

RESEARCH

Open Access



# Which evidence-based teaching practices change over time? Results from a university-wide STEM faculty development program

Mallory A. Jackson<sup>†</sup>, Sungmin Moon<sup>†</sup>, Jennifer H. Doherty and Mary Pat Wenderoth<sup>\*†</sup> 

## Abstract

**Background:** There is overwhelming evidence that evidence-based teaching improves student performance; however, traditional lecture predominates in STEM courses. To provide support as faculty transform their lecture-based classrooms with evidence-based teaching practices, we created a faculty development program based on best practices, Consortium for the Advancement of Undergraduate STEM Education (CAUSE). CAUSE paired exploration of evidence-based teaching with support for classroom implementation over two years. Each year for three years, CAUSE recruited cohorts of faculty from seven STEM departments. Faculty met biweekly to discuss evidence-based teaching and receive feedback on their implementation. We used the PORTAAL observation tool to document evidence-based teaching practices (PORTAAL practices) across four randomly chosen class sessions each term. We investigated if the number of PORTAAL practices used or the amount of practices increased during the program.

**Results:** We identified identical or equivalent course offerings taught at least twice by the same faculty member while in CAUSE ( $n = 42$  course pairs). We used a one-way repeated measures within-subjects multivariate analysis to examine the changes in average use of 14 PORTAAL practices between the first and second timepoint. We created heat maps to visualize the difference in number of practices used and changes in level of implementation of each PORTAAL practice. Post-hoc within-subjects effects indicated that three PORTAAL practices were significantly higher and two were lower at timepoint two. Use of prompting prior knowledge and calling on volunteers to give answers decreased, while instructors doubled use of prompting students to explain their logic, and increased use of random call by almost 40% when seeking answers from students. Heat maps indicated increases came both from faculty's adoption of these practices and increased use, depending on the practice. Overall, faculty used more practices more frequently, which contributed to a 17% increase in time that students were actively engaged in class.

**Conclusions:** Results suggest that participation in a long-term faculty development program can support increased use of evidence-based teaching practices which have been shown to improve student exam performance. Our findings can help prioritize the efforts of future faculty development programs.

**Keywords:** STEM faculty development, Evidence-based teaching, PORTAAL practices

## Introduction

Discipline-based education researchers have amassed overwhelming evidence that active learning, also referred to as evidence-based teaching, improves student learning in higher education (Derting & Ebert-May, 2010;

\*Correspondence: [mpw@uw.edu](mailto:mpw@uw.edu)

<sup>†</sup>Mallory A. Jackson and Sungmin Moon contributed equally to this work  
Department of Biology, University of Washington, Seattle, WA 98195, USA

Freeman et al., 2011; Knight & Wood, 2005; Prince, 2004; Ruiz-Primo et al., 2011). Many factors contribute to students' success in their coursework, including but not limited to scaffolding of course material, homework assignments, students' motivation, and family responsibilities (Eddy et al., 2015a). However, there is ample data to support the impact of classroom practices on student course and exam performance (Freeman et al., 2014; Moon et al., 2021). Implementing evidence-based teaching practices for more than 67% of class time was also found to significantly narrow differences in performance between students from minoritized and non-minoritized groups (Theobald et al., 2020). Therefore, there is broad national consensus that building the capacity of higher education to implement evidence-based teaching practices is a critical area of action for tackling the educational and workplace challenges for STEM in the twenty-first century (Alberts, 2008; American Association for the Advancement of Science, 2011; Anderson et al., 2011; Association of American Medical Colleges-Howard Hughes Medical Institute, 2009; Committee on the Status, Contributions, and Future Directions of Discipline-Based Education Research et al., 2012; Cooper et al., 2015; Harvey et al., 2016; Olson & Riordan, 2012).

Though these calls for change began over two decades ago (Anderson et al., 2011; Ebert-May et al., 1997; Handelsman et al., 2004), studies using classroom observation tools to assess the level of student-centered teaching methods in STEM classes indicate that traditional lecture predominates (Drinkwater et al., 2017; Ferrare, 2019; Lund et al., 2015; Stains et al., 2018). Faculty offer numerous reasons for their lack of implementation of these effective teaching practices, including lack of training in these teaching methods, lack of time to reorganize and transform their courses, lack of incentive to engage in the time-consuming and challenging process of change (Brownell & Tanner, 2012; Carbone et al., 2019; Henderson & Dancy, 2007; Kim et al., 2019; Shadle et al., 2017), and a concern that course content will be sacrificed in the process (Petersen et al., 2020). Changing faculty's teaching methods, therefore, remains a daunting task that proceeds at a glacial pace (Drinkwater et al., 2017; Ferrare, 2019; Henderson et al., 2011; Lund et al., 2015; Stains et al., 2018).

Numerous faculty development efforts have been implemented to provide support and guidance as STEM faculty attempt to transform their lecture-based classrooms with evidence-based teaching practices (Borda et al., 2020; Callens et al., 2019; Derting et al., 2016; Herman et al., 2018; McConnell et al., 2020; Viskupic et al., 2019). Originally, many of these programs brought teams of faculty from institutions of higher education from across the United States together for week-long faculty

development workshops (Baker et al., 2014; Ebert-May & Weber, 2006; Pfund et al., 2009). More recent efforts have focused within institutions and across their STEM faculty in an effort to change the culture of teaching at that institution (Herman et al., 2018; McConnell et al., 2020).

The Consortium for the Advancement of Undergraduate STEM Education, CAUSE, is a faculty development program at a major R1 institution in the Northwest that fits into the latter category. The CAUSE faculty development program combines attention to the individual faculty member's professional growth, research-based best practices of faculty development programs, and a systems thinking approach. Each of these components will be explained below.

### CAUSE framework

At the level of the individual instructor, the CAUSE program is guided by Clarke and Hollingsworth's interconnected model of professional growth (Clarke & Hollingsworth, 2002) which provides both a model and mechanism by which to promote individual change (Additional file 1: Supplementary Material 1). This model identifies four domains of influence for change: the personal (beliefs about teaching), the external (awareness of alternatives), the domain of practice (implementing new teaching practices), and the domain of outcomes (changes in student performance). Because each domain influences and is influenced by the others, each must be addressed for sustained professional growth. In addition, it is also critical to embed faculty in a well-designed network by providing long-term support that includes performance evaluation and feedback and membership in a community of practice (Kezar, 2011).

Henderson and colleagues (2011) conducted a comprehensive review of 191 articles published from 1995 to 2008 on faculty development efforts in STEM to discern the difference between effective and ineffective programs. Henderson and colleagues found that two early approaches in faculty development turned out to be clearly inadequate: top-down demands from administrators and disseminating "best curriculums" (Henderson et al., 2011; Turpen et al., 2016). The key features of professional development programs for faculty to incorporate more evidence-based practices were, (1) long term support of faculty lasting at least a semester, (2) design elements that recognize and acknowledge an understanding of a university as a complex system, and (3) efforts focused on encouraging faculty to be reflective about their teaching (Henderson et al., 2011).

Austin was commissioned by the National Research Council to summarize the barriers and strategies that exist for faculty to integrate evidence-based approaches into their teaching practices (Austin, 2011). Austin's

major insight was to approach these challenges using systems thinking. Systems thinking recognizes that faculty are indeed individuals, but they are also members of departments that are embedded in colleges or schools, which themselves comprise the larger institution—the university. Both Henderson and colleagues’ and Austin’s work supports the claim that for change to be effective, faculty development programs must acknowledge and incorporate the impact other levels of organization have on the individual (Austin, 2011; Henderson et al., 2011).

