### Enactment of Teaching Practices

Observations of graduate trainees' spring teaching provided evidence for how trainees enacted strategies they learned in STEP-UP. The observation protocol was designed to focus on the observable STEP-UP practices that were also the focus of the autumn and spring surveys. Graduate trainees demonstrated their use of a number of instructional strategies aimed at supporting student confidence and emotional well-being which are two primary aims of STEP-UP. The primary teaching strategies evident in observations were active learning strategies and asking provoking questions, as well as gently throwing attention to students and giving instructions for active learning activities (Gentle "throws" are those transitions in attention that give the next speaker useful amounts of time and encouragement to speak without delaying the overall conversation; this technique is used widely in broadcast journalism to position speakers for smooth responses). Graduate trainees employed some

**Table 4** Results from retrospective pre-survey taken by trainees before and after STEP-UP autumn course (n=13). Percentages in the first two columns represent total positive responses (confident, 4, and very confident, 5). The final column shows the change from pre to post scores. (Scale: Not at all Confident = 1, to Very Confident = 5.)

Teaching Strategies	Before Autumn Course	After Autumn Course	Pre/Post Change
Activating passive lecture material using active learning techniques tied to deliberate practice	15%	100%	85%
Throwing attention to another speaker gently	8%	85%	77%
Framing active learning activities	23%	92%	69%
Asking provoking questions (both probing and pressing)	15%	77%	62%
Encouraging inclusive student participation	0%	54%	54%
Facilitating difficult conversations	0%	38%	38%
Using public speaking voice within a classroom to direct attention and foster discussion	69%	77%	8%
Supporting Students Socioemotionally			
Assessing students using methods likely to build trust	0%	77%	77%
Communicating to students that instructors have their perceived best interests at heart	15%	77%	62%
Connecting emotionally with students	23%	69%	46%
Assessment			
Assessing students in ways that challenge higher-level thinking	8%	85%	77%
Assessing students fairly and equitably	8%	69%	62%
Design + Preparation			
Creating a syllabus	0%	85%	85%
Designing a class session starting from learning goals	8%	85%	77%
Conceptualizing classroom teaching moves + design choices through educational theory	0%	69%	69%
Writing an effective learning goal	8%	69%	62%
Reflection			
Understanding research data around active learning	23%	92%	69%
Productively pulling meaning out of student-written teaching evaluations	31%	69%	38%

**Table 5** Results from far-post survey taken by graduate trainees after their spring teaching experience. Percentages represent the trainees who indicated that their confidence increased during spring teaching. (Scale: decreased, stayed the same, increased)

Teaching activity in which trainees' confidence increased	%
Framing active learning activities	100%
Asking provoking questions (both probing and pressing)	83%
Assessing students using methods likely to build trust	83%
Creating a syllabus	83%
Designing a class session starting from learning goals	83%
Writing an effective learning goal	83%
Activating passive lecture material using active learning techniques tied to deliberate practice	75%
Encouraging inclusive student participation	75%
Communicating to students that instructors have their perceived interests at heart	75%
Assessing students in ways that challenge higher-level thinking	75%
Conceptualizing classroom teaching moves + design choices through educational theory	75%
Understanding research data around active learning	75%
Facilitating difficult conversations	58%
Assessing students fairly and equitably	58%
Throwing attention to another speaker gently	50%
Using public speaking voice within a classroom to direct attention and foster discussion	50%
Connecting emotionally with students	42%
Productively pulling meaning out of student-written teaching evaluations	33%

strategies for supporting students socioemotionally and effectively assessing students, often through the use of discussion questions.

**Active Learning Strategies** Graduate trainees implemented two categories of *active learning strategies* most frequently in the observations of teaching sessions. The most often used strategy was discussion of an instructor-posed question for students to discuss with a neighbor or small group. This was often followed by a gentle throw to a group or a region of the room to share student responses, though instructors often asked for a response from the whole room and called on student volunteers. Most examples of the strategy were a modified version of think-pair-share (e.g., turn to your neighbor or group, often with think time first). Polls (online or thumbs/hands) were also a common follow-up. The second strategy trainees commonly used was *organized group work*. When graduate trainees used this strategy, students often engaged in analysis and then presented figures, paper summaries, etc. In one class session focused on data visualization, small groups analyzed progressively more complex data visualizations, came up with improvements, and then practiced skills in groups that they would need for a final project.

Trainees frequently supported students by *giving explicit instructions for active learning*. Most often, graduate trainees used this to help students understand what was going to happen next, the format of the activity, and what the expectations for participation were. In some instances, the instructor gave some sense of purpose



(e.g., you'll need to do this on your midterm exam or in the final assignment; this is a debate that scientists in my lab are having, etc.).

**Supporting Students Socioemotionally** In the observations, trainees regularly made efforts to connect with students, often by expressing enthusiasm for material, connecting their professional experiences to the course work, and or making current debates in the field or in their own labs explicit. Graduate trainees also took opportunities to communicate to students that instructors had their perceived interests at heart, primarily by noting material that was complex, reassuring students of what they did not need to know, acknowledging and explaining how instructors were listening to student input, explaining why institutional or class structures were the way they were, and examples of how instructors designed the course to reduce the overall cognitive load for students. Student exit tickets noted that these strategies had a positive impact on student learning.

Graduate trainees posed questions to assess students' understanding in ways that were likely to build trust between instructor and student. Across the observations, trainees frequently made use of "gentle throws" to ask questions to groups or sections of the room. Graduate trainees consistently responded to students' contributions to class in positive ways, encouraging and affirming student responses and regularly revoicing and elaborating on student answers. Importantly, they deftly handled incorrect or partially correct student responses, identifying which parts of a student response were correct and processing the answer out loud with students to break down any partially or wholly incorrect responses in a productive and encouraging way. In several cases, the trainees used a polling app that allowed students to respond anonymously but also see their peers' responses, then engaged the whole class in breaking down the correct answers.

**Questioning** Finally, graduate trainees' questions to students, both pressing and probing, were very common in observations. As the trainees noted in the focus group, they had many opportunities to practice using provoking questions. Trainees used complex questions throughout each session and also specifically identified recall questions or rote questions meant to help everyone go through a process, or set of steps, or to ensure a shared baseline understanding or review from previous course work. This strategy was often used in place of instructor lecture; students were encouraged to contribute and hear from each other.

There were some instances in the observations in which it was apparent the trainee was reflecting on their practice in the moment, evidenced by posing a question to students and then immediately revising the question to be a more meaningful question requiring higher-level thinking. For example, in an interaction with a small group in a breakout room, one trainee helped students decide on which statistical tests would be appropriate for their hypothesis about crow behavior. The trainee went back and forth between asking students a number of probing questions to help them narrow down their ideas to even more specific statistical tests (one-tailed vs two tailed t-test) and giving them "just in time" information to explain or further elaborate on their choices.

You are correct, it is a chi-square test. So yeah... do you know the difference between a — or let me backup — uh...what do you know about the differences between these two tests?

The difference between the quality of the trainee's initial question and their revised question, and specifically the higher-level thinking required to answer the new question, is evidenced by the yes/no answer the first question would yield compared with the thorough explanation a student gave about the goodness of fit.

### **Trainee Reflections on Enactment of Teaching Practices**

In the focus group, graduate trainees described a number of examples of supporting students' emotional well-being and self-efficacy. One area of heavy support they noticed were COVID-related: absences, inability to contribute to group projects due to illness, and gaps in knowledge (presumably because of missed classes or difficulty completing classwork during the height of COVID responses). Trainees noted that their interactions with students in these areas were opportunities to practice supporting students' well-being broadly, but also provided instructors with positive feedback about students' experiences of the course. One trainee described how students who were absent with COVID contacted instructors directly. Students were apologetic and expressed disappointment at missing class because they were enjoying the work so much. In another instance not related to COVID, a student emailed their instructors to let them know that they would miss class due to a panic attack. The graduate trainees noted that the act of sharing this information alone was evidence of student comfort in the course. They went on to have several positive interactions with this student over the quarter: the student went beyond the requirements for missing class, seemed to take more pride in their project, and came to office hours to learn more about getting a degree.

