2-2018

Improving undergraduate STEM education: The efficacy of discipline-based professional development

Cathryn A. Manduca
cmanduca@carleton.edu

Ellen R. Iverson

Michael Luxenberg

R. Heather Macdonald
College of William and Mary, rhmacd@wm.edu

David A. McConnell

See next page for additional authors

Recommended Citation

Manduca, Cathryn A.; Iverson, Ellen R.; Luxenberg, Michael; Macdonald, R. Heather; McConnell, David A.; Mogk, David W.; and Tewksbury, Barbara J., Improving undergraduate STEM education: The efficacy of discipline-based professional development (2018). *SCIENCE ADVANCES*, 3(2). 10.1126/sciadv.1600193

Authors

Cathryn A. Manduca, Ellen R. Iverson, Michael Luxenberg, R. Heather Macdonald, David A. McConnell, David W. Mogk, and Barbara J. Tewksbury

SOCIAL SCIENCES

Improving undergraduate STEM education: The efficacy of discipline-based professional development

Cathryn A. Manduca,[1]* Ellen R. Iverson,[1] Michael Luxenberg,[2] R. Heather Macdonald,[3] David A. McConnell,[4] David W. Mogk,[5] Barbara J. Tewksbury[6]

We sought to determine whether instructional practices used by undergraduate faculty in the geosciences have shifted from traditional teacher-centered lecture toward student-engaged teaching practices and to evaluate whether the national professional development program On the Cutting Edge (hereinafter Cutting Edge) has been a contributing factor in this change. We surveyed geoscience faculty across the United States in 2004, 2009, and 2012 and asked about teaching practices as well as levels of engagement in education research, scientific research, and professional development related to teaching. We tested these self-reported survey results with direct observations of teaching using the Reformed Teaching Observation Protocol, and we conducted interviews to understand what aspects of Cutting Edge have supported change. Survey data show that teaching strategies involving active learning have become more common, that these practices are concentrated in faculty who invest in learning about teaching, and that faculty investment in learning about teaching has increased. Regression analysis shows that, after controlling for other key influences, faculty who have participated in Cutting Edge programs and who regularly use resources on the Cutting Edge website are statistically more likely to use active learning teaching strategies. Cutting Edge participants also report that learning about teaching, the availability of teaching resources, and interactions with peers have supported changes in their teaching practice. Our data suggest that even one-time participation in a workshop with peers can lead to improved teaching by supporting a combination of affective and cognitive learning outcomes.

INTRODUCTION

Transforming instruction in undergraduate science, technology, engineering, and mathematics (STEM) classrooms from a dominantly lecture-based, content-focused format to one in which students engage in learning concepts and processes of science is regarded as critical to the economic and cultural health of our nation (*1*, *2*). This shift requires efficient ways to promote transitions from a teacher-centered, lecture-based class model to active learning, student-engaged environments. One widely used strategy to facilitate this transformation is to provide professional development opportunities for faculty to learn about alternative approaches to teaching that have been shown to reduce attrition, improve student learning (*3*, *4*), and reduce the achievement gap among student populations (*5*, *6*). Professional development programs may focus on general attributes of change or be designed for faculty within their discipline. There is considerable debate about the efficacy of professional development programs (*7*–*9*) because the practices they promote have not been extensively adopted by faculty (*10*). Here, we characterize changes in undergraduate teaching practices in the geosciences using data collected from three national surveys between 2004 and 2012, a series of faculty interviews, and observations of classroom teaching in more than 200 classrooms. We discuss how these changes are related to participation in a national professional development program, On the Cutting Edge (hereinafter Cutting Edge).

Established in 2002, Cutting Edge (*11*, *12*) engages faculty in a community where they can share and discuss their teaching with peers and learn from experts. From 2002 to 2012, Cutting Edge was the only national-scale professional development program operating in the geosciences. As a result, we are able to explore the relationship between a specific national professional development program and teaching practices in a particular discipline.

Cutting Edge provides a unique opportunity to test the impact of a particular program on a substantial segment of all faculty who teach within a single STEM discipline at the undergraduate level in the United States. From 2002 to 2012, Cutting Edge offered 118 workshops and professional development events that were attended by more than 2000 unique U.S. geoscience faculty and 800 postdoctoral fellows and graduate students. We estimate that this represents approximately 20% of U.S. geoscience faculty. [The American Geosciences Institute (AGI) (*13*) estimates that there were approximately 8500 tenured or tenure-track geoscience faculty employed within U.S. geoscience departments at 4-year universities in 2013 and that this number had been relatively constant since 2008. Geoscience courses were also taught at approximately 400 two-year colleges in 2013, where there are relatively few geoscience faculty at each institution (*13*). We estimate that there are approximately 10,000 geoscience faculty teaching geoscience courses at U.S. higher-education institutions.] The workshop participants and project staff have developed a >5000-page website that summarizes workshop findings and provides access to a peer-reviewed collection of teaching activities and other related online resources (*11*, *12*). The website is viewed by approximately 1 million unique visitors per year, further expanding the program's reach.