**CAUSE faculty development program**

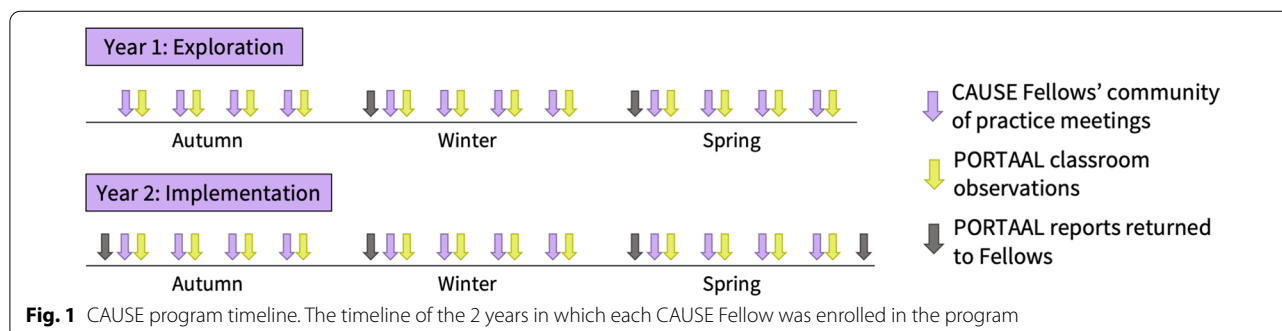
The CAUSE program was designed to incorporate the above recommendations and encourage and support faculty teaching STEM courses at a research-intensive university to use more evidence-based teaching practices (Table 1). The program recruited cohorts of three faculty from seven different STEM departments to participate in the program as CAUSE Fellows. By recruiting three faculty from each department for each of three cohorts, we attempted to create a critical mass of faculty in each

department to sustain efforts over time. This form of community is a critical component of effective faculty development programs as it creates a sense of comradery and support as the Fellows embark on implementing evidence-based teaching practices in their classes (Elliott et al., 2016; Newman, 2017). The cross-discipline nature of the community was positioned to impact the culture of STEM teaching and create a new set of norms across STEM departments.

Each CAUSE Fellow committed to a two-year program consisting of an *Exploration* phase followed by an *Implementation* phase. During the *Exploration* phase, the Fellows attended biweekly meetings for a total of four meetings each quarter (Fig. 1). These meetings were facilitated by a faculty member who was engaged in biology education research and experienced with using evidence-based teaching practices. At the biweekly meetings, CAUSE Fellows read and discussed seminal papers to increase awareness of evidence-based teaching practices. They also discussed what they saw when they visited large classes on campus that were

**Table 1** Features of CAUSE from the primary literature

Research and theoretical models for faculty development programs		CAUSE faculty development program
Interconnected model of professional growth (Clarke & Hollingsworth, 2002)	Beliefs about teaching	Read Instructor Mindset paper (Canning et al., 2019)
	Awareness of alternatives	<i>Exploration Phase:</i> Reading the literature
	Implementing new teaching practices	<i>Implementation Phase:</i> Practice in classroom
	Changes in student performance	Achievement gap reports each quarter
Best practices for faculty development (Henderson et al., 2011)	Long-term support	Two-year program commitment
	Recognize university as a complex system	Work at department level and select departmental leads
	Encourage faculty reflection	Provide quarterly PORTAAL and achievement gap reports
Systems thinking University as an ecosystem (Austin, 2011)	Structural: department culture	Department leads recruited faculty to program
	Human resources: community of practice	Biweekly group discussions
	Political: assessment data	PORTAAL and achievement gap reports
	Symbolic: travel awards and Provost recognition	\$2,000 for travel to national conference to present work on transformation of course



implementing evidence-based teaching practices. The *Implementation* phase was designed to provide support as faculty selected one or two teaching practices to implement in their own classes. The biweekly meetings in this phase focused on barriers and successes during this process. Upon completion of the two-year CAUSE program, Fellows were eligible to receive \$2000 to attend an academic professional meeting, where they were encouraged to present a poster on how they transformed their undergraduate STEM course from a traditional lecture format to an evidence-based teaching format.

To monitor the classroom teaching practices of CAUSE Fellows over time, we used the Practical Observation Rubric to Assess Active Learning (PORTAAL; Eddy et al., 2015b). PORTAAL identifies readily implemented teaching practices that have evidence in the literature to increase academic achievement of students. Unlike other classroom observation tools, every second of an entire class session is coded for the presence of teaching practices, whether they are instructor-centered or student-centered. As PORTAAL identifies a subset of evidence-based teaching practices, we will refer to these teaching methods as PORTAAL practices. To provide feedback, CAUSE Fellows received a PORTAAL report after each term of teaching that consisted of their average value for each practice as well as a histogram showing the level of implementation of each teaching practice for all Fellows for comparison (Additional file 1: Supplementary Materials 2 and 3). PORTAAL reports were distributed and implications for modifications to teaching methods were discussed during biweekly meetings.

As the fourth domain of Clarke and Hollingsworth's interconnected model of professional growth is the domain of outcomes (Clarke & Hollingsworth, 2002), we collected and analyzed exam scores from each class a Fellow taught. Using this information, we created a report each term for each Fellow showing differences in academic performance between students based on binary gender, first-generation status, race/ethnicity, and economically or educationally disadvantaged status.

### Research questions

We were interested in documenting the patterns of change that STEM faculty had in the adoption and implementation of specific PORTAAL practices during their time in the CAUSE faculty development program. Our hypothesis was that CAUSE, a program based on best practices of faculty development, in the context of a theory of change, would encourage and support faculty to adopt PORTAAL practices. We posed two research questions:

1. Did implementation of the PORTAAL practices change from the first to last time instructors taught the same or equivalent course during CAUSE?
2. Were the changes in implementation of PORTAAL practices due to Fellows adopting these practices, increased use of practices over time, or a combination of adopting practices and increased use?

## Methods

### Data collection

#### Courses and instructors

The CAUSE program recruited three cohorts of faculty ( $n=47$ ) at a major R1 institution from seven STEM departments: biology, chemistry, computer science, mathematics, physics, psychology, and public health. This R1 institution is on the quarter system. Faculty in CAUSE collectively taught 157 courses during the program. Information on gender and academic track of all Fellows is included in supplementary information (Additional file 1: Supplementary Material 4). To determine how Fellows changed their teaching practices over the course of their long-term participation in CAUSE, we identified identical or equivalent course offerings that were taught at least twice by the same faculty member during the time they were in CAUSE. A visual representation of the courses that were taught at least twice by the same instructor can be found in Additional file 1: Supplementary Material 5.

From 157 total courses, we identified 42 pairs of identical or equivalent course offerings. Of the 42 course pairs, 35 were identical courses. Equivalent courses were defined as courses in the same introductory series in chemistry, organic chemistry, math, or physics; courses of the same division (either upper or lower) and relatively the same class size within the same department; courses that were both non-majors (or open enrollment) within the same department. Of the seven courses that were not identical, three courses were in the same introductory series (e.g., chemistry or physics), while four were of the same course level, class size, and had similar course content within that discipline. The final data set included 35 unique instructors, as seven instructors taught two pairs of equivalent courses. Characteristics of the 42 pairs of courses are included in Additional file 1: Supplementary Material 6.

Human subjects research was conducted with the approval of the Institutional Review Board at the University of Washington (STUDY00002830).

#### Classroom observation data

Our classroom observation procedure was the same as in the study by Moon and colleagues (2021). We randomly

selected and coded recordings of four class sessions of each course for each timepoint. Most recordings were made by the university’s lecture capture technology, and a few were filmed by the research team when lecture capture technology was not available. To select which recording to include in this study, the 10-week quarter was divided into four equal time periods excluding the first and last week of the quarter, with one day randomly selected from each time period. All videos were coded by trained researchers using the classroom observation tool, PORTAAL (Eddy et al., 2015b). The PORTAAL coding rubric is found in supplementary information (Additional file 1: Supplementary Material 7).

Over the course of the project, six coders were trained by a member of the research team to code classroom videos using PORTAAL. For each coding assignment, the coders were paired to observe the same set of videos, and pairs were shuffled each time a new set of videos was assigned. Novice coders were first paired with a more experienced coder. The two coders individually coded the entire class video that was selected for analysis. Coders then met and discussed any disagreements in coding to come to complete consensus on all codes for each video. The most common coding disagreements were Bloom’s level of the activity, time differences of a few seconds when recording the start and end of an activity, or tallying the number of students who responded during debrief if student voices were difficult to hear. These disagreements may have occurred a few times for each 50–110-min video. On rare occasions, a code could not be reconciled between the coders. In this instance, a

member of the research team provided a final decision on how to code the disagreement.

Though PORTAAL identified 21 teaching practices from the literature that were documented to increase student academic performance (Eddy et al., 2015a), not all practices are commonly seen in most courses. Therefore, prior to analysis, we selected 14 PORTAAL practices that were frequently used by CAUSE Fellows and had minimal redundancy for our analyses (Table 2). We determined the average value for each of the PORTAAL practices by totaling the number or duration of each practice and dividing by the number of videos coded. The average PORTAAL values were then standardized to a 50-min period which is the length of most class sessions at this institution. These average values were used in all the subsequent analyses.