Trainees noted how their use of active learning strategies seemed to have a positive impact on students as well (as will be explored in further detail in the next section). One trainee noted that students who were less vocal during class completed impressive presentations and had positive email exchanges with the instructor about their growth over the quarter. Another trainee noted that walking around to hear small group discussions allowed them to bring some of those conversations back to the large group and that students were much more willing to engage. Finally, a graduate trainee noted that rather than just reading their final presentation slides, as is typical, their undergraduate students incorporated active learning strategies into their final presentations.

Supplemental C and Supplemental D describe highlights from trainee's reflections on their teaching, specifically their stories about a teaching strategy they used successfully. Supplemental C summarizes the main themes across the stories 12 graduate trainees told as they reflected on evidence that their use of the strategy was successful, the course features that supported their success, and their reflection on this experience, including conclusions they are drawing and potential next steps after this success. Supplemental D highlights illustrations, or mini-case studies, of how trainees interpreted and enacted what they learned in the autumn course, reflected

on evidence of student outcomes, and came to conclusions about the results for their teaching and their students.

*Research Question #3: Which of these practices are most valuable in terms of undergraduate students' learning and engagement?*

To understand the undergraduate experience in each course, students completed exit tickets at least three times during the course. Overall, undergraduate reflections provided additional evidence that the strategies graduate trainees employed successfully resulted in a more participatory and inclusive course environment for students. Almost all undergraduates in the spring courses felt that their class provided opportunities for their active participation, which most (72%) described as “more opportunity to participate than other similar courses.” Students identified specific contributing factors such as the focus on active learning, the increased interaction between students in smaller breakout groups, and checking in as a large group.

Almost half of student responses (44%) described their class as “more equitable and inclusive than other courses.” Most of the remaining half of students said that it was similar to other courses. Many students also noted that the department has an equitable and inclusive environment generally. Again, students pointed out features of the course, such as small breakout groups, that helped them feel encouraged to contribute and provided multiple avenues to understand scientific concepts.

When asked about active participation opportunities compared to other science classes, students very consistently said that there were more varied, useful, and meaningful opportunities for active participation in the course taught by a STEP-UP trainee. For example, one student said, “[The instructors] allow for open responses from the class, and I enjoy how casual the class is. I feel like I could mention a topic, and the instructors will even navigate the class in a direction towards a point I made. It makes me feel important.” Another said,

I would say they absolutely provide an opportunity to engage at every point in the material. Not everyone can speak, but there are always several opportunities to consult resources, discuss in small groups, and then present. This cycle, especially incorporating the opportunity to really apply, research, and cross-reference makes for a very engaging class! –Undergraduate Student

Students identified a number of examples of what worked for them in class. Most often students described examples of active learning strategies, including group work, varied participation structures, specific resources, or instructor-provided scaffold such as worksheets and content understanding.

More than half of students in the sample (56%) reported that what they identified as working for them was directly related to something the instructor did (see Fig. 3). For example, most students described how the instructors set up the course (e.g., set up group work, set up clear slides, explained concepts clearly) was a factor that ultimately helped them most in the course. One student said that they “appreciated being given a heads up prior to being called on.” They went on to say that it was related to something the instructor did and “it allowed my group time to collect our thoughts and prepare to answer the questions that was presented to us.” This student further described this practice as having a positive impact on their emotional well-being because it “relieved some anxiety that my group would have had.”

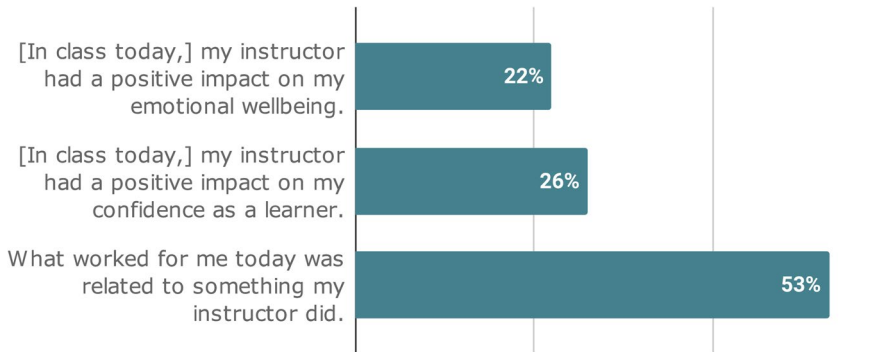
About 22% of students reported that the instructor had a positive impact on their emotional well-being. Similarly, about a quarter of students in the sample (26%) reported that the instructor had a positive impact on their confidence as a learner.

Though not all students report that their instructor had a positive impact on their well-being and/or confidence, those that did were often very emphatic or effusive in their explanation and told important stories about their experiences. For example, a couple of students reported how surprising it was to be referred to by name, or to have a trainee go beyond their expectations to meet an accommodation they needed. Below is a preliminary list of student-identified instructor moves that contributed to their confidence and emotional well-being:

- Acknowledging common stressors, accommodating absences and student concerns; stating explicitly that the material is difficult
- Inviting all kinds of responses and positively receiving right and wrong answers. Stating explicitly that wrong answers are good and welcome
- Taking the time to explain things, wrong answers, Poll Everywhere answers, etc.
- Acknowledging students personally—using their names, making eye contact, responding to individual emails
- Creating a generally positive environment: (words students used) relaxing, calm, supportive, fun, engaging, safe, low pressure, encouraging
- Participation structures: calling on groups or letting people know they'll be called on; providing multiple ways to participate, earn points, make-up work
- Implementing student accommodation for a break for the whole class; including disability statements in the syllabus
- Acknowledging feedback from students
- Supplemental E highlights sample student reflections from the subset of students who reported that the instructor had a positive impact on their well-being and confidence as a learner. These quotes illustrate some of the ways that students were impacted as a result of graduate trainees' enactment of STEP-UP teaching practices.

*Research Question #4: Which of these practices are most valuable to graduate trainees?*

To understand alumni perspectives on their experience with STEP-UP, including which practices seemed most valuable to them as they entered the job market, we conducted interviews with three alumni who have taken teaching-focused positions since their graduation. All three reported that their experience with STEP-UP was valuable in that they learned about teaching strategies, but also started to situate their teaching in a broader framework that gave them credibility and improved their self-efficacy. They reported these gains in relation to their teaching and their job search experience. Alumni appreciated that their work in the program was grounded in educational research and evidence-based strategies, multiple opportunities to rehearse and practice teaching, a cohort model, and opportunities to reflect on their teaching. Instructor supports were crucial. Alumni described how the program mentor modeled socioemotional and academic/professional support for graduate trainees



**Fig. 3** This graph shows the results from several multiple choice survey items answered by undergraduate students of STEP-UP instructors as part of post-session exit tickets. These exit tickets were filled out online ( $n=644$ )

(e.g., named strategies when in use), intentionally built community among the trainees, and supported them during the job search.

STEP-UP informed and influenced their career searches and perceived marketability. According to the alumnus, the program was especially consequential for the job search because the experience gave them language for their work and helped professionalize their experience as educators. They described this professionalization of teaching as a career or job responsibility in contrast to their R1 graduate experience more generally, which they reported consistently devalued teaching. The experience supported alumni's ability to navigate the interview process, as they learned about the different kinds of institutions and job possibilities and how to recognize jobs that would value their approach to teaching. Two alumni said they would not have gotten the job without STEP-UP and the third said that it heavily influenced their application/process. Alumni described how the program mentor advised on application packages, including how to present their teaching, writing teaching statements, facilitating practice teaching sessions for sample lessons, and providing a reference.