The design of the Cutting Edge program arose from the leaders' experience with teaching and professional communities and was informed by research on teaching and learning. Rather than teaching a specific set of instructional skills, Cutting Edge workshops focus on engaging participants in a process of learning that leads to improvements in their teaching (*14*–*16*). Central design principles aim to foster

[1]Science Education Resource Center, Carleton College, Northfield, MN 55057, USA. [2]Professional Data Analysts Inc., Minneapolis, MN 55414, USA. [3]Department of Geology, College of William and Mary, Williamsburg, VA 23187, USA. [4]Marine, Earth, and Atmospheric Sciences Department, North Carolina State University, Raleigh, NC 27695, USA. [5]Department of Earth Sciences, Montana State University, Bozeman, MT 59717, USA. [6]Department of Geosciences, Hamilton College, Clinton, NY 13323, USA.
*Corresponding author. Email: cmanduca@carleton.edu

learning from peers (*16*, *17*), introduce ideas from cognitive science and education research (*8*, *18*), and engage participants in reflecting on applications to their own work (*19*). By facilitating networking at the workshops, supporting ongoing communications with email lists, and involving all in the creation of an online collection of resources, Cutting Edge incorporates participants into a geoscience education community (*20*).

In 2002, we initiated an evaluation of the Cutting Edge program to determine its impact. Then, in 2004, 2009, and 2012, we administered nationwide surveys (*21*, *22*) to college and university geoscience faculty. They answered questions about instructional and assessment practices in either an introductory or a majors course that they taught, as well as questions about the context of their professional work. We conducted interviews with 120 faculty at several points during the program (*22*, *23*) to investigate the impact of the workshops and websites. To evaluate the self-reports of instructional practice described in the surveys, we conducted classroom observations of more than 200 faculty using the Reformed Teaching Observation Protocol (RTOP) (*24*, *25*). Here, we consider these data in the context of current models for stimulating and supporting changes in teaching practice and address four research questions:

(1) Has there been measurable change in undergraduate geoscience instruction from teacher-centered lecture to student-engaged teaching practices?

(2) What role does learning about teaching play in supporting these pedagogical changes?

(3) Is faculty participation in Cutting Edge associated with use of student-engaged teaching practices?

(4) What impacts do participants recognize as coming from the workshops?

Theories have been developed to support studies of professional development in the fields of human resource development (*26*, *27*), continuing medical education (*28*), and educational development in both K-12 (*29*–*31*) and higher education (*9*, *15*, *16*). Although a unified framework has not developed, both learning and transfer of learning to the professional setting emerge as two distinct components necessary to affect performance. Characteristics of the individual, the nature of the learning opportunity, and the features of the professional setting where changes will be applied are important mediators in these models. We examine teaching practices in undergraduate geoscience classrooms and their relationship to activities that can be characterized as learning about teaching, including participation in the Cutting Edge workshop program and use of its website, in the context of these theories of professional development. We see evidence that the characteristics of individuals, particularly motivation and self-efficacy, are important, and we discuss the design of professional development opportunities with respect to both learning and support for transfer.

RESULTS

The three national surveys that we conducted in 2004, 2009, and 2012 investigated three questions that lie at the heart of teaching: (i) What methods do faculty use to teach undergraduate courses? (ii) How do faculty learn about the content and methods that they use in their teaching? (iii) How do faculty share with their colleagues what they learn about teaching? The surveys were administered by email to a list of all identifiable geoscience faculty in the United States. Of the approximately 10,000 geoscience faculty who teach undergraduates, 2207 faculty participated in the 2004 survey, 2874 participated in the 2009 survey, and 2466 participated in the 2012 survey. Of the 2009 survey respondents, approximately 565 completed the 2004 survey, and approximately 1111 of the 2012 survey respondents completed at least one previous administration (fig. S1). Our ability to identify faculty in 2-year colleges was substantially enhanced before the third administration of the survey in 2012; as a result, five times as many 2-year college faculty members (308 respondents) were included in the 2012 administration.

Survey data show movement away from lecture toward student-engaged teaching practices

We used two approaches in our faculty surveys to investigate whether there was movement away from traditional lecture. In each of the three survey administrations, faculty chose a specific introductory or majors course that they had taught in the previous 2 years and estimated (i) the amount of class time spent on student activities, questions, and discussion and (ii) the frequency with which they made use of specific teaching strategies during the lecture portion of the course. Both approaches indicate a self-reported movement away from lecture in both introductory courses and courses for majors between 2004 and 2012.

Estimates of class time spent on student activities, questions, and discussion.

We asked faculty to report the percentage of time spent on student activities, questions, and discussions (interactive time) in the lecture portion of the course. We classified the data into two groups: (i) more than 20% of class time was interactive, and (ii) 20% or less of class time was interactive. We chose 20% to differentiate significant use of interactive teaching because 20% represents a natural division in the data reported across survey administrations (fig. S2 and table S1). Twenty percent also aligns with reporting from similar surveys of STEM teaching practices (*32*, *33*). This classification is further supported by descriptions of active teaching (*34*, *35*), including a study in which the researchers observed group work occupying more than 20% of time in classes that they classified as using peer instruction or collaborative learning instructional styles (*36*).