**Data analysis**

***RQ 1: Did implementation of the PORTAAL practices change from the first to last time instructors taught the same or equivalent course during CAUSE?***

We used a one-way repeated measures within-subjects multivariate analysis ( $n = 42$ ) to examine the changes in PORTAAL practices for instructors. Analyses were conducted in SPSS (Version 25.0). Fellows who taught the same course at least twice were included in the data set. The majority of Fellows only taught the same or equivalent course twice; however, a few taught the same course as many as five times during CAUSE (Additional file 1: Supplementary Materials 5 and 8). Though we sought to look for patterns of change each time a Fellow taught

**Table 2** PORTAAL practices

PORTAAL practice	Instance (I)/ Duration (D)
(1) Students working alone [Alone]	I
(2) Providing alternative answers to a question [Alt_Ans]	I
(3) Students explaining answers [Exp_Ans]	I
(4) High Bloom’s in-class activities [HB]	I
(5) Instructor provides answer to a question [Ins_Ans]	I
(6) Instructor explains the answer to a question [Ins_Exp]	I
(7) Instructor prompts students to use prior knowledge [PK]	I
(8) Instructor use of positive feedback [Pos_FB]	I
(9) Instructors prompting logic from students [Prom_Log]	I
(10) Instructors randomly calling on students for answers [RC_Ans]	I
(11) Students working in small groups [SG]	I
(12) Student time in debrief of a question [ST_DB]	D
(13) Total student time during class [TST]	D
(14) Student volunteer or whole class answers a question [V_WC_Ans]	I

Unit of instance:  $n$  times. Unit of duration: minutes. Code names of each practice are included in brackets



the same course, we did not have a large enough sample size to conduct this level of analysis. Therefore, we had to limit our analysis to only the first and last teachings. We will refer to these pairs of courses as “courses at timepoint one” (the first time they taught the course in CAUSE) and “courses at timepoint two” (the last time they taught the course while in CAUSE).

We are aware that first adopters of a new technique are often more willing to change than late adopters; therefore, we controlled for the CAUSE cohort (1, 2, or 3). As faculty had very different teaching schedules, some taught the same course multiple times between the first and second timepoint; therefore, we controlled for whether the same course was taught in the interim. As time between teachings of a course may impact implementation of new practices, we also controlled for the number of quarters between the first and second timepoint. We also included course level (upper or lower division) as a covariate to control for any effects of differential implementation of evidence-based teaching by course level.

It was important to account for the variability across instructors, equivalent courses, and instructor/course pairs as our data included two timepoints of observations, seven instructors who taught multiple pairs of courses, and seven pairs of equivalent courses that were taught by multiple instructors. We, therefore, tested all possible random intercept structures using instructors, equivalent courses, and instructor/course pairs, then compared the model fit using Akaike’s Information Criterion (AIC). We considered the differences between the two AIC values:

$$\Delta i = AIC_i - AIC_{\min}$$

where  $AIC_i$  is the AIC value of the  $i$ th model, and  $AIC_{\min}$  is the lowest AIC value.

The model with the lowest AIC was considered the best fit model, unless  $\Delta i$  was less than ten, in which case the model with the fewest parameters was the best fit (Burnham & Anderson, 2004). We found that the model with the lowest AIC included an instructor random intercept (Additional file 1: Supplementary Material 9), which was subsequently included in our repeated measures analysis.

**RQ 2: Were the changes in implementation of PORTAAL practices due to Fellows adopting these practices, increased use of practices over time, or a combination of adopting practices and increased use?**

To explore the patterns of adoption of PORTAAL practices and the trends in changes in use between timepoints one and two, we used the average value of each PORTAAL practice across the four observations for each course in two different ways. First, we converted

the average values to a binary code to denote the presence or absence of each PORTAAL practice in each course. Any average value other than zero was coded as “1” and zeros were coded as “0”. We created four categories of possible types of change in use based on the binary codes: “Always Used”, “Began Using”, “Stopped Using”, and “Never Used”. If a CAUSE Fellow used a PORTAAL practice at both timepoints (i.e., the binary code for that practice was “1” at both timepoints), their code for the type of change in use of that practice would be “Always Used”. We coded the changes in use as previously described for all PORTAAL practices in all the paired courses. We then created heat maps using the ggplot2 package in R (R Core Team, 2021; Wickham, 2016) to discern patterns of changes in types of use of PORTAAL practices for each paired course by department.

Next, we used the average PORTAAL scores across the four observations for each course at timepoints one and two to calculate the difference in average use. We calculated these values by subtracting the average PORTAAL value at timepoint two from the average value of the same practice at timepoint one. We then created a heat map of these difference scores for each practice in which the intensity and hue of the color represents the amount of positive or negative change between timepoints. Each Fellow was assigned a three-digit code which was used to maintain confidentiality when labeling the heat maps. Departments and course numbers were also de-identified, though the de-identified course numbering correctly reflects the course level of each paired course (e.g., 101 refers to a 100-level course).

## Results

### RQ 1: Did implementation of PORTAAL practices change?

When pooling all 14 practices, there was no significant difference in mean PORTAAL values across time ( $F_{(14, 23)} = 1.207$ ,  $p = 0.334$ ). However, post-hoc analyses for PORTAAL practices found that three practices were implemented significantly more at the second timepoint and two practices were significantly less (Table 3). At timepoint two, Fellows doubled their use of prompting students to explain their logic about in-class questions, increased use of random call by almost 40% when asking students for answers, and increased the total class time when students were actively engaged in the course by 17%. The two practices that significantly decreased from the first to second timepoint were the number of times the instructor called on an individual student volunteer or the whole class to give an answer and the number of times instructors reminded students to use their prior knowledge (Table 3).

**Table 3** Pairwise comparisons of the PORTAAL practices

PORTAAL practice	Mean (Time 1)	Mean (Time 2)	Mean difference (T2-T1)	Std. error	p value
Working alone	1.856	1.924	0.069	0.184	0.712
Alternative answers	0.574	0.654	0.08	0.143	0.582
Explaining answers	3.569	4.395	0.827	0.516	0.118
High Bloom's activities	2.347	2.461	0.115	0.237	0.631
Instructor answers	3.774	4.115	0.341	0.348	0.332
Instructor explains	5.088	4.984	- 0.104	0.271	0.703
Prior knowledge**	0.417	0.156	- 0.261**	0.083	0.003
Positive feedback	3.124	3.111	- 0.013	0.260	0.961
Prompting logic*	1.099	2.201	1.103*	0.432	0.015
Random call answers*	1.967	2.738	0.771*	0.347	0.033
Small groups	4.066	4.234	0.168	0.315	0.597
Student time in debrief	5.479	6.551	1.072	0.604	0.084
Total student time*	13.185	15.413	2.228*	0.937	0.023
Volunteer or whole class answers*	9.453	7.895	- 1.559*	0.740	0.042

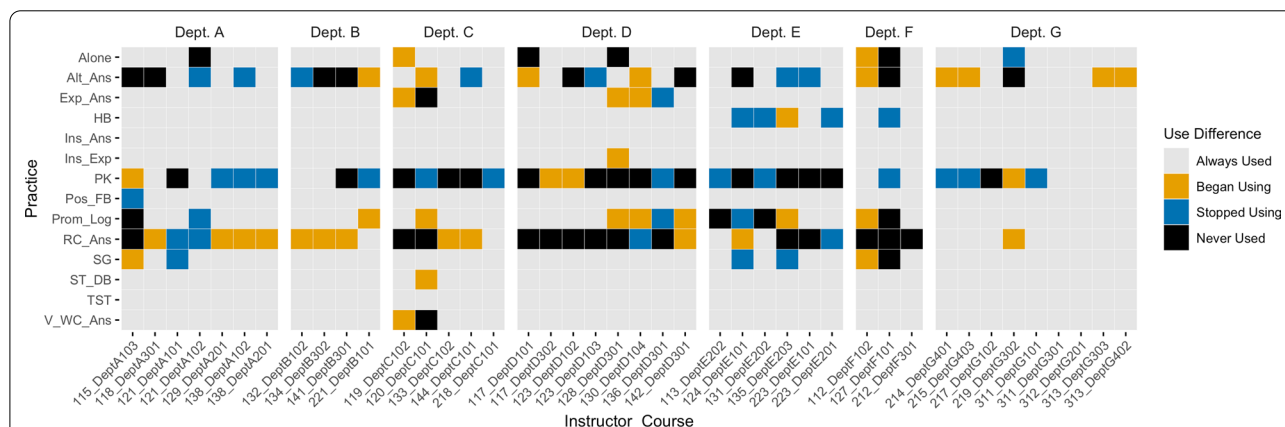
p values are Bonferroni adjusted for multiple comparisons. \*p < 0.05. \*\*p < 0.01. \*\*\*p < 0.001