In the interviews, alumni described how they are using what they learned directly in their teaching every day, especially in how they plan, reflect, and adjust their teaching. They are teaching with intentionality, continually assessing where they are in their teaching in relation to a broader framework through regular reflection on what good interaction with students looks like and asking themselves where there are opportunities in their teaching. They described meaningful examples of strategies they have tried out, reflected on the results (sometimes with data), and made adjustments with a goal in mind. One alumnus described requiring students to come to office hours before the first test. The alumnus made this decision after several students who failed the first test came to office hours and then drastically improved their subsequent test scores.

The first two semesters' students would appear that they were doing fine in the class and just bomb the first exam...so I was thinking, what could I do to increase the chance that they're going to come see me before that first exam,

that they're gonna participate in class before that exam. And that's when I made those mandatory meetings before the first exam and a lot of them think it's [about] me getting to know them and them getting to know me... I want them to feel comfortable with me because sometimes you think that the professor is someone you can't talk to.

This alumnus reflected on student data and student's experience in office hours, recognized the impact of student comfort and relationship building on student test scores, and adapted their course requirements to better support students.

Two examples of how alumni described teaching with intentionality that was informed by STEP-UP were highlighted in their interviews; in their teaching, they described working toward a specific goal/construct and constructing a set of overlapping strategies that support that goal. These two examples highlight how the alumni are focused on empathizing with and supporting the *students' experience*. In the first example, alumni described employing strategies for setting up their class to encourage engagement in the moment and helping students learn how to learn in a safe space. These strategies included using online polling apps for students to register questions anonymously and recording class sessions with explicit instruction for how to rewatch the video and try out the content after class. In another example, alumni described strategies for giving regular feedback to students as part of students' building expertise. These strategies included using both formative and summative assessment, limiting homework so that there is enough time to give meaningful feedback, and providing a relatively rigorous process for allowing revisions, including student reflection on the process itself.

## Discussion

The goal of this research was to improve professional development for early-career faculty (i.e., STEM doctoral students) by describing the relationships between features in a teaching course, graduate trainee practices in the classroom, and outcomes for undergraduates in those classrooms. STEP-UP implemented an evidence-based theory of change by supporting graduate trainees to engage undergraduate students by enacting active-learning techniques and well-being in their courses. STEP-UP provided opportunities for graduate trainees to practice evidence-based, culturally responsive teaching strategies in a cohort of peers with meaningful feedback (External Domain). New knowledge about teaching strategies and opportunities for comfortable, low stress practice in the External Domain encouraged professional experimentation during training and while teaching (Domain of Practice). Similarly, positive teaching perspectives and evidence-based strategies in the External Domain influenced changes in the attitudes and beliefs held by graduate trainees about teaching (Personal Domain). Changes in knowledge, beliefs, and attitudes around self-efficacy and experimentation (Personal Domain) led to enactment and reflection of teaching practices (Domain of Practice). The combined changes in the External, Personal, and Domain of Practice resulted in undergraduate student reports of engagement and well-being, appreciation of supportive classroom climate, connections

with peers, and participation structures that students found to be inclusive in graduate trainee courses (Domain of Consequence).

### **Increased Teaching Self-efficacy**

In this study, we aimed to describe the features of a practice-based teacher development course that helped graduate student trainees increase enactment of and reflection on best practices in teaching (Research Question #1). The most valuable features of STEP-UP were opportunities to practice teaching strategies with peers and mentors and discuss teaching strategies that were new to them. Through these activities and practice, we found that graduate trainees' teaching self-efficacy increased and that trainees with increases in self-efficacy went on to practice reflection and enactment of teaching strategies. This result is evidence for social cognitive theory which states that those with higher teaching self-efficacy perform better at teaching (Bandura et al., 1999). While we did not measure teaching performance of the graduate trainees, the mere fact that they practiced reflection and enactment indicates their growth as teachers. Other data sources in this study also reveal that their teaching practices were recognized and appreciated by the undergraduate students they taught.

In the spring focus groups, trainees reported that STEP-UP increased their confidence in teaching, resulting in personal pride that encouraged graduate trainees to experiment further with teaching practices. For some trainees, the experience reinforced their love of teaching, while for others it provided experience that allowed them to discern if they would like to teach more and at what level. Several graduate trainees reported these changes amidst a research supervisory environment that was unsupportive of or antagonistic to their teaching, yet both trainees and alumni continue to recognize the value of their efforts and growth. Although graduate trainees in STEP-UP increased their teaching self-efficacy, literature suggests that departmental culture regarding teaching can be a significant factor in graduate students' perceptions of teaching and furthermore, a poor departmental culture surrounding teaching can lead to a lack of teaching self-efficacy and ultimately poor teaching performance (Burke & Hutchins, 2007; DeChenne et al., 2015). Instilling self-efficacy throughout the program, both by using a rigorous practice-based training program and by providing opportunities for well-supported experimentation, should be seen as a key aspect of future efforts especially for those situated in a campus or departmental culture that is unsupportive of teaching efforts.

### **Active Learning and Supporting Student Well-being**

We intended to describe teaching practices that were transferred from the practice-based teacher education program STEP-UP to the domain of practice (Research Question #2). Across data collected for this study, we found that graduate trainees were able to implement two teaching practices most explicitly: supporting students

socioemotionally and using active learning techniques especially through Socratic questioning practices.

Trainees were provided numerous practice opportunities during STEP-UP for teaching methods from microskills (statements of support and connection, for example) to macroskills (design of assignments with structures that put students in positive positions for collaboration). When observed, graduate trainees regularly made efforts to connect with undergraduate students through supportive actions and enthusiasm. For example, graduate trainees were observed reflecting and discussing their professional successes and failures. Ovid et al. (2021) found that one way to create a more inclusive classroom environment was through incorporating positive non-content talk. Although that study found a low percentage (14%) of students remembered when instructors shared personal stories, those stories were viewed positively by undergraduate students. In addition, sharing stories that humanize, create relevancy, and increase engagement for students may motivate them in the classroom (DeSurra & Church, 1994; Dewsbury & Brame, 2019; Freeman et al., 2007; Stolk et al., 2021; Trujillo & Tanner, 2014).

Active learning practices ranged from microskills (giving students a stated purpose for an activity) to medium-scale (using questioning strategies to probe or press student understanding) to macroskills (building classroom activities that provided opportunities for active-learning). The most common way that trainees enacted active learning practices was through questioning techniques. While formal assessment was not typically visible in the observations, we believe that the presence of questioning practices has the potential to lead to equitable and high-quality formative assessment of student learning (Morris et al., 2021). Taking a broad stance on what counts as assessment (as an equity move) to include instructor-posed questions provides evidence across the trainee observations of fair and equitable assessment through the following practices: asking questions that encourage higher-level thinking, gentle throws to groups, giving think time, random call, asking probing questions, revoicing student responses, giving positive feedback, reframing incorrect/partially correct responses. Additionally, there were many examples of trainee questions, small group discussions, and assignments that involved higher-level thinking. Throughout observations and discussions, trainees were clearly practicing and reflecting on occasions in which they positively handled incorrect student responses during active learning, which helped undergraduates to process their answers in an encouraging way.