The percentage of faculty reporting more than 20% of class time spent on student activities, questions, and discussions shows an increase over time in the population as a whole, from 34% in 2004, to 38% in 2009, to 51% in 2012 (Fig. 1 and Table 1). The Mantel chi-square for trend shows a linear association of the outcome across the three study periods [$\chi^2(1) = 86.6$; $P < 0.001$].

We analyzed the data by institution type to determine whether increases were due to the larger number of 2-year college faculty responses in the 2012 survey. Although a larger percentage of 2-year college faculty report courses with more than 20% interactive class time, a significant increase is also shown in faculty responses from other institutional types (table S2). This increase persists across subgroups defined by faculty type (table S3), teaching strategy (table S4), and course level (table S5).

Use of different teaching strategies during the lecture portion of a course.

We asked faculty to indicate the frequency with which they used the following teaching strategies in the lecture portion of the course: traditional lecture, lecture with demonstration, lecture in which questions posed by the instructor are answered by individual students, lecture in which questions posed by the instructor are answered simultaneously by the entire class, small-group discussion or think-pair-share, whole-class discussions, and in-class exercises. Responses indicating using a strategy on a weekly basis or in nearly every class were classified as frequent use; those indicating never using a strategy or using it once or several times were classified as infrequent use. An exploratory factor analysis of these responses from the combined 2004 and 2009 data suggests three principal groupings in the data on the basis of decreasing levels of student engagement:

(1) Active learning: faculty reporting frequent use of small-group discussion, whole-class discussion, or in-class exercises with or without the use of any other methods

(2) Active lecture: faculty reporting frequent use of demonstrations and/or posing questions with or without traditional lecture

(3) Traditional lecture: faculty reporting infrequent use of strategies other than traditional lecture

Across the three survey administrations, the data show a significant increase in the proportion of respondents who report using active learning teaching strategies (42% in 2004, 47% in 2009, and 57% in 2012; Fig. 1 and Table 1) and a decrease in the proportion of respondents who report using only traditional lecture methods (19, 13, and 11%, respectively; Table 1). Mantel chi-square tests indicate a linear trend across the 3 years [$\chi^2(1) = 89.9$; $P < 0.01$].

We computed Spearman's rank-order correlation coefficients using the 2012 data to determine the relationship between teaching strategy and other teaching practices that support student engagement (table S6). The data show a moderate correlation between teaching strategy and responses that indicate the following: (i) student exploration of concepts before receiving instruction [$r_s(1657) = 0.29$, $P < 0.001$]; (ii) student-student discussion of ideas, problems, or course content [$r_s(1659) = 0.36$, $P < 0.001$]; and (iii) students having an impact on the direction of the class [$r_s(1663) = 0.29$, $P < 0.001$]. The use of self-reported active learning teaching strategies correlates moderately with reporting of more than 20% interactive class time [$r_s(1711) = 0.46$, $P < 0.001$].

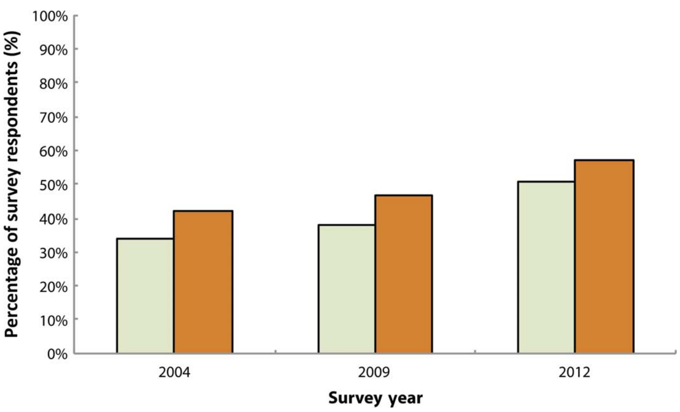

Fig. 1. Percentage of 2004, 2009, and 2012 survey respondents who report spending more than 20% of class time on student activities, questions, and discussions (light green bars) and percentage of respondents who report active learning (orange bars).

Faculty who invest in learning about teaching are more likely to use practices that support student engagement

To understand the relationship between teaching practices and professional behavior, we conducted a cluster analysis of responses to survey items addressing faculty engagement and participation in the geoscience community, including publications and presentations about teaching and research, participation in workshops, teaching or research talks attended, and communication with colleagues about course content and pedagogy. Using a clustering algorithm, we identified three groups of faculty within both the 2012 data alone and the combined 2004 and 2009 data:

(1) Education-focused faculty who reported significant activity related to improving teaching (their own and/or others)

(2) Geoscience research–focused faculty who reported significant geoscience research activity

(3) Teaching faculty who reported lower levels of activity in both geoscience research and activity related to improving teaching

Education-focused faculty made up the smallest percentage of 2012 survey respondents (20%), as compared to geoscience research–focused faculty (37%) and teaching faculty (43%) (table S7).