**RQ 2: What is the pattern of change?**

We created two heat maps to explore the type of changes in use of each PORTAAL practice (i.e., if Fellows adopted practices between timepoints one and two) and changes in the average level of implementation of PORTAAL practices. Figure 2 indicates if Fellows always used the PORTAAL practice (i.e., at both timepoints one and two), began using new practices at the second timepoint, stopped using practices at the second timepoint, or never used the practice. Across all paired courses, 17 courses (40%) adopted one new PORTAAL practice, five courses (12%) adopted two practices, four courses (10%) adopted three practices and one course taught by

Fellow 114 adopted four new practices. In total, 27 of the courses (64%) were using at least one new practice between timepoints one and two (Table 4). The percent of courses within a department that adopted at least one new practice varied across STEM departments from 100% of courses in Department B to only 33% of courses in Departments E and F (Table 4). The percent of each type of use across all PORTAAL practices within each department is found in Additional file 1: Supplementary Material 10.

Two PORTAAL practices were always used (gray boxes) at both timepoints across all courses in all departments: “Instructor answers” and “total student



**Fig. 2** Changes in presence or absence of PORTAAL practices from T1 to T2. Each column represents the change in use based on the binary values of each PORTAAL practice between timepoints one and two for a paired course. Each x-axis label refers to a Fellow and the de-identified course they taught. The de-identified course numbers reflect the course level. If the same course was taught by multiple Fellows, the de-identified course number is the same. See Table 2 for the long descriptions of the code names of the PORTAAL practices shown on the y-axis

**Table 4** Number and percentage of courses per department that adopted at least one new PORTAAL practice

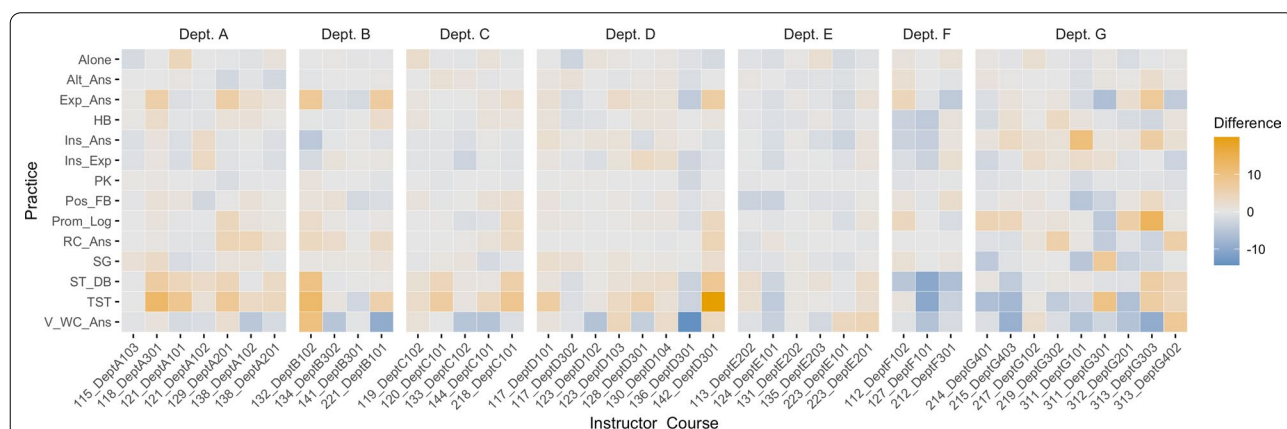
Department	Number of courses	Number of courses that adopted a new PORTAAL practice	Percent of courses that adopted a new PORTAAL practice
A	7	5	71%
B	4	4	100%
C	5	4	80%
D	8	6	75%
E	6	2	33%
F	3	1	33%
G	9	5	56%
Total	42	27	64%

time” (Fig. 2). Adopting a new PORTAAL practice was not evenly distributed across practices. “Random call answers” was adopted by the greatest number of courses (12 courses) followed by explaining “alternative answers” to questions (9 courses) and instructors’ use of “prompting logic” when asking students to solve in-class problems (7 courses). Stopping use of a PORTAAL practice (blue boxes) showed similar heterogeneity. Instructors in 13 courses stopped prompting students to use their “prior knowledge” when working on in-class problems and seven stopped asking students to explain alternate answers to questions. Use of “prior knowledge” was also the practice that was never used (black boxes) in 15 courses, while “random call answers” was never used in 14 courses. The percent of practices Fellows always used, began using, stopped using, and never used is found in Additional file 1: Supplementary Material 11. A heat map of the binary use values at each timepoint is found in Additional file 1: Supplementary Material 12.

To discern how Fellows changed the average level of implementation of PORTAAL practices, we created a heat map showing the difference in average values of each practice between the two timepoints (Fig. 3). Gray represents no change in the average value of the PORTAAL practice over time. A heat map showing the average values of the PORTAAL practices for each paired course at each timepoint is found in supplementary information (Additional file 1: Supplementary Material 13).

We found that changes in average implementation of all PORTAAL practices varied by CAUSE Fellow, course, and department. For example, Fellow 136 decreased use of PORTAAL practices more than Fellow 142, though they taught the same course in the same department. This variation also occurred at the department level. For example, Department A had more positive trends in implementation of PORTAAL practices than Department E. In addition, we plotted histograms of the average PORTAAL values at each timepoint for the practices that were significant in the repeated measures analysis to examine if the changes were due to a few Fellows greatly increasing the use of these practices. We found that the changes appear to be due to several Fellows changing their use of these practices rather than a few outliers (see Additional file 1: Supplementary Material 14).

The significant increase in “prompting logic” appears to be spread across most departments but may have been most pronounced in Department G, where all but one of the courses increased use of this practice. Across all 42 paired courses, 17 courses showed an increased average use of “random call answers”. Of these, only five of the courses were using random call at the first timepoint (see Additional file 3). We investigated if the significant increase in use of “random call answers” and significant



**Fig. 3** Changes in average values of PORTAAL practices from T1 to T2. Each column represents the difference in average value of each PORTAAL practice between timepoints one and two for a paired course. Each x-axis label refers to a Fellow and the paired course they taught. The de-identified course numbers reflect the course level. If the same course was taught by multiple Fellows, the de-identified course number is the same. See Table 2 for the long descriptions of the code names of the PORTAAL practices shown on the y-axis



decrease in “volunteers or whole class answers” was due to Fellows shifting from calling on volunteers to random call. We found that 11 paired courses fit this trend (see Additional file 1: Supplementary Material 15 and Additional file 3), while only six of the paired courses increased both “random call” and “volunteer or whole class answers”.

The practices “total student time” and “volunteer or whole class answers” were used by nearly all Fellows at both timepoints. The significant increase in time that students were actively engaged in class was, therefore, due to increased implementation by Fellows who were already using some student-centered classroom activities. This trend occurred predominantly in Departments A, B, C, and D.

Of all the paired courses, 21 implemented the practice “prior knowledge” fewer times on average at the second timepoint. Only eight of these courses continued to use “prior knowledge”, which suggests that the significant decrease in this practice may be due to Fellows stopping their use of this practice. However, this was the least frequently implemented practice, so these results should be interpreted with caution.

## Discussion

Our study found that faculty who participated in the CAUSE faculty development program changed their teaching behaviors and incorporated more evidence-based teaching practices over their time in CAUSE. Compared to the first time a Fellow taught a course, while in CAUSE, Fellows more frequently prompted students to use their logic when answering questions, used random call when asking students to report out on their answers, and increased the amount of time students were working in class at the second timepoint. Fellows decreased their use of volunteers or the whole class answering questions and the use of prompting students to use their prior knowledge when solving problems in class. These results suggest that one of the most successful impacts of CAUSE was encouraging and allowing instructors to think of ways to shift the time spent in class from instructor-centered to student-centered, and to move towards a more equitable classroom by decreasing the use of calling on volunteers during class.