In the exit ticket survey responses, undergraduate students found these trainee moves to be particularly compelling and reported that these practices created a collaborative and safe learning environment. This enabled students to take risks and contribute meaningfully in class without fear of being subjected to negative feedback or ostracization. Studies have shown that active learning can benefit student learning and performance (Freeman et al., 2014; Theobald et al., 2020). Although there is debate about who active learning helps most, and which students may be left behind (Cooper & Brownell, 2016; England et al., 2017; Gin et al., 2020), our study indicates that graduate trainees in STEP-UP



have prioritized student well-being while guiding active learning. We posit that this combination is essential for undergraduate students to feel comfortable in an active learning environment which asks them to be vulnerable by offering their ideas to their peers in the class.

### **In-class Participation and Emotional Support from Instructor**

We intended to describe which of these practices were most valuable in terms of undergraduate students' learning and engagement (Research Question #3). Across several types of undergraduate-facing data, the two most salient outcomes were that undergraduates had many opportunities to participate meaningfully in class and that those undergraduates felt supported emotionally by their instructors. While comparing perceptions to other courses was beyond the scope of this study, undergraduates consistently reported that they felt able to participate comfortably at frequencies that seem to indicate success of graduate trainee teaching methods surpassing undergraduates' perceived norms. While it may not be surprising to see supportive teaching in a positive classroom environment where graduate trainees were explicitly oriented and trained towards these practices (Rozhenkova et al., 2023), it is notable that these outcomes are often associated with the most experienced professors (Ambrose, 2010; Stronge, 2013; Whitaker, 2020; Wilson, 2004). Undergraduates participated and were routinely given opportunities to analyze their own understanding of the material presented to them (Winne & Azevedo, 2014). Metacognitive practices have been correlated with improved learning outcomes and retention in STEM especially for students historically and currently excluded in academia (Hansen et al., 2023; Knight et al., 2022; McKinney et al., 2021; Seymour, 1995). Furthermore, the intentionality of emotional support displayed by graduate trainees is likely to accentuate undergraduates' use of active learning opportunities as they perceive safety in doing so, as well as to help undergraduates identify themselves as people likely to succeed in STEM.

This research was conducted in part during the COVID pandemic which started in 2020. While the method of instruction for graduate trainees did not change (but took a hiatus in the depth of the pandemic), one cohort of instructors was forced to quickly switch to online teaching. While not enough data for well-supported conclusions was collected about this cohort alone, the positive undergraduate and instructor feedback which matched other cohorts suggests that trainees were able to apply skills and perspectives to their new teaching environment. Anecdotally, they may have been faster and more adept at shifting curriculum, perhaps due to the advantage of habits and minds and lenses on education that focused on the student experience in their recent training course. More research will be needed to delve into the challenges and opportunities for trainee learning and practice.

For those considering creating a teaching professional development program for graduate students in the life sciences, this study offers a roadmap to design features that are likely to positively support both graduate students and undergraduate students. Programs that provide opportunities for authentic teaching experimentation after rigorous practice in active, supportive teaching methods are likely to see

similar signs of success in the experiences of their graduate trainees and undergraduates. We suggest that basing the design of such programs in practice-based teacher development (Stroupe et al., 2020), creating an environment of intentionally instilling self-efficacy for trainees, and explicitly modeling and providing opportunities to practice teaching moves that demonstrate socioemotional support are key features in this arena.

## Limitations

Tempering these positive findings is the knowledge that our research did not assess individual undergraduate students in their own personal contexts (such as demographics like race/ethnicity and gender); this kind of deeper case-based qualitative work will be needed to understand examples of particular teaching moves that graduate trainees might use to explicitly catalyze undergraduate progress for specific student groups. We intended to assess which of these practices are most valuable for trainees as they become early-career faculty. Theoretically, practices perceived as successful should translate to both their Domains of Consequence (career outcomes, like positions achieved) and Personal Domains (knowledge and beliefs that they take into those careers, like attitudes towards future development in their own teaching). Our limited assessment with STEP-UP alumni gives initial clues that graduate trainees are becoming student-centered teachers who are likely to both advocate for students and to continue improving in their teaching. However, it is important to emphasize that all alumni that chose to respond for interviews are all employed by teaching institutions; hence, these alumni may value teaching more than those that did not respond. Much of the design and research of this program occurred during a global pandemic, which changed instructional modes and stressed participants in uncontrollable ways. Importantly, alumni report on teaching moves and design ideas that have rapidly progressed beyond those that they explicitly practiced as part of STEP-UP, indicating likely continued cycles of enactment and reflection that are likely to drive future interconnected professional growth.

We would also like to note that this study took place at one institution and so the results cannot necessarily be generalized to other institutions. The impacts and benefits observed by this research are unlikely to remain for less-robust experiences for which graduate trainees spend less time learning and practicing. We hope that our recommendations for developing a teaching professional development program could be used to help jumpstart other programs with similar goals of teaching graduate students evidence-based teaching strategies and promoting inclusion in the classroom.

## Conclusion

The problem of developing teaching experts from STEM graduate students is complex, important, and continues to be worthy of investment. Tackling it requires developing dual expertise in both science and teaching strategies, but the potential benefits for science are profound. This study examines how graduate students can be supported to develop their teaching practice and self-efficacy in ways that positively affect the students that they teach. The findings from this study, which were prevalent across different types of data, are aligned for the goals most often discussed around improving college STEM education. Additionally, we provide a model for how a Design-Based Research approach might inform other teaching professional development experiences for science graduate students, especially in the interest of scaling best practices. We present this work in hopes of contributing to the design of future programs that improve on this model of STEM teaching development.

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s41979-023-00109-6>.

**Acknowledgements** The STEP-UP program was originally conceived and driven by three fantastic graduate students: Michelle Hays, Sam Cutler, and Sophie Archambeault. Elaine Klein was a crucially important contributor to the research development and implementation of this program. We thank Nina Salama, Bill Mahoney, Celeste Berg, Rich Gardner, and Rebecca Price for working on STEP-UP over the years. We appreciate Teddy Maley and Kelley Hennessey for their editing expertise, and Carolyn Cohen for the spiral journal technique. The dozens of talented STEM instructors who have come through the STEP-UP program continue to inspire us as they work with students and colleagues in the ways we'd always hoped for their generation.

**Funding** This work was supported by the Innovations in Graduate Education program under NSF#1855841.

**Data Availability** Analysis of participant data is available in Supplementary files. Data supporting these files are not publicly available in order to protect patient privacy.