Although we recognize that some faculty are active in both improving teaching and conducting geoscience research and that all three types of faculty participate in workshops, an analysis of the cluster variable means in all three survey years shows strong differences in behavior across these three types of faculty (table S8). Education-focused faculty are strongly distinguished from other faculty types by the number of teaching-related talks and workshops that they attended in the last 2 years. Geoscience research–focused faculty are strongly distinguished by the number of articles and presentations on research that they have produced in the past 2 years (Fig. 2).

Analysis of teaching strategy by faculty type in 2012 (Fig. 3 and table S9) shows that although active learning teaching strategies are reported by at least half of the faculty from all types, these strategies are reported more frequently by the education-focused faculty. Education-focused faculty also more frequently report that more than 20% of class time was interactive (table S9). When compared to other faculty types (table S10), education-focused faculty are also statistically significantly more

Table 1. Interactive class time and teaching strategies versus survey year.

Survey year	Class time spent on student activities, questions, and discussions: Less than or equal to 20% (*n*)	More than 20% (*n*)	Total (*n*)	Teaching strategy: Traditional lecture (*n*)	Active lecture (*n*)	Active learning (*n*)	Total (*n*)
2004	1035	531	1566	300	611	657	1568
2009	1255	775	2030	256	816	945	2017
2012	873	893	1766	185	560	996	1741

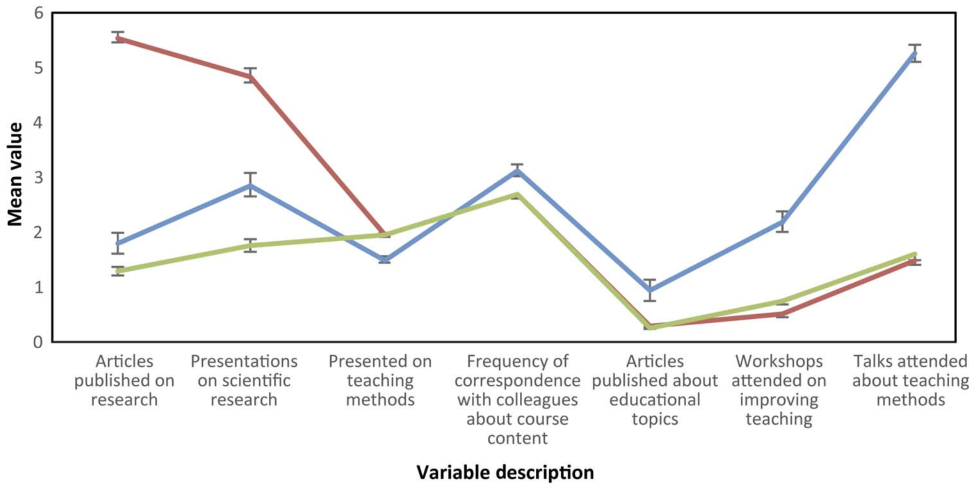

Fig. 2. Pattern of cluster variable mean values for the three faculty types demonstrating strong differences in reported behavior between groups; data were combined across all three survey years (2004, 2009, and 2012). The blue line represents education-focused faculty, the red line represents geoscience research–focused faculty, and the green line represents teaching faculty. Error bars represent 95% confidence interval (table S8).

likely to report other behaviors that support student engagement in learning than geoscience research–focused faculty or teaching faculty:

(1) Guided data analysis (71% education-focused faculty, 61% geoscience research–focused faculty, and 60% teaching faculty; χ^2 = 44.62; $P < 0.001$)

(2) Student data collection to solve a problem (53, 46, and 44%; χ^2 = 30.66; $P < 0.001$)

(3) Problems of national or global interest (64, 52, and 52%; χ^2 = 53.51; $P < 0.001$)

(4) Problems of local community interest (46, 34, and 32%; χ^2 = 49.23; $P < 0.001$)

(5) Study skills or time management reflection (45, 26, and 35%; χ^2 = 15.42; $P < 0.001$)

(6) Learning success reflection (56, 39, and 40%; χ^2 = 59.79; $P <$ 0.001)

(7) Problem-solving strategies reflection (52, 43, and 39%; χ^2 = 52.24; $P < 0.001$)

(8) Student motivation (79, 66, and 66%; χ^2 = 74.58; $P < 0.001$)

(9) Low-stakes opportunities for testing (70, 55, and 58%; χ^2 = 28.95; $P < 0.001$)

(10) Time for students to get to know each other (65, 48, and 51%; χ^2 = 84.13; $P < 0.001$)

(11) Teamwork and collaborative learning (76, 56, and 60%; χ^2 = 86.22; $P < 0.001$)

Faculty classified as education-focused come from all types of institutions and from all experience levels (tables S11 and S12). On average, education-focused faculty have teaching loads and class sizes similar to those of teaching faculty (table S11). Geoscience research–focused faculty come from all types of institutions, although they are the most common type of faculty only at doctoral-granting institutions. On average, geoscience research–focused faculty report spending about half as much time teaching in class or laboratory as other groups (table S11).