These gains may have been because the CAUSE program was specifically designed to incorporate many of the previously identified aspects of successful faculty development programs. The CAUSE program was a long-term program, provided a community of support both within and across STEM departments, increased awareness and practice of evidence-based teaching practices, and provided feedback to Fellows on both changes in

their teaching methods and how student performance in their courses changed over their time in CAUSE.

### Changes in specific PORTAAL practices

Five PORTAAL practices showed significant changes between the first to second timepoint. Below, we further describe each of these PORTAAL practices and offer possible explanations for how the CAUSE faculty development program may have impacted Fellows’ use of these practices.

#### *Prompting logic*

CAUSE Fellows doubled the number of times per class session they prompted students to explain logic between their first and second time teaching the same course during CAUSE. Faculty added on average one more use of this PORTAAL practice per class. Over the course of 30 class sessions in a 10-week quarter, this would translate to students hearing their instructor remind them to use logic and reasoning when answering questions approximately 60 times rather than just 30 times. Such repetition may further instill in students the value and habit of using logic in their problem solving. Seventy percent of courses always used this practice, and by timepoint two, Fellows in seven more courses added this practice to their teaching repertoire, while only two stopped using it. We coded the use of “prompting logic” when instructors asked students to explain their logic or reasoning behind the answer to a question. This often occurred at the beginning of a clicker question or discussion question when students were working in groups and the instructor reminded students to not focus solely on the correct answer, but rather to make sure that they can explain to their peers the reasoning of how they came to that answer. Similarly, when faculty asked students to report on their answers, they could prompt them to indicate the logic behind the solution to the problem. We hypothesize that increased use of prompting logic may be due to the ease of implementing this PORTAAL practice. Incorporating the practice of prompting logic does not require much planning or time investment by the instructor, as the instructor just needs to remember to ask the students to explain their reasoning. If the instructor’s goal is to increase the amount of time that students are working or answering questions in class, reminding students to discuss their logic with their peers, and asking for students to explain the logic to the class is an effective way to make the class more student-centered.

When given an in-class question to answer, students often focus on quickly guessing the correct answer or trying to eliminate incorrect answers if it is a multiple-choice question. For deep learning to occur, students need to take the time to apply their knowledge to reason

through to the correct answer (Kahneman, 2011). However, research shows that students need to be reminded to take the time needed to analyze the question and formulate an answer based on logic (Knight et al., 2013). This PORTAAL practice was emphasized in CAUSE meetings when faculty discussed best practices in use of multiple choice clicker or in-class short answer questions (Chien et al., 2016; Lewin et al., 2016; Smith & Knight, 2020; Smith et al., 2011). Our previous research has also supported the finding that use of prompting logic is positively correlated with student exam performance (Moon et al., 2021).

#### **Random call**

The use of random call per class session increased about 40% between the first and second timepoint. Initially, random call was used about twice per class and this increased to almost three times. Again, this would translate into a change from approximately 60 random calls per quarter to more than 80 over the same time period which would allow both more voices and a greater diversity of voices to be heard over the 10-week quarter. Interestingly, random call was the most frequently adopted PORTAAL practice at the second timepoint but was also never used in 14 of the 42 paired courses. This suggests that even with the focus on this practice during meetings and support from other Fellows, some participants in CAUSE were still hesitant to use random call. We coded “random call answers” when an instructor used a randomized list of names or seat numbers to call on individual students or groups to give their answers to an in-class formative assessment question.

Random call was one of the practices that was emphasized during the biweekly meetings. The CAUSE Fellows read and discussed three papers that show how using random call can contribute to more inclusive classrooms (Dallimore et al., 2010; Eddy et al., 2014; Waugh & Andrews, 2020). Furthermore, the Fellows often shared stories about how they implemented random call in their classrooms. These stories may have been encouraging for faculty who were not yet using random call, and this sense of support from the community of CAUSE Fellows may have motivated others to implement this practice. Another explanation could be that using random calling is a natural next step for instructors who have scaffolded their lesson plans with interactive clicker questions that are debriefed by students providing answers. In this case, the instructor can use a random list of student names or groups to call on to answer the question. We hypothesize that the primary reasons we found an increase in the average number of random calls per class in our data set were the emphasis during CAUSE meetings on creating equitable classrooms and the support from instructors

in CAUSE who were already using random call and were willing to share their insights.

The challenge in any classroom is hearing from all students and not just those sitting in the front two rows. Though some are concerned that calling on students to answer in-class questions may be anxiety provoking for students (Cooper et al., 2018), others have found that when used in an intentional manner and employing specific guidelines, random call can help to make the classroom more equitable (Dallimore et al., 2010, 2013; Waugh & Andrews, 2020). During the CAUSE biweekly meetings, particular attention was paid to how to implement the method to decrease student anxiety. The CAUSE program emphasized the same key components that Waugh and Andrews found were important to the college biology faculty they interviewed: instructors should explain to students why they are using random call, allow students to talk before calling on them, select a group of students to answer, and ask students to report out the collective idea of their group (Waugh & Andrews, 2020).

#### **Total student time**

The CAUSE Fellows in this study had an almost 17% increase in the amount of time students were talking, thinking, or working in class. “Total student time” was made up of all the time per 50-min class that students spent thinking alone, working in groups, or answering questions. It was the inverse of time the instructor is giving instructions, lecturing, or answering questions. On average, CAUSE Fellows increased total student time from 26% of class time to 31% of class time. We hypothesize this increase in total student time may be attributed to the fact that PORTAAL practices are designed to incorporate more time for students to engage with the material on their own or with a peer. The significant increase in the number of times that faculty prompted the use of logic could increase the amount of time that students used to explain the logic underlying their answer to either their peer group or the class. Changes in total student time in debrief trended toward significance and provides some support that this practice also contributed to increased total student time.

Active learning has been defined as “instructional activities involving students in doing things and thinking about what they are doing” (Bonwell & Eison, 1991) and “a process that requires students to select, organize, and integrate information, either independently or in groups” (Committee on the Status, Contributions, and Future Directions of Discipline-Based Education Research et al., 2012). Though these practices can take many forms, at their foundation is increased student engagement with course material as they move

from an orientation of passive listening and reception of material to constructing understanding. A theory of learning that encapsulates these ideas is the Interactive, Constructive, Active, Passive (ICAP) framework (Chi & Wylie, 2014). In the ICAP framework, Active is defined as students taking notes or manipulating an object as part of a task rather than being involved in activities traditionally associated with active learning, while Passive refers to students sitting quietly and listening. Chi and Wylie found support that Constructive activities maximize student learning as they produce deep understanding with potential for transfer, while the Interactive mode produces deeper understanding with potential to innovate novel ideas (Chi & Wylie, 2014). Implementing more PORTAAL practices that are Interactive and Constructive may have contributed to an overall increase in the total amount of time that students were working alone, discussing in small groups, and reporting their answers to the whole class.

#### ***Volunteer and whole class answers***

Faculty calling on “volunteers or the whole class to answer” was the most frequently used PORTAAL practice at the first timepoint. This was not surprising as many instructors have a history of calling on volunteers to answer questions, which contributes to the interactive nature of the classroom. There was a 16% decrease in use of this practice at timepoint two. We hypothesize that this change may be due to the increase in use of random call rather than calling on volunteers, as when one method of calling on students increases, the other usually decreases. In our study, we found that instructors often used both methods for gathering answers from students, even at timepoint two. We were, however, very encouraged to see use of random call increase significantly, as use of volunteers has shown a gender and racial bias in the classroom (Eddy et al., 2014; Reinholz & Shah, 2018; Reinholz et al., 2020).

The significant increase in random call answers and decrease in volunteers and whole class answers we observed may have been due to reciprocal use of these practices, as we found that 26% of the paired courses had both an increase in use of random call and a decrease in volunteer or whole class answers. We hypothesize that faculty may replace calling on volunteers or the whole class with using random call. However, we never found an instance, where faculty stopped calling on volunteers. This may suggest that faculty find the practice of calling on volunteers useful for soliciting answers from students and, therefore, may continue using both practices in their classrooms.

#### ***Prompting prior knowledge***

Prompting students to use prior knowledge was the least-implemented PORTAAL practice at both timepoints. “Prompting prior knowledge” was coded when an instructor reminded students to use knowledge from a previous class session or course to answer a question. A recent paper by Arthurs and Kowalski reinforces the valuable contribution this practice has for student learning (Arthurs & Kowalski, 2021). Though we detected a significant decrease in the implementation of this practice, this result should be interpreted cautiously due to its limited use by CAUSE Fellows. Future research should investigate use of this practice in STEM classrooms and its correlation with student learning.