## Declarations

**Competing Interests** The authors declare that they have no competing interests.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- AAAS. (2009) Vision and change in undergraduate biology education. [https://www.researchgate.net/publication/248290185\\_Vision\\_and\\_Change\\_in\\_Undergraduate\\_Biology\\_Education\\_A\\_Call\\_to\\_Action](https://www.researchgate.net/publication/248290185_Vision_and_Change_in_Undergraduate_Biology_Education_A_Call_to_Action). Accessed 25 Aug 2023
- Addy, T. M., Dube, D., Mitchell, K. A., & SoRelle, M. (2021). *What inclusive instructors do: Principles and practices for excellence in college teaching*. Stylus Publishing, LLC.
- Ambrose, S. A. (Ed.). (2010). *How learning works: Seven research-based principles for smart teaching* ((1st ed. ed.). Jossey-Bass.
- Anderson, T., & Shattuck, J. (2012). Design-based research: A decade of progress in education research? *Educational Researcher*, 41(1), 16–25.
- Auerbach, A. J. J., & Andrews, T. C. (2018). Pedagogical knowledge for active-learning instruction in large undergraduate biology courses: A large-scale qualitative investigation of instructor thinking. *International Journal of STEM Education*, 5(1), 1–25.
- Bandura, A., Freeman, W. H., & Lightsey, R. (1999). *Self-efficacy: The exercise of control*. Springer.
- Banilower, E. R., Smith, P. S., Malzahn, K. A., Plumley, C. L., Gordon, E. M., & Hayes, M. L. (2018). *Report of the 2018 NSSME+*. Horizon Research, Inc. <http://horizon-research.com/NSSME/2018-nssme/research-products/reports/technical-report>
- Bannan-Ritland, B. (2003). The role of design in research: The integrative learning design framework. *Educational Researcher*, 32(1), 21–24.
- Barab, S., & Squire, K. (2004). Design-based research: Putting a stake in the ground. *The Journal of the Learning Sciences*, 13(1), 1–14.
- Baumgartner, E. (2007). A professional development teaching course for science graduate students. *Journal of College Science Teaching*, 36(6), 16–21.
- Beals, R., Zimny, S., Lyons, F., & Bobbitt, O. (2021). Activating social capital: How peer and socio-emotional mentoring facilitate resilience and success for community college students. *Frontiers in Education*, 6, 1–16.
- Bean, J. C., & Melzer, D. (2021). *Engaging ideas: The professor's guide to integrating writing, critical thinking, and active learning in the classroom*. John Wiley & Sons.
- Bennett, J., Lattuca, L., Redd, K., & York, T. (2020). Strengthening pathways to faculty careers in STEM: Recommendations for systemic change to support underrepresented groups. In *Lessons from the APLU INCLUDES Project*. Association of Public and Land-Grant Universities.
- Bielik, T., Delen, I., Krell, M., & Assaraf, O. B. Z. (2023). Characterising the literature on the teaching and learning of system thinking and complexity in STEM education: A bibliometric analysis and research synthesis. *Journal for STEM Education Research*, 6(2), 199–231. <https://doi.org/10.1007/s41979-023-00087-9>
- Bouwma-Gearhart, J., Lenz, A., & Ivanovitch, J. (2019). The interplay of postsecondary science educators' problems of practice and competencies: Informing better intervention designs. *Journal of Biological Education*, 53(4), 365–377.
- Bradforth, S. E., Miller, E. R., Dichtel, W. R., Leibovich, A. K., Feig, A. L., Martin, J. D., Bjorkman, K. S., Schultz, Z. D., & Smith, T. L. (2015). University learning: Improve undergraduate science education. *Nature*, 523(7560). <https://doi.org/10.1038/523282a>
- Brame, C. (2016). *Active learning*. Vanderbilt University Center for Teaching.
- Brower, A. M., Carlson-Dakes, C. G., & Barger, S. S. (2007). A learning community model of graduate student professional development for teaching excellence in Higher Education. *CIRTL Publication*, 1–29.
- Brownell, S. E., & Tanner, K. D. (2012). Barriers to faculty pedagogical change: Lack of training, time, incentives, and...Tensions with professional identity? *CBE—Life Sciences Education*, 11, 339–346. <https://doi.org/10.1187/cbe.12-09-0163>
- Burke, L. A., & Hutchins, H. M. (2007). Training transfer: An integrative literature review. *Human Resource Development Review*, 6(3), 263–296.
- Byars-Winston, A. M., Branchaw, J., Pfund, C., Leverett, P., & Newton, J. (2015). Culturally diverse undergraduate researchers' academic outcomes and perceptions of their research mentoring relationships. *International Journal of Science Education*, 37(15), Article 15. <https://doi.org/10.1080/09500693.2015.1085133>
- Calabrese Barton, A., & Tan, E. (2019). Designing for rightful presence in STEM: The role of making present practices. *Journal of the Learning Sciences*, 28(4–5), 616–658.

- Callens, M. V., Kelter, P., Motschenbacher, J., Nyachwaya, J., Ladbury, J. L., & Semanko, A. M. (2019). Developing and implementing a campus-wide professional development program. *Journal of College Science Teaching*, 49(2), 68–75.
- Carless, D., & Winstone, N. (2020). Teacher feedback literacy and its interplay with student feedback literacy. *Teaching in Higher Education*. <https://doi.org/10.1080/13562517.2020.1782372>
- Chen, S., Binning, K. R., Manke, K. J., Brady, S. T., McGreevy, E. M., Betancur, L., Limeri, L. B., & Kaufmann, N. (2021). Am I a science person? A strong science identity bolsters minority students' sense of belonging and performance in college. *Personality and Social Psychology Bulletin*, 47(4), 593–606.
- Clarke, D., & Hollingsworth, H. (2002). Elaborating a model of teacher professional growth. *Teaching and Teacher Education*, 18(8), 947–967.
- Connolly, M. R., Lee, Y.-G., & Savoy, J. N. (2018). The effects of doctoral teaching development on early-career STEM scholars' college teaching self-efficacy. *CBE—Life Sciences Education*, 17(1), ar14.
- Cooper, K. M., & Brownell, S. E. (2016). Coming out in class: Challenges and benefits of active learning in a biology classroom for LGBTQIA students. *CBE—Life Sciences Education*, 15(3), ar37.
- Darling-Hammond, L. (1998). Teacher learning that supports student learning. *Educational Leadership*, 55(5), 6–11.
- DeChenne, S. E., Koziol, N., Needham, M., & Enochs, L. (2015). Modeling sources of teaching self-efficacy for science, technology, engineering, and mathematics graduate teaching assistants. *CBE—Life Sciences Education*, 14(3), ar32. <https://doi.org/10.1187/cbe.14-09-0153>
- Denzin, N. K. (1978). Triangulation: A case for methodological evaluation and combination. *Sociological Methods*, 339–357.
- Denzin, N. K. (2012). Triangulation 2.0. *Journal of Mixed Methods Research*, 6(2), 80–88.
- Derry, S. J., Pea, R. D., Barron, B., Engle, R. A., Erickson, F., Goldman, R., Hall, R., Koschmann, T., Lemke, J. L., & Sherin, M. G. (2010). Conducting video research in the learning sciences: Guidance on selection, analysis, technology, and ethics. *The Journal of the Learning Sciences*, 19(1), 3–53.
- Deshler, J. M., Hauk, S., & Speer, N. (2015). Professional development in teaching for mathematics graduate students. *Notices of the AMS*, 62(6), 638–643.
- DeSurra, C. J., & Church, K. A. (1994). Unlocking the classroom closet: Privileging the marginalized voices of gay/lesbian college students. In *Paper presented at the Annual Meeting of the Speech Communication Association*.
- Dewsbury, B., & Brame, C. J. (2019). Inclusive teaching. *CBE—Life Sciences Education*, 18(2), 2. <https://doi.org/10.1187/cbe.19-01-0021>
- Dewsbury, B. M. (2017). On faculty development of STEM inclusive teaching practices. *FEMS Microbiology Letters*, 364(18), 18. <https://doi.org/10.1093/femsle/fnx179>
- Drew, C. (2011). *Why science majors change their minds (it's just so darn hard)*. The New York Times Company.
- Du, W., Liu, D., Johnson, C. C., Sondergeld, T. A., Bolshakova, V. L., & Moore, T. J. (2019). The impact of integrated STEM professional development on teacher quality. *School Science and Mathematics*, 119(2), 105–114.
- England, B. J., Brigati, J. R., & Schussler, E. E. (2017). Student anxiety in introductory biology classrooms: Perceptions about active learning and persistence in the major. *PLoS One*, 12(8), e0182506.
- Erdmann, R., Miller, K., & Stains, M. (2020). Exploring STEM postsecondary instructors' accounts of instructional planning and revisions. *International Journal of STEM Education*, 7, 1–17.
- Feldon, D. F., Peugh, J., Timmerman, B. E., Maher, M. A., Hurst, M., Strickland, D., Gilmore, J. A., & Stiegelmeier, C. (2011). Graduate students' teaching experiences improve their methodological research skills. *Science*, 333(6045), 1037–1039.
- Flores, G. M., Bañuelos, M., & Harris, P. R. (2023). “What are you doing here?”: Examining minoritized undergraduate student experiences in STEM at a minority serving institution. *Journal for STEM Education Research*. <https://doi.org/10.1007/s41979-023-00103-y>
- Ford, C. L., & Yore, L. D. (2012). Toward convergence of critical thinking, metacognition, and reflection: Illustrations from natural and social sciences, teacher education, and classroom practice. *Metacognition in Science Education*, 251–271. [https://doi.org/10.1007/978-94-007-2132-6\\_11](https://doi.org/10.1007/978-94-007-2132-6_11)