Faculty investment in learning about teaching is increasing

Participation by both teaching faculty and geoscience research–focused faculty in education workshops and presentations increased significantly between 2009 and 2012 (table S13). Whereas 100% of education-focused faculty attended three or more education talks in the 2 years before the 2012 survey, the percentage of geoscience research–focused and teaching faculty reporting attendance at three or more talks increased from 23 and 28%, respectively, to 37 and 43% between 2009 and 2012. Similarly, the percentage of faculty who reported attending one or more teaching workshops in the 2 years prior to the 2012 survey increased from 25 and 36% to 38 and 52%.

Participation in Cutting Edge is associated with teaching that promotes student engagement

We developed regression models for the 2012 data to investigate the impact of the Cutting Edge program on the use of interactive class time (table S14) and teaching strategies (table S15). Controlling for all variables in the final models, Cutting Edge workshop participants who also make use of the website are 1.5 times more likely to spend more than 20% of class time on activities, questions, or discussion and 1.2 times more likely to report teaching using an active learning teaching strategy, as opposed to an active lecture teaching strategy, than respondents who had neither used the website nor attended a Cutting Edge workshop. The presence of a laboratory (with or without a discussion section) was a significant predictor of interactive class time and teaching strategies when compared to the absence of a laboratory. The only other significant covariate retained in the active learning model was faculty type: Education-focused faculty were more likely to report active learning than teaching faculty. In contrast, the likelihood of reporting more than 20% of class time on activities other than lecture is significantly higher not only for education-focused faculty (versus teaching faculty) but also for early-career faculty (versus senior faculty), for faculty at 2-year colleges (versus 4-year colleges), and in small and medium classes (versus large classes) and courses for majors (versus introductory courses).

For the analyses described above, we used the 2012 survey data set because it is the most current and because it includes the broadest set of Cutting Edge participants. A regression model developed using 2009 data yielded similar results—Cutting Edge participants who attended workshops and made use of the website were 1.5 times more likely to spend more than 20% of class time on activities than those who did not participate with Cutting Edge. We did not develop

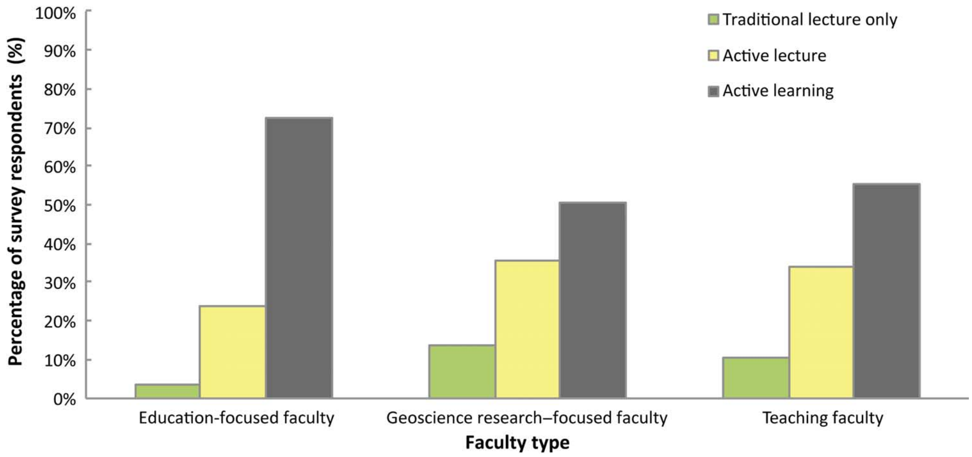

Fig. 3. Percentage of 2012 survey respondents (n = 1642) within each cluster (education-focused faculty, geoscience research–focused faculty, and teaching faculty) classified as reporting traditional lecture (green), active lecture (yellow), and active learning teaching strategies (gray).

a model using the 2009 data to predict active learning relative to active lecture.

Because survey findings using self-reported data may not accurately reflect practice (*7*, *37*), we collected additional data to test whether observed teaching practices among a subpopulation of geoscience instructors corresponded to the reported behavior of the whole population. We conducted 205 in-person observations of classroom teaching using the RTOP (*25*). Faculty type could be determined from survey responses for more than half (n =110) of these observations. Cutting Edge workshop participants who also use the website were observed to teach using more elements of active learning than instructors who had no involvement with Cutting Edge or who only used the online resources. They earned higher RTOP scores, and this finding confirmed the survey findings in aggregate (Fig. 4 and table S16), that is, that participation in Cutting Edge was associated with more extensive use of student engagement activities in class.