#### ***Patterns of change***

The heat map of differences in the use of practices between the two timepoints provides insight to the patterns of adoption of PORTAAL practices (Fig. 2). Across all paired courses, Fellows in 64% of courses adopted at least one new practice, and one Fellow incorporated four new practices in their course. Patterns of adoption of new practices were department specific. For instance, in Department B, every course added a new practice, while in Department E, only two of six courses added a new PORTAAL practice. Department G also had a low percentage of courses that adopted a new practice (56%), but unlike in Department E, 90% of the PORTAAL practices were always used across courses (Additional file 1: Supplementary Material 10). Therefore, Department G was likely experiencing a ceiling effect for practice adoption. Department G may also represent the added benefit of having a community of faculty who regularly use evidence-based teaching practices. A critical mass of faculty can provide support for both learning about evidence-based practices and how best to implement them. The lack of a critical mass of faculty using PORTAAL practices in Department E may explain why only two Fellows added a new practice. It is possible that the culture in this department is not as supportive of evidence-based teaching practices.

For the PORTAAL practices that changed significantly from timepoint one to timepoint two, the heat maps provide insight to if the changes were due to adoption or increased use in the same courses. Depending on the practice, there was considerable variation in the level of adoption and change in average implementation of the practice. The significant changes in some practices appear to be driven by adoption by Fellows, while others are driven more so by Fellows increasing or decreasing the amount that they used the practices.

An example of a PORTAAL practice that showed a significant increase from timepoints one to two and seems to be driven by widespread adoption by CAUSE Fellows is random call. Based on the heat maps, most Fellows were using this practice by the second timepoint. As random call was one of the least frequently implemented practices during Fellows' first time teaching during CAUSE, it suggests that the faculty development program successfully supported implementation of this new evidence-based teaching practice.

Alternatively, some practices were implemented across nearly all courses at both timepoints, and the significant changes in their average values are not due to adoption or Fellows who stopped using these practices. For example, "total student time" was used in all courses at both timepoints but we found a significant increase in its average use over time. This suggests that Fellows were finding ways to make their courses more student-centered during their time in the program. Though the overall trend showed that faculty were increasing the amount of time students spent working alone, in small groups, or providing answers, the change was not the same across departments. Department G and Department F showed a decreasing trend in total student time. Fellows often commented in the biweekly meetings that they were having trouble pacing their student activities, which caused them to fall behind on content and curtail some activities. In contrast, in Department A the average amount of total student time doubled, yet it was still only half of the average amount of total student time in Department G (Additional file 2). This suggests that even though departments are using more student-centered teaching practices, there is still room for considerable growth in many STEM classrooms.

### Limitations

In this study, we only coded four class videos for each course. We realize that there may be variation in how faculty teach from day to day, particularly as faculty begin to implement new teaching practices. Therefore, coding more observations could have better represented faculty's teaching, as is suggested in recent research (Sbeglia et al., 2021; Stains et al., 2018).

As we only coded four class sessions over the course of the term, it is possible that at timepoint two we missed other PORTAAL practices that Fellows added but used less often. Similarly, though we have labeled one set of use, "Stopped Using", it is possible that Fellows were still using this practice but not on the days that we selected to code classroom videos. Therefore, it is possible that we underestimated the number of PORTAAL practices added and overestimated the number of PORTAAL

practices that Fellows stopped using based on these observations.

Teaching schedules also prevented us from doing repeated measures analysis for all the faculty in the CAUSE program, because faculty either did not teach often enough or taught different classes each quarter. This was particularly true for tenure-track Fellows, which explains why there are predominately teaching-track faculty in this study. As teaching is the primary responsibility of teaching-track faculty and, therefore, the basis of their promotions, it is possible that our data was biased towards instructors who were more motivated to improve their teaching by incorporating evidence-based practices.

### Conclusions and implications for faculty development programs

The CAUSE faculty development program operationalized the theoretical foundations of faculty change by addressing attention to the individual faculty members' professional growth, best practices in faculty development, and a systems thinking approach (Table 1). Some of the unique aspects of the CAUSE program are that it used PORTAAL as an objective classroom observation tool and provided quarterly feedback on teaching and student exam performance to each CAUSE Fellow. Quarterly academic reports on how the various groups of students in their classes performed on exams kept the faculty informed as to the impact of their teaching on all students in their class.

During their tenure in CAUSE, Fellows doubled their use of prompting logic when asking students questions, a practice that may help students develop reasoning skills that could transfer to improved exam performance (Kahneman, 2011). CAUSE Fellows also decreased calling on volunteers and the whole class for answers while increasing their use of random calls. The changes in these practices may contribute to increased equity in the classroom (Dallimore et al., 2010; Eddy et al., 2014). Similarly, the increase in time that students are actively engaged in building their understanding in class can foster a more student-centered and inclusive course. We realize that a class has a finite time period which may limit the number of PORTAAL practices an instructor can use. However, our findings show that multiple faculty had high rates of implementation of many of the PORTAAL practices, indicating that it is possible to incorporate a variety of practices during each class session. With the increased use of PORTAAL practices, we can expect that student academic performance should also increase as previous research has shown a positive correlation between the level of



evidence-based teaching and student exam performance (Eddy et al., 2015b; Freeman et al., 2014; Moon et al., 2021; Theobald et al., 2020).

Change, however, is slow and “the journey begins with one step.” In the case of CAUSE Fellows, that one step is prompting students to use logic and reasoning and using some form of random calling on students. Faculty development programs have a daunting task ahead if they are to answer the national calls to transform STEM teaching in higher education. Transforming traditional lecture to evidence-based teaching practices is necessary if we are to make our classrooms more inclusive and ensure a more diverse and productive STEM force in the future. Our findings can help prioritize the future efforts of faculty development programs to encourage faculty to adopt the teaching practices we found faculty were more likely to embrace, and to realize that patterns of change can look quite different across instructors, courses, and departments.

#### Abbreviations

CAUSE: Consortium for the Advancement of Undergraduate STEM Education; ICAP: Interactive, Constructive, Active, Passive; PORTAAL: Practical Observation Rubric to Assess Active Learning; STEM: Science, technology, engineering, and mathematics.

#### Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s40594-022-00340-4>.

**Additional file 1.** Supplementary figures and tables.

**Additional file 2.** Descriptive statistics of the 14 PORTAAL practices for each department at each time point.

**Additional file 3.** Average PORTAAL values for each paired course.

#### Acknowledgements

We thank the undergraduate researchers, E. Swanberg, P.J. Kambhiranond, A. Immel, C. Meng, and J. McAleer, who coded the class recordings. We also appreciate feedback on analyses from the Biology Education Research Group.

#### Authors' contributions

MAJ was responsible for the acquisition, analysis, and interpretation of data, and drafted the work. SM was responsible for the acquisition, analysis, and interpretation of data. JHD was responsible for analysis and interpretation of data, and substantively revised the work. MPW was responsible for conception and design of the work, the interpretation of data, and drafting the work. All authors read and approved the final manuscript.

#### Funding

This project was supported by the National Science Foundation (DUE 1725149). The funding body had no role in the design of the study and collection, analysis, and interpretation of data, and in writing the manuscript. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

#### Availability of data and materials

All data generated or analyzed during this study are included in this published article and its additional information files.

#### Declarations

##### Competing interests

The authors declare that they have no competing interests.