- 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*, *111*, 8410–8415.
- Freeman, T. M., Anderman, L. H., & Jensen, J. M. (2007). Sense of belonging in college freshmen at the classroom and campus levels. *The Journal of Experimental Education*, *75*(3), 203–220.
- Gannon, K. (2018). *The case for inclusive teaching*. The Chronicle of Higher Education Inc.
- Gasser, L., Grütter, J., Buholzer, A., & Wettstein, A. (2018). Emotionally supportive classroom interactions and students' perceptions of their teachers as caring and just. *Learning and Instruction*, *54*, 82–92. <https://doi.org/10.1016/j.learninstruc.2017.08.003>
- Gess-Newsome, J., Southerland, S. A., Johnston, A., & Woodbury, S. (2003). Educational reform, personal practical theories, and dissatisfaction: The anatomy of change in college science teaching. *American Educational Research Journal*, *40*(3), 731–767.
- Gilmore, J., Maher, M. A., Feldon, D. F., & Timmerman, B. (2014). Exploration of factors related to the development of science, technology, engineering, and mathematics graduate teaching assistants' teaching orientations. *Studies in Higher Education*, *39*(10), 1910–1928.
- Gin, L. E., Guerrero, F. A., Cooper, K. M., & Brownell, S. E. (2020). Is active learning accessible? Exploring the process of providing accommodations to students with disabilities. *CBE—Life Sciences Education*, *19*(4), es12.
- Glaser, B. G., & Strauss, A. L. (1968). *The discovery of grounded theory: Strategies for qualitative research*. Weidenfeld and Nicolson.
- Gouldthorpe, J. L., & Israel, G. D. (2013). *Capturing change: Comparing pretest-posttest and retrospective evaluation methods*. University of Florida IFAS Extension. <https://journals.flvc.org/edis/article/download/120541/118963>
- Green, S. K., Johnson, R. L., Kim, D.-H., & Pope, N. S. (2007). Ethics in classroom assessment practices: Issues and attitudes. *Teaching and Teacher Education*, *23*(7), 999–1011.
- Handelsman, J. (2007). *Scientific teaching*. W.H. Freeman and Co.
- Handelsman, J., Ebert-May, D., Beichner, R., Bruns, P., Chang, A., DeHaan, R., Gentile, J., Lauffer, S., Stewart, J., Tilghman, S. M., & Wood, W. B. (2004). Education. Scientific teaching. *Science*, *304*, 521–522. <https://doi.org/10.1126/science.1096022>
- Handelsman, J., Elgin, S., Estrada, M., Hays, S., Johnson, T., Miller, S., Mingo, V., Shaffer, C., & Williams, J. (2022). Achieving STEM diversity: Fix the classrooms. *Science*, *376*(6597), 1057–1059.
- Hansen, M. J., Palakal, M. J., & White, L. (2023). The importance of STEM sense of belonging and academic hope in enhancing persistence for low-income, underrepresented STEM students. *Journal for STEM Education Research*. <https://doi.org/10.1007/s41979-023-00096-8>
- Henderson, C., & Dancy, M. H. (2007). Barriers to the use of research-based instructional strategies: The influence of both individual and situational characteristics. *Physical Review Special Topics-Physics Education Research*, *3*, 010103. <https://doi.org/10.1103/PhysRevSTPER.3.020102>
- Jiménez, T. R., & Orozco, M. (2021). Prompts, not questions: Four techniques for crafting better interview protocols. *Qualitative Sociology*, *44*, 507–528.
- Jones, S. M., & Kahn, J. (2017). *The evidence base for how we learn: Supporting students' social, emotional, and academic development. Consensus statements of evidence from the Council of Distinguished Scientists*. National Commission on Social, Emotional, and Academic Development & the Aspen Institute [https://www.assets.aspeninstitute.org/content/uploads/2017/09/SEAD-%20Research-Brief-9.12\\_updated-web.pdf](https://www.assets.aspeninstitute.org/content/uploads/2017/09/SEAD-%20Research-Brief-9.12_updated-web.pdf)
- Killpack, T. L., & Melón, L. C. (2016). Toward inclusive STEM classrooms: What personal role do faculty play? *CBE—Life Sciences Education*, *15*(3), es3.
- Knight, J. K., Weaver, D. C., Pepper, M. E., & Hazlett, Z. S. (2022). Relationships between prediction accuracy, metacognitive reflection, and performance in introductory genetics students. *CBE—Life Sciences Education*, *21*(3), ar45.
- Kuh, G. D., Jankowski, N., Ikenberry, S. O., & Kinzie, J. L. (2014). *Knowing what students know and can do: The current state of student learning outcomes assessment in US colleges and universities*. National Institute for Learning Outcomes Assessment [www.learningoutcomesassessment.org/wp-content/uploads/2019/02/2013SurveyReport.pdf](http://www.learningoutcomesassessment.org/wp-content/uploads/2019/02/2013SurveyReport.pdf).
- Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge University Press.