Cutting Edge participants report that learning about teaching, availability of teaching resources, and interactions with peers support changes in teaching practice

In 2013, Rockman et al was contracted to conduct an independent analysis of 71 transcripts from interviews with Cutting Edge workshop participants who had also submitted a full response to the survey in any of the years it was administered. The interviews were conducted as part of several studies by the project evaluators between 2002 and 2012. These studies probed the impact of the Cutting Edge program on participants and their teaching practices (*21*, *22*). In most cases, the interview data predate the survey responses, with 39 cases (55%) using 2012 survey data for categorization. We therefore do not know whether the faculty type assigned using the 2012 survey data is an accurate description of a participant's practices at the time that the interview was conducted. The sample includes a higher percentage of education-focused faculty (45%) than either the survey population as a whole (20%) or the Cutting Edge population (32%) (table S17).

Rockman et al coded the interview responses for evidence of practices queried in the survey (the coding guide is presented in the Supplementary Materials). Responses reporting specific changes in teaching practice included descriptions indicative of active learning or active lecture teaching strategies coded as

(1) Pedagogies related to student engagement and development of metacognitive skills

(2) Activities where students are engaged in independent thinking, reflection, and problem solving

(3) Activities that have places for students to assess their own learning

(4) Activities that engage students in data collection and analysis, observation, and experimentation

(5) Lecture with demonstration

(6) Lecture that encourages class discussion

(7) Usage of small-group discussion or think-pair-share

Nearly three-quarters of the interviewees (51 cases) reported on specific changes they had made to their teaching practice. In nearly half of the interviews (34 cases), faculty stated how the Cutting Edge experience had given them the knowledge and skills necessary to make these changes. In the interviews, faculty reported two ways in which the Cutting Edge experience supported the transfer of workshop learning to changes in practice. First, more than half of the participants (39 cases) stated that they were able to use Cutting Edge web resources to continue to build on their new understanding of teaching and learning. They described how they were currently using activities and approaches gained through the workshops or found through the Cutting Edge website in their teaching, although interviews followed participation in Cutting Edge workshops by as many as 3 years.

Second, faculty noted that interactions with colleagues during a Cutting Edge workshop were critical to putting changes in teaching into practice. In 73% of the interviews (52 cases), faculty stated how discussions with participants helped them leave the workshop with renewed motivation and enthusiasm to incorporate changes into their teaching practices. Faculty described how conversations about teaching at the workshop gave them ideas related to implementation and assessment of new strategies. They valued the opportunity to talk with others who were equally interested in improved teaching. More than half of the faculty (33 cases) reported on the value of these conversations at two or more points during the interview. In half of the interviews (24 cases), participants also stated that collegial interactions gave them greater confidence in their abilities to broadly use active learning strategies and to implement specific methods or activities. In a few cases, faculty stated that occasional communications by email or phone

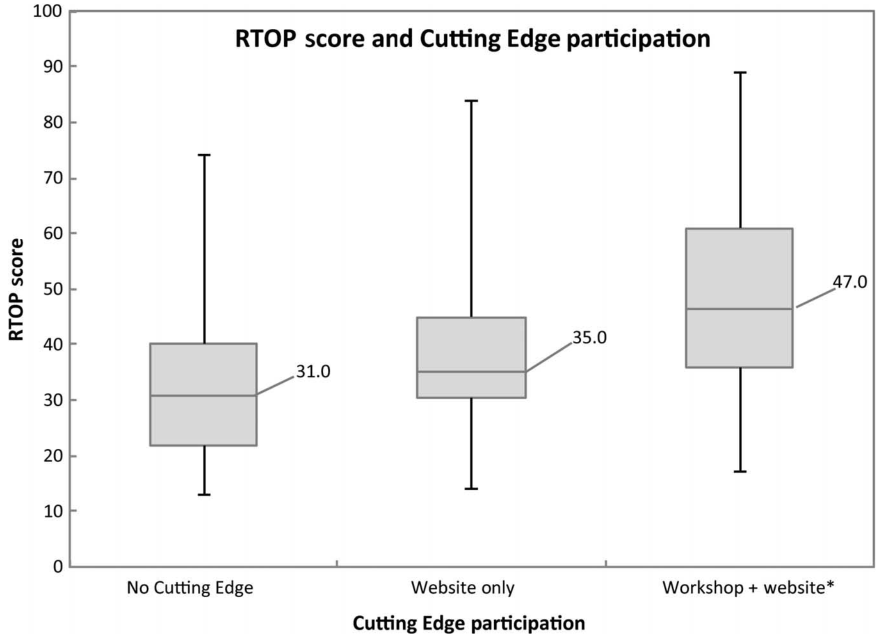

Fig. 4. RTOP scores obtained from observations of faculty who had (i) neither participated in a Cutting Edge workshop nor used the website, (ii) used the website only, or (iii) both used the website and participated in a Cutting Edge workshop. The full range of scores for each group is indicated by a thin vertical line. The interquartile range (approximately 50% of scores) is represented with a box with a horizontal line delineating the median. *Bonferroni post hoc tests from an analysis of variance (ANOVA) test ($F = 22.6$; $P < 0.001$) indicate a significant difference in mean RTOP scores between group 3 [those who use the Cutting Edge website and attend workshops ($M = 48.2$, SD = 16.2)] and the other two groups [group 1 ($M = 33.1$, SD = 13.6); $P < 0.001$; group 2 ($M = 37.2$, SD = 13.1); $P < 0.001$].

following the workshop had been useful in supporting their changes in practice.