Received: 12 October 2021 Accepted: 22 February 2022

Published online: 04 March 2022

#### References

- Alberts, B. (2008). Considering science education. *Science*, 319(5870), 1589–1589. <https://doi.org/10.1126/science.1157518>
- American Association for the Advancement of Science. (2011). Vision and change in undergraduate biology education: A call to action. *Washington, DC*
- Anderson, W. A., Banerjee, U., Drennan, C. L., Elgin, S. C. R., Epstein, I. R., Handelsman, J., Hatfull, G. F., Losick, R., O'Dowd, D. K., Olivera, B. M., Strobel, S. A., Walker, G. C., & Warner, I. M. (2011). Changing the culture of science education at research universities. *Science*, 331(6014), 152–153. <https://doi.org/10.1126/science.1198280>
- Arthurs, L. A., & Kowalski, C. M. (2021). Engaging students' prior knowledge during instruction improves their learning of groundwater and aquifers. *Journal of Geoscience Education*. <https://doi.org/10.1080/10899995.2021.2004536>
- Association of American Medical Colleges-Howard Hughes Medical Institute. (2009). *Scientific foundations for future physicians* (Report of the AAMC-HHMI Committee, p. 46). Association of American Medical Colleges-Howard Hughes Medical Institute
- Austin, A. E. (2011). Promoting evidence-based change in undergraduate science education. In *Fourth committee meeting on status, contributions, and future directions of discipline-based education research*
- Baker, L. A., Chakraverty, D., Columbus, L., Feig, A. L., Jenks, W. S., Pilarz, M., Stains, M., Waterman, R., & Wesemann, J. L. (2014). Cottrell Scholars Collaborative New Faculty Workshop: professional development for new chemistry faculty and initial assessment of its efficacy. *Journal of Chemical Education*, 91(11), 1874–1881. <https://doi.org/10.1021/ed500547n>
- Bonwell, C. C., & Eison, J. A. (1991). *Active learning: creating excitement in the classroom*. 1991 ASHE-ERIC Higher Education Reports. ERIC Clearinghouse on Higher Education, The George Washington University, One Dupont Circle, Suite 630, Washington, DC 20036–1183
- Borda, E., Schumacher, E., Hanley, D., Geary, E., Warren, S., Ipsen, C., & Stredicke, L. (2020). Initial implementation of active learning strategies in large, lecture STEM courses: lessons learned from a multi-institutional, interdisciplinary STEM faculty development program. *International Journal of STEM Education*. <https://doi.org/10.1186/s40594-020-0203-2>
- 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(4), 339–346. <https://doi.org/10.1187/cbe.12-09-0163>
- Burnham, K. P., & Anderson, D. R. (2004). Multimodel inference: understanding AIC and BIC in model selection. *Sociological Methods & Research*, 33(2), 261–304. <https://doi.org/10.1177/0049124104268644>
- 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: successes and challenges. *Journal of College Science Teaching*, 49(2), 68–75.
- Canning, E. A., Muenks, K., Green, D. J., & Murphy, M. C. (2019). STEM faculty who believe ability is fixed have larger racial achievement gaps and inspire less student motivation in their classes. *Science Advances*, 5(2), eaau4734. <https://doi.org/10.1126/sciadv.aau4734>
- Carbone, A., Drew, S., Ross, B., Ye, J., Phelan, L., Lindsay, K., & Cottman, C. (2019). A collegial quality development process for identifying and addressing barriers to improving teaching. *Higher Education Research & Development*, 38(7), 1356–1370. <https://doi.org/10.1080/07294360.2019.1645644>
- Chi, M. T. H., & Wylie, R. (2014). The ICAP framework: linking cognitive engagement to active learning outcomes. *Educational Psychologist*, 49(4), 219–243. <https://doi.org/10.1080/00461520.2014.965823>



- Chien, Y.-T., Chang, Y.-H., & Chang, C.-Y. (2016). Do we click in the right way? A meta-analytic review of clicker-integrated instruction. *Educational Research Review*, 17, 1–18. <https://doi.org/10.1016/j.edurev.2015.10.003>
- Clarke, D., & Hollingsworth, H. (2002). Elaborating a model of teacher professional growth. *Teaching and Teacher Education*, 18(8), 947–967. [https://doi.org/10.1016/S0742-051X\(02\)00053-7](https://doi.org/10.1016/S0742-051X(02)00053-7)
- Committee on the Status, Contributions, and Future Directions of Discipline-Based Education Research, Board on Science Education, Division of Behavioral and Social Sciences and Education, & National Research Council. (2012). *Discipline-based education research: Understanding and improving learning in undergraduate science and engineering*. National Academies Press
- Cooper, K. M., Downing, V. R., & Brownell, S. E. (2018). The influence of active learning practices on student anxiety in large-enrollment college science classrooms. *International Journal of STEM Education*, 5(1), 23. <https://doi.org/10.1186/s40594-018-0123-6>
- Cooper, M. M., Caballero, M. D., Ebert-May, D., Fata-Hartley, C. L., Jardeleza, S. E., Krajcik, J. S., Laverty, J. T., Matz, R. L., Posey, L. A., & Underwood, S. M. (2015). Challenge faculty to transform STEM learning. *Science*, 350(6258), 281–282. <https://doi.org/10.1126/science.aab0933>
- Dallimore, E. J., Hertenstein, J. H., & Platt, M. B. (2010). Class participation in accounting courses: factors that affect student comfort and learning. *Issues in Accounting Education*, 25(4), 613–629. <https://doi.org/10.2308/iaee.2010.25.4.613>
- Dallimore, E. J., Hertenstein, J. H., & Platt, M. B. (2013). Impact of cold-calling on student voluntary participation. *Journal of Management Education*, 37(3), 305–341. <https://doi.org/10.1177/1052562912446067>
- Derting, T. L., & Ebert-May, D. (2010). Learner-centered inquiry in undergraduate biology: positive relationships with long-term student achievement. *CBE Life Sciences Education*, 9(4), 462–472. <https://doi.org/10.1187/cbe.10-02-0011>
- Derting, T. L., Ebert-May, D., Henkel, T. P., Maher, J. M., Arnold, B., & Passmore, H. A. (2016). Assessing faculty professional development in STEM higher education: sustainability of outcomes. *Science Advances*, 2(3), e1501422. <https://doi.org/10.1126/sciadv.1501422>
- Drinkwater, M. J., Matthews, K. E., & Seiler, J. (2017). How is science being taught? Measuring evidence-based teaching practices across undergraduate science departments. *CBE Life Sciences Education*, 16(1), ar18. <https://doi.org/10.1187/cbe.15-12-0261>
- Ebert-May, D., Brewer, C., & Allred, S. (1997). Innovation in large lectures: teaching for active learning. *BioScience*, 47(9), 601–607. <https://doi.org/10.2307/1313166>
- Ebert-May, D., & Weber, E. P. (2006). FIRST—What's next? *CBE Life Sciences Education*, 5(1), 27–28. <https://doi.org/10.1187/cbe.05-12-0133>
- Eddy, S. L., Brownell, S. E., Thummaphan, P., Lan, M.-C., & Wenderoth, M. P. (2015). Caution, student experience may vary: social identities impact a student's experience in peer discussions. *CBE Life Sciences Education*, 14(4), ar45. <https://doi.org/10.1187/cbe.15-05-0108>
- Eddy, S. L., Brownell, S. E., & Wenderoth, M. P. (2014). Gender gaps in achievement and participation in multiple introductory biology classrooms. *CBE Life Sciences Education*, 13(3), 478–492. <https://doi.org/10.1187/cbe.13-10-0204>
- Eddy, S. L., Converse, M., & Wenderoth, M. P. (2015). PORTAAL: a classroom observation tool assessing evidence-based teaching practices for active learning in large science, technology, engineering, and mathematics classes. *CBE Life Sciences Education*, 14(2), ar23. <https://doi.org/10.1187/cbe.14-06-0095>
- Elliott, E. R., Reason, R. D., Coffman, C. R., Gangloff, E. J., Raker, J. R., Powell-Coffman, J. A., & Ogilvie, C. A. (2016). Improved student learning through a faculty learning community: how faculty collaboration transformed a large-enrollment course from lecture to student centered. *CBE Life Sciences Education*, 15(2), ar22. <https://doi.org/10.1187/cbe.14-07-0112>
- Ferrare, J. J. (2019). A multi-institutional analysis of instructional beliefs and practices in gateway courses to the sciences. *CBE Life Sciences Education*, 18(2), ar26. <https://doi.org/10.1187/cbe.17-12-0257>
- 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(23), 8410–8415. <https://doi.org/10.1073/pnas.1319030111>
- Freeman, S., Haak, D., & Wenderoth, M. P. (2011). Increased course structure improves performance in introductory biology. *CBE Life Sciences Education*, 10(2), 175–186. <https://doi.org/10.1187/cbe.10-08-0105>
- 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). Scientific teaching. *Science*, 304(5670), 521–522. <https://doi.org/10.1126/science.1096022>
- Harvey, C., Eshleman, K., Koo, K., Smith, K. G., Paradise, C. J., & Campbell, A. M. (2016). Encouragement for faculty to implement Vision and Change. *CBE Life Sciences Education*, 15(4), e57. <https://doi.org/10.1187/cbe.16-03-0127>
- Henderson, C., Beach, A., & Finkelstein, N. (2011). Facilitating change in undergraduate STEM instructional practices: an analytic review of the literature. *Journal of Research in Science Teaching*, 48(8), 952–984. <https://doi.org/10.1002/tea.20439>
- 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(2), 020102. <https://doi.org/10.1103/PhysRevSTPER.3.020102>
- Herman, G. L., Greene, J. C., Hahn, L. D., Mestre, J. P., Tomkin, J. H., & West, M. (2018). Changing the teaching culture in introductory STEM courses at a large research university. *Journal of College Science Teaching*, 47(6), 32–38.
- Kahneman, D. (2011). *Thinking, fast and slow*. Macmillan.
- Kezar, A. (2011). What is the best way to achieve broader reach of improved practices in higher education? *Innovative Higher Education*, 36(4), 235–247. <https://doi.org/10.1007/s10755-011-9174-z>
- Kim, A. M., Speed, C. J., & Macaulay, J. O. (2019). Barriers and strategies: implementing active learning in biomedical science lectures. *Biochemistry and Molecular Biology Education*, 47(1), 29–40. <https://doi.org/10.1002/bmb.21190>
- Knight, J. K., Wise, S. B., & Southard, K. M. (2013). Understanding clicker discussions: Student reasoning and the impact of instructional cues. *CBE Life Sciences Education*, 12(4), 645–654. <https://doi.org/10.1187/cbe.13-05-0090>
- Knight, J. K., & Wood, W. B. (2005). Teaching more by lecturing less. *Cell Biology Education*, 4(4), 298–310. <https://doi.org/10.1187/05-06-0082>
- Lewin, J. D., Vinson, E. L., Stetzer, M. R., & Smith, M. K. (2016). A campus-wide investigation of clicker implementation: the status of peer discussion in STEM classes. *CBE Life Sciences Education*, 15(1), ar6. <https://doi.org/10.1187/cbe.15-10-0224>
- Lund, T. J., Pilarz, M., Velasco, J. B., Chakraverty, D., Rosploch, K., Undersander, M., & Stains, M. (2015). The best of both worlds: Building on the COPUS and RTOP observation protocols to easily and reliably measure various levels of reformed instructional practice. *CBE Life Sciences Education*, 14(2), ar18. <https://doi.org/10.1187/cbe.14-10-0168>
- McConnell, M., Montplaisir, L., & Offerdahl, E. (2020). Meeting the conditions for diffusion of teaching innovations in a university STEM department. *Journal for STEM Education Research*, 3(1), 43–68.
- Moon, S., Jackson, M. A., Doherty, J. H., & Wenderoth, M. P. (2021). Evidence-based teaching practices correlate with increased exam performance in biology. *PLoS ONE*, 16(11), e0260789. <https://doi.org/10.1371/journal.pone.0260789>
- Newman, T. (2017). Making an impact: utilising faculty learning communities to enhance teaching and learning. In J. McDonald & A. Cater-Steel (Eds.), *Communities of Practice* (pp. 423–435). Springer Singapore. [https://doi.org/10.1007/978-981-10-2879-3\\_20](https://doi.org/10.1007/978-981-10-2879-3_20)
- Olson, S., & Riordan, D. G. (2012). Engage to excel: producing one million additional college graduates with degrees in science, technology, engineering, and mathematics. Report to the president. *Executive Office of the President*.
- Petersen, C. I., Baepfer, P., Beitz, A., Ching, P., Gorman, K. S., Neudauer, C. L., Roza-itis, W., Walker, J. D., & Wingert, D. (2020). The tyranny of content: "Content coverage" as a barrier to evidence-based teaching approaches and ways to overcome it. *CBE Life Sciences Education*, 19(2), ar17. <https://doi.org/10.1187/cbe.19-04-0079>
- Pfund, C., Miller, S., Brenner, K., Bruns, P., Chang, A., Ebert-May, D., Fagen, A. P., Gentile, J., Gossens, S., Khan, I. M., Labov, J. B., Pribbenow, C. M., Susman, M., Tong, L., Wright, R., Yuan, R. T., Wood, W. B., & Handelsman, J. (2009). Summer Institute to improve university science teaching. *Science*, 324(5926), 470–471. <https://doi.org/10.1126/science.1170015>
- Prince, M. (2004). Does active learning work? A review of the research. *Journal of Engineering Education*, 93(3), 223–231. <https://doi.org/10.1002/j.2168-9830.2004.tb00809.x>