- Lawson, A. E., Banks, D. L., & Logvin, M. (2007). Self-efficacy, reasoning ability, and achievement in college biology. *Journal of Research in Science Teaching*, 44(5), 706–724. <https://doi.org/10.1002/tea.20172>
- LeCompte, M. D., Preissle, J., & Tesch, R. (1993). *Ethnography and qualitative design in educational research*. Academic Press.
- Lee, J., Kim, T., & Su, M. (2021). Reassessing school effectiveness: Multi-objective value-added measures (MOVAM) of academic and socioemotional learning. *Studies in Educational Evaluation*, 68, 100972.
- Leshner, A. I. (2018). Student-centered, modernized graduate STEM education. *Science*, 360(6392), 969–970.
- Leshner, A. I., & Scherer, L. (2019). Critical steps toward modernizing graduate STEM education. *Issues in Science and Technology*, 35(2), 46–49.
- Lockwood, S. A., Miller, A. J., & Cromie, M. M. (2014). Preparing future biology faculty: An advanced professional development program for graduate students. *The American Biology Teacher*, 76(1), 17–21.
- Love Stowell, S. M., Churchill, A. C., Hund, A. K., Kelsey, K. C., Redmond, M. D., Seiter, S. A., & Barger, N. N. (2015). Transforming graduate training in STEM education. *The Bulletin of the Ecological Society of America*, 96(2), 317–323.
- Mack, K., & Winter, K. (2018). Teaching to increase diversity and equity in STEM. *2018 CoNECD-The Collaborative Network for Engineering and Computing Diversity Conference*. American Society for Engineering Education.
- McKinney, J., Chang, M.-L., & Glassmeyer, D. (2021). Why females choose STEM majors: Understanding the relationships between major, personality, interests, self-efficacy, and anxiety. *Journal for STEM Education Research*, 4(3), 278–300. <https://doi.org/10.1007/s41979-021-00050-6>
- McMullen, M. B., Buzzelli, C., & Yun, N. R. (2015). Pedagogy of care for well-being. In *The Routledge International Handbook of Philosophies and Theories of Early Childhood Education and Care* (pp. 283–292). Routledge.
- Merriam, S. B., & Tisdell, E. J. (2015). *Qualitative research: A guide to design and implementation*. John Wiley & Sons.
- Moreira, R. G., Butler-Purry, K., Carter-Sowell, A., Walton, S., Juranek, I. V., Chaloo, L., Regisford, G., Coffin, R., & Spaulding, A. (2019). Innovative professional development and community building activity program improves STEM URM graduate student experiences. *International Journal of STEM Education*, 6, 1–16.
- Moriña, A. (2017). Inclusive education in higher education: Challenges and opportunities. *European Journal of Special Needs Education*, 32(1), 3–17.
- Morris, R., Perry, T., & Wardle, L. (2021). Formative assessment and feedback for learning in higher education: A systematic review. *Review of Education*, 9(3), e3292.
- Morrison, J., Frost, J., Gotch, C., McDuffie, A. R., Austin, B., & French, B. (2021). Teachers' role in students' learning at a project-based STEM high school: Implications for teacher education. *International Journal of Science and Mathematics Education*, 19, 1103–1123.
- National Academies of Sciences, Engineering, and Medicine. (2018). *How people learn II: Learners, contexts, and cultures*. National Academies Press.
- National Research Council. (2003). *Evaluating and improving undergraduate teaching in science, technology, engineering, and mathematics*. The National Academies Press. <https://doi.org/10.17226/10024>
- National Research Council. (2015). *Reaching students: What research says about effective instruction in undergraduate science and engineering*. The National Academies Press. <https://doi.org/10.17226/18687>
- National Research Council. (2018). *Graduate STEM education for the 21st century*. The National Academies Press. <https://doi.org/10.17226/24622>
- O'Leary, E. S., Shapiro, C., Toma, S., Sayson, H. W., Levis-Fitzgerald, M., Johnson, T., & Sork, V. L. (2020). Creating inclusive classrooms by engaging STEM faculty in culturally responsive teaching workshops. *International Journal of STEM Education*, 7(1), 1–15. <https://doi.org/10.1186/s40594-020-00230-7>
- Ovid, D., Rice, M. M., Luna, J. V., Tabayooyong, K., Lajevardi, P., & Tanner, K. D. (2021). Investigating student perceptions of Instructor Talk: Alignment with researchers' categorizations and analysis of remembered language. *CBE—Life Sciences Education*, 20(4), ar61.

- Pfeifer, M. A., Cordero, J. J., & Stanton, J. D. (2023). What I wish my instructor knew: How active learning influences the classroom experiences and self-advocacy of STEM majors with ADHD and specific learning disabilities. *CBE—Life Sciences Education*, 22(1), ar2.
- Polizzi, S. J., Zhu, Y., Reid, J. W., Ofem, B., Salisbury, S., Beeth, M., Roehrig, G., Mohr-Schroeder, M., Sheppard, K., & Rushon, G. T. (2021). Science and mathematics teacher communities of practice: Social influences on discipline-based identity and self-efficacy beliefs. *International Journal of STEM Education*, 8(1), 30. <https://doi.org/10.1186/s40594-021-00275-2>
- Price, E., Lau, A. C., Goldberg, F., Turpen, C., Smith, P. S., Dancy, M., & Robinson, S. (2021). Analyzing a faculty online learning community as a mechanism for supporting faculty implementation of a guided-inquiry curriculum. *International Journal of STEM Education*, 8(1), 17. <https://doi.org/10.1186/s40594-020-00268-7>
- Priftanji, D., Hill, J. D., & Ashby, D. M. (2020). Managing difficult conversations. *American Journal of Health-System Pharmacy*, 77(21), 1723–1726.
- Reeves, T. D., Marbach-Ad, G., Miller, K. R., Ridgway, J., Gardner, G. E., Schussler, E. E., & Wischusen, E. W. (2016). A conceptual framework for graduate teaching assistant professional development evaluation and research. *CBE—Life Sciences Education*, 15(2), es2.
- Reyes, M. R., Brackett, M. A., Rivers, S. E., White, M., & Salovey, P. (2012). Classroom emotional climate, student engagement, and academic achievement. *Journal of Educational Psychology*, 104(3), 700.
- Rivera, S. (2018). A summer institute for STEM graduate teaching assistants. *Journal of College Science Teaching*, 48(2), 28–32.
- Robinson, W. S. (1951). The logical structure of analytic induction. *American Sociological Review*, 16, 812–818.
- Robnett, R. D., Nelson, P. A., Zurbriggen, E. L., Crosby, F. J., & Chemers, M. M. (2018). Research mentoring and scientist identity: Insights from undergraduates and their mentors. *International Journal of STEM Education*, 5(41), 1–14. <https://doi.org/10.1186/s40594-018-0139-y>.
- Rosenberg, M. B., Hilton, M. L., & Dibner, K. A. (2018). Indicators for monitoring undergraduate STEM education. In *Consensus study report*. National Academies Press.
- Rozhenkova, V., Snow, L., Sato, B. K., Lo, S. M., & Buswell, N. T. (2023). Limited or complete? Teaching and learning conceptions and instructional environments fostered by STEM teaching versus research faculty. *International Journal of STEM Education*, 10(1), 51. <https://doi.org/10.1186/s40594-023-00440-9>
- Ruzek, E. A., Hafen, C. A., Allen, J. P., Gregory, A., Mikami, A. Y., & Pianta, R. C. (2016). How teacher emotional support motivates students: The mediating roles of perceived peer relatedness, autonomy support, and competence. *Learning and Instruction*, 42, 95–103.
- Sawyer, R. K. (2006). *The Cambridge Handbook of the Learning Sciences* ((2nd ed.) ed.). Cambridge University Press.
- Schussler, E. E., Read, Q., Marbach-Ad, G., Miller, K., & Ferzli, M. (2015). Preparing biology graduate teaching assistants for their roles as instructors: An assessment of institutional approaches. *CBE—Life Sciences Education*, 14(3), ar31.
- Schwartz, D. L., Tsang, J. M., & Blair, K. P. (2016). *The ABCs of how we learn: 26 scientifically proven approaches, how they work, and when to use them* (1st ed.). W. W. Norton & Company.
- Scott, E. E., Wenderoth, M. P., & Doherty, J. H. (2020). Design-based research: A methodology to extend and enrich biology education research. *CBE—Life Sciences Education*, 19(3), es11. <https://doi.org/10.1187/cbe.19-11-0245>
- Seymour, E. (1995). The loss of women from science, mathematics, and engineering undergraduate majors: An explanatory account. *Science Education*, 79(4), 437–473.
- Seymour, E., & Hunter, A.-B. (Eds.). (2019). *Talking about leaving revisited: Persistence, relocation, and loss in undergraduate STEM education*. Springer. <https://doi.org/10.1007/978-3-030-25304-2>
- Shepard, L. A. (2021). Ambitious teaching and equitable assessment: A vision for prioritizing learning, not testing. *American Educator*, 45(3), 28.
- Shulman, L. S. (1987). Learning To Teach. *AAHE Bulletin*, 40(3), 5–9.
- Shultz, M., Nissen, J., Close, E., & Van Dusen, B. (2022). The role of epistemological beliefs in STEM faculty's decisions to use culturally relevant pedagogy at Hispanic-Serving Institutions. *International Journal of STEM Education*, 9(1), 32. <https://doi.org/10.1186/s40594-022-00349-9>
- Siegel, M. A., Wissehr, C., & Halverson, K. (2008). Sounds like success: A framework for equitable assessment. *The Science Teacher*, 75(3), 43–46.