Nearly half of the interviews (31 cases) reported shifts in attitudes about teaching. Faculty stated how they now realized that teaching is a learned skill supported by a research-based literature. They described this shift as an "aha moment" or transformative, suggesting a substantive change in their teaching beliefs. Faculty discussed this shift in conjunction with the description of their more student-centered teaching approaches. Some of these faculty reported that this new understanding affirmed and validated their efforts toward improved teaching.

Limitations of the study

The study design is limited, in that there was no randomized assignment. The population for the study is based on responses to a national survey that was administered as a census sample to all identifiable geoscience faculty. We did not do a study of nonresponders to determine how representative the responses are of the population at each time frame. Thus, the responses at each point in time do not allow us to distinguish changes in the aggregate behavior over time from changes in the responding population. Although we chose the smaller subsample populations for observation and interviews to include faculty representing different dimensions of interest, caution must be exercised in any generalizations because of the small population of faculty who were observed or interviewed relative to the population of U.S. geoscience faculty.

The timing of the different methods of investigation introduces challenges in assessing internal consistency between methods. Most of the individuals from our qualitative cases participated in an interview 2 to 7 years earlier. This sequencing confounds the factors of time and methods. The RTOP observations were conducted closer to the time of a participant's survey response and therefore do not present the same problem.

Because individuals report on their teaching in different courses, the study design does not support comparison of self-reported individual performance through time. Therefore, we are unable to obtain a direct pre-post measure of Cutting Edge influence on an individual's teaching. The influence of preexisting elements, such as personal preference for active learning teaching methods, can only be inferred by comparison among groups. Additionally, there may be a bias in the responses tending to overestimate improved teaching. All of the invitations for the study methods (survey, observations, and interviews) were proffered by faculty associated with the Cutting Edge program. Consequently, the respondents may preferentially include faculty who personally know these inviters. These faculty may be more likely either to have improved their teaching or to have overreported their attention to teaching. We did compare the demographics of the 2004 respondents to the demographics of the overall population, and comparison showed minimal bias (*20*). Although the number of Cutting Edge workshop participants has grown and their representation in the survey population has increased, the response rate for this group has not increased (table S18), arguing against a significant response bias.

DISCUSSION

Student-engaged teaching practices are associated with learning about teaching

In 2005, Macdonald *et al.* (*21*) reported that the 2004 survey showed widespread but not extensive use of engaged teaching methods by geoscience faculty. Our study shows that the use of those methods has increased since the first survey administration. A similar result is documented in the Higher Education Research Institute (HERI) survey of faculty across all disciplines during the interval 2004–2014, in which fewer faculty reported extensive lecturing in all classes (*38*). Disaggregation of the HERI data shows that this movement was smaller in STEM classrooms than in social science and humanities classrooms. However, the percentage of physical science faculty (physical science includes geoscience in the HERI report) using extensive lecturing in all or most courses dropped from approximately 80 to 66%. At the same time, the reported use of student inquiry to drive learning increased from 25 to 35% (*38*). The role of professional development experiences is unexplored in these studies.

In our study, we find that although engaged teaching methods are used by some members of all three faculty groups, these methods are reported most frequently by education-focused faculty—the group of faculty whose behavior is typified by attending workshops and talks about teaching. That is, the adoption of teaching practices that have been demonstrated to improve student learning is more frequently associated with instructors who invested more in learning about teaching. This population forms approximately 20% of our sample and is drawn from a broad cross section of types of institutions.

In their study of faculty at two institutions of higher education, Condon *et al.* (*39*) observed that faculty engaged in a long-term, self-guided process of learning about teaching and improving their practice. These faculty capitalized on a variety of professional development opportunities to learn and integrated this learning with information and ideas from other sources. They drew on this growing body of knowledge and skills over time to solve the challenges they faced in fostering learning in their courses. This behavior is very similar to the way that scientists, including science faculty, learn in their disciplinary specialties to inform their research. We suggest that our data are recording this type of behavior in education-focused geoscience faculty nationwide.

One of the important factors in models and theories addressing professional learning and its application to practice is the motivation of the learner. Although motivation to learn and motivation to transfer learning to practice are frequently considered separately (*26*, *28*, *40*), Naquin and Holton (*41*) explored a combined construct—motivation to improve work through learning (MTIWL)—that is aligned with the behavior observed by Condon *et al.* (*39*). We suggest that MTIWL is a valuable construct to understand the correlation between professional development and improved teaching in our data, because it combines both the motivation to seek out learning opportunities and the motivation to use this knowledge to change teaching. We suggest that education-focused faculty in our sample have more MTIWL as applied to their teaching than other faculty in our sample. Their repeated participation in professional development opportunities indicates that they are actively seeking opportunities to learn about teaching, and we infer that the stronger teaching methods they report reflect the use of this learning in their teaching.