- R Core Team. (2021). *R: A language and environment for statistical computing* (4.1.1) [Computer software]. R Foundation for Statistical Computing. <https://www.R-project.org/>
- Reinholz, D. L., & Shah, N. (2018). Equity analytics: a methodological approach for quantifying participation patterns in mathematics classroom discourse. *Journal for Research in Mathematics Education*, 49(2), 140–177. <https://doi.org/10.5951/jresmetheduc.49.2.0140>
- Reinholz, D. L., Stone-Johnstone, A., & Shah, N. (2020). Walking the walk: using classroom analytics to support instructors to address implicit bias in teaching. *International Journal for Academic Development*, 25(3), 259–272. <https://doi.org/10.1080/1360144X.2019.1692211>
- Ruiz-Primo, M. A., Briggs, D., Iverson, H., Talbot, R., & Shepard, L. A. (2011). Impact of undergraduate science course innovations on learning. *Science*, 331(6022), 1269–1270. <https://doi.org/10.1126/science.1198976>
- Sbeglia, G. C., Goodridge, J. A., Gordon, L. H., & Nehm, R. H. (2021). Are faculty changing? How reform frameworks, sampling intensities, and instrument measures impact inferences about student-centered teaching practices. *CBE Life Sciences Education*, 20(3), ar39. <https://doi.org/10.1187/cbe.20-11-0259>
- Shadle, S. E., Marker, A., & Earl, B. (2017). Faculty drivers and barriers: laying the groundwork for undergraduate STEM education reform in academic departments. *International Journal of STEM Education*, 4(1), 8. <https://doi.org/10.1186/s40594-017-0062-7>
- Smith, M. K., & Knight, J. K. (2020). Clickers in the biology classroom: strategies for writing and effectively implementing clicker questions that maximize student learning. In J. J. Mintzes & E. M. Walter (Eds.), *Active learning in college science: the case for evidence-based practice* (pp. 141–158). Springer International Publishing. [https://doi.org/10.1007/978-3-030-33600-4\\_10](https://doi.org/10.1007/978-3-030-33600-4_10)
- Smith, M. K., Wood, W. B., Krauter, K., & Knight, J. K. (2011). Combining peer discussion with instructor explanation increases student learning from in-class concept questions. *CBE Life Sciences Education*, 10(1), 55–63. <https://doi.org/10.1187/cbe.10-08-0101>
- Stains, M., Harshman, J., Barker, M. K., Chasteen, S. V., Cole, R., DeChenne-Peters, S. E., Eagan, M. K., Esson, J. M., Knight, J. K., Laski, F. A., Levis-Fitzgerald, M., Lee, C. J., Lo, S. M., McDonnell, L. M., McKay, T. A., Michelotti, N., Musgrove, A., Palmer, M. S., Plank, K. M., ... Young, A. M. (2018). Anatomy of STEM teaching in North American universities. *Science*, 359(6383), 1468–1470. <https://doi.org/10.1126/science.aap8892>
- 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., ... Freeman, S. (2020). Active learning narrows achievement gaps for underrepresented students in undergraduate science, technology, engineering, and math. *Proceedings of the National Academy of Sciences*, 117(12), 6476–6483. <https://doi.org/10.1073/pnas.1916903117>
- Turpen, C., Dancy, M., & Henderson, C. (2016). Perceived affordances and constraints regarding instructors' use of Peer Instruction: Implications for promoting instructional change. *Physical Review Physics Education Research*, 12(1), 010116. <https://doi.org/10.1103/physrevphyseduces.12.010116>
- Viskopic, K., Ryker, K., Teasdale, R., Manduca, C., Iverson, E., Farthing, D., Bruckner, M. Z., & McFadden, R. (2019). Classroom observations indicate the positive impacts of discipline-based professional development. *Journal for STEM Education Research*, 2(2), 201–228. <https://doi.org/10.1007/s41979-019-00015-w>
- Waugh, A. H., & Andrews, T. C. (2020). Diving into the details: constructing a framework of random call components. *CBE Life Sciences Education*, 19(2), ar14. <https://doi.org/10.1187/cbe.19-07-0130>
- Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. <https://ggplot2.tidyverse.org>

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Submit your manuscript to a SpringerOpen® journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)