- Sithole, A., Chiyaka, E. T., McCarthy, P., Mupinga, D. M., Bucklein, B. K., & Kibirige, J. (2017). Student attraction, persistence and retention in STEM programs: Successes and continuing challenges. *Higher Education Studies*, 7(1), 46–59.
- Snow, C., & O'Connor, C. (2016). Close reading and far-reaching classroom discussion: Fostering a vital connection. *Journal of Education*, 196(1), 1–8.
- Stains, M., Harshman, J., Barker, M. K., Chasteen, S. V., Cole, R., DeChenne-Peters, S. E., Eagan, M. K., Jr., Esson, J. M., Knight, J. K., & Laski, F. A. (2018). Anatomy of STEM teaching in North American universities. *Science*, 359(6383), 1468–1470.
- Stephens, N. M., Markus, H. R., & Fryberg, S. A. (2012). Social class disparities in health and education: Reducing inequality by applying a sociocultural self model of behavior. *Psychological Review*, 119(4), 723.
- Stolk, J. D., Gross, M. D., & Zastavker, Y. V. (2021). Motivation, pedagogy, and gender: Examining the multifaceted and dynamic situational responses of women and men in college STEM courses. *International Journal of STEM Education*, 8(1), 35. <https://doi.org/10.1186/s40594-021-00283-2>
- Stronge, J. (2013). *Evaluating what good teachers do: Eight research-based standards for assessing teacher excellence*. Routledge.
- Stroupe, D., Hammerness, K., & McDonald, S. (2020). *Preparing science teachers through practice-based teacher education*. Harvard Education Press.
- Tanner, K., & Allen, D. (2006). Approaches to biology teaching and learning: On integrating pedagogical training into the graduate experiences of future science faculty. *CBE—Life Sciences Education*, 5, 1–6. <https://doi.org/10.1187/cbe.05-12-0132>
- Tanner, K. D. (2013). Structure matters: Twenty-one teaching strategies to promote student engagement and cultivate classroom equity. *CBE—Life Sciences Education*, 12, 322–331. <https://doi.org/10.1187/cbe.13-06-0115>
- Tharani, A., Husain, Y., & Warwick, I. (2017). Learning environment and emotional well-being: A qualitative study of undergraduate nursing students. *Nurse Education Today*, 59, 82–87.
- Theobald, E. J., Hill, M. J., Tran, E., Agrawal, S., Arroyo, E. N., Behling, S., Chambwe, N., Cintrón, D. L., Cooper, J. D., Dunster, G., Grummer, J. A., Hennessey, K., Hsiao, J., Iranon, N., Jones, L., Jordt, H., Keller, M., Lacey, M. E., Littlefield, C. E., et al. (2020). Active learning narrows achievement gaps for underrepresented students in undergraduate science, technology, engineering, and math. *Proceedings of the National Academy of Sciences*, 117(12), 6476. <https://doi.org/10.1073/pnas.1916903117>
- Thomas, C., & Zolkoski, S. (2020). Preventing stress among undergraduate learners: The importance of emotional intelligence, resilience, and emotion regulation. *Frontiers in Education*, 5, 94.
- Thomas, D. R. (2006). A general inductive approach for analyzing qualitative evaluation data. *American Journal of Evaluation*, 27(2), 237–246.
- Trujillo, G., & Tanner, K. D. (2014). Considering the role of affect in learning: Monitoring students' self-efficacy, sense of belonging, and science identity. *CBE—Life Sciences Education*, 13, 6–15. <https://doi.org/10.1187/cbe.13-12-0241>
- Ufnar, J., & Shepherd, V. (2021). The sustained Scientist in the Classroom Partnership: A model for innovative STEM graduate training programs. *Journal of STEM Outreach*, 4(1), 1–15.
- Van Es, E. A., & Sherin, M. G. (2002). Learning to notice: Scaffolding new teachers' interpretations of classroom interactions. *Journal of Technology and Teacher Education*, 10(4), 571–596.
- Von Vacano, C., Ruiz, M., Starowicz, R., Olojo, S., Luna, A. Y. M., Muzzall, E., Mendoza-Denton, R., & Harding, D. J. (2022). Critical faculty and peer instructor development: Core components for building inclusive STEM programs in higher education. *Frontiers in Psychology*, 13.
- Wan, T., Geraets, A. A., Doty, C. M., Saitta, E. K. H., & Chini, J. J. (2020). Characterizing science graduate teaching assistants' instructional practices in reformed laboratories and tutorials. *International Journal of STEM Education*, 7(1), 30. <https://doi.org/10.1186/s40594-020-00229-0>
- Wenger, E. (1998). Communities of practice: Learning as a social system. *Systems Thinker*, 9(5), 2–3.
- Whitaker, T. (2020). *What great teachers do differently: Nineteen things that matter most*. CRC Press.
- White, K., Beach, A. L., Finkelstein, N., & Henderson, C. (2021). *Transforming Institutions: Accelerating systemic change in higher education*. University of Massachusetts Amherst Libraries.
- Wiggins, G. (1998). *Educative assessment*. Jossey-Bass.
- Wiggins, G. P., & McTighe, J. (1998). *Understanding by design*. Association for Supervision and Curriculum Development.
- Williams, A. E., & O'Dowd, D. K. (2021). Seven practical strategies to add active learning to a science lecture. *Neuroscience Letters*, 743, 135317.

- Wilson, D. F. (2004). *Supporting teachers, supporting pupils: The emotions of teaching and learning*. Psychology Press.
- Windschitl, M., Thompson, J., & Braaten, M. (2020). *Ambitious science teaching*. Harvard Education Press.
- Winne, P. H., & Azevedo, R. (2014). Metacognition. In K. Sawyer (Ed.), *Cambridge handbook of the learning sciences*. Cambridge University Press.
- Yik, B. J., Raker, J. R., Apkarian, N., Stains, M., Henderson, C., Dancy, M. H., & Johnson, E. (2022). Evaluating the impact of malleable factors on percent time lecturing in gateway chemistry, mathematics, and physics courses. *International Journal of STEM Education*, 9(1), 15. <https://doi.org/10.1186/s40594-022-00333-3>
- Yorke, L., Rose, P., Bayley, S., Wole, D., & Ramchandani, P. (2021). The importance of students' socio-emotional learning, mental health and wellbeing in the time of COVID-19. *Rise Insights*, 25, 1–11.
- Zee, M., & Koomen, H. M. (2016). Teacher self-efficacy and its effects on classroom processes, student academic adjustment, and teacher well-being: A synthesis of 40 years of research. *Review of Educational Research*, 86(4), 981–1015.
- Zeichner, K. M. (1987). Preparing reflective teachers: An overview of instructional strategies which have been employed in preservice teacher education. *International Journal of Educational Research*, 11(5), 565–575.
- Zemenick, A. T., Turney, S., Webster, A. J., Jones, S. C., & Weber, M. G. (2022). Six principles for embracing gender and sexual diversity in postsecondary biology classrooms. *BioScience*, 72(5), 481–492.
- Zhou, X., Shu, L., Xu, Z., & Padrón, Y. (2023). The effect of professional development on in-service STEM teachers' self-efficacy: A meta-analysis of experimental studies. *International Journal of STEM Education*, 10(1), 37. <https://doi.org/10.1186/s40594-023-00422-x>

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## Authors and Affiliations

Shelley Stromholt<sup>1</sup> · Benjamin Wiggins<sup>2,3</sup>  · Bailey Von der Mehden<sup>4</sup>

✉ Benjamin Wiggins  
bwiggins@shoreline.edu

<sup>1</sup> Aspect Research and Evaluation, Seattle, WA 98105, USA

<sup>2</sup> Shoreline Community College, Shoreline, WA 98133, USA

<sup>3</sup> Department of Pharmacology, University of Washington, 8848 Burke Ave N, Seattle, WA 98195, USA

<sup>4</sup> Department of Ecology and Evolutionary Biology, University of Tennessee, Knoxville, TN 37996, USA