Viewed in this light, an association between learning and improved teaching practice is not congruent with the belief that professional development has no impact on faculty who are motivated to change their courses ("they would have used these methods anyway"—a frequently heard criticism in studies of the impact of professional development programs). Rather, it appears that professional development is better understood as providing important opportunities to learn for those who are motivated to learn and that professional development is used by these individuals to improve practice.

Our data suggest that additional work to establish the relationship between MTIWL, repeat participation in professional development opportunities of both short and long duration, and teaching practice is warranted and could substantially move forward the ability to transform teaching in higher education. It is possible that improvements in teaching practice over time, as recorded in our data and in the HERI data, reflect increased faculty participation in professional development around teaching and their use of this new knowledge in their teaching. Our study design does not support direct testing of this important hypothesis.

The average number of teaching workshops and talks attended was reported to increase over time in the geoscience research–focused and teaching faculty groups. This finding might reflect any combination of increased availability of professional development offerings, an increased value of learning about teaching by faculty, and changes in the work environment and performance expectations. More research is warranted to determine whether the documented increase is associated with an increase in MTIWL within this population and what, if any, role the large-scale, long-lasting Cutting Edge program played in this increase. Regardless of cause, the self-reported increase in attendance at teaching workshops and talks is an encouraging finding, suggesting that faculty investment in learning about teaching is increasing.

Teaching practices and impact of the Cutting Edge program

Regression analyses of the survey data indicate that Cutting Edge workshop participants of all faculty types are more likely to report engaged teaching practices than those who do not participate in the program. Cutting Edge participants within the teaching faculty group, the group receiving the smallest amount of professional development, show significantly more engaged teaching, suggesting that even limited participation has an effect. Cutting Edge participants within the education-focused faculty group–the population of faculty who participate most frequently in professional development activities beyond Cutting Edge–also exhibit more characteristics of engaged teaching than nonparticipants within this group, suggesting that Cutting Edge participation also benefits those with substantial incoming knowledge. Our observation study confirms that Cutting Edge workshop participants, as a group, have stronger teaching practices than nonparticipants, as measured by the RTOP protocol. We do not have sufficient information or sample size to subdivide the RTOP population by faculty type.

In both of our regression models, the outcome variable asks only about teaching in the lecture portion of the course. One might thus expect the results to be insensitive to the presence or absence of a laboratory or discussion section in a course. This is not the case. The presence of laboratory and/or discussion sections is retained as a significant variable in both models. This may indicate that respondents misread the question and are reporting on all aspects of a course, or it may indicate that the presence of a laboratory and discussion changes the dynamics and instructional style in the lecture portion of a course.

The active learning model retained no significant variables other than faculty type, indicating that reporting of active learning strategies versus active lecture strategies does not depend on variables such as class size, course level (introductory or majors), or institution type. Active learning classrooms are characterized by the use of discussion and in-class activities. The result suggests that these methods are being adapted in classrooms of all types and sizes.

In contrast, the model for interactive class time retained several other significant variables, including variables that are descriptive of the course, the faculty member, and the institution. Consistent with perceptions of courses for majors, faculty describing courses for majors are more likely to report more than 20% interactive class time, as are faculty describing small- or medium-sized courses (less than 80 people). Thus, although faculty report use of discussions and activities in all types of settings, the amount of time spent on activities other than lecture is dependent on course size and type.

Early-career faculty who have been teaching for 7 years or less, as well as 2-year college faculty, are more likely to report more than 20% interactive class time. This finding deserves further exploration using observations of teaching practice. The use of less lecture by early-career faculty could reflect a wide variety of influences, including increased opportunities for graduate students to learn about teaching, the impact of early-career mentoring through Cutting Edge or other programs, changing hiring practices, or faculty values and expectations.

Although a number of teaching practices that engage students are exhibited more frequently by education-focused faculty, our data show relatively weak correlations among the different indicators of engaged teaching. Similarly, the qualitative data indicate that individuals are making a wide variety of changes as a result of their participation in Cutting Edge. These data are not surprising given the broad set of topics addressed by the workshop program and the absence of a strong focus on promoting a single teaching strategy or set of outcomes. These results are consistent with a model of evolutionary change or continuous improvement in teaching (*42*, *43*).

Our qualitative data provide insight into the aspects of Cutting Edge program design that may underpin the strong teaching performance of participants. Three impacts are reported: (i) information and resources from Cutting Edge supported changes in teaching, (ii) the program catalyzed a change in attitude or belief about teaching, and (iii) interactions with colleagues resulted in renewed energy for change and, in some cases, increased confidence and self-efficacy. These results are based on the qualitative data obtained from individuals classified as members of each of the three faculty types but are dominated by responses obtained from education-focused faculty.

Models and theories describing professional learning draw a distinction between learning at a workshop and transfer of learning to the work environment (*26*, *27*, *38*, *40*, *44*). The first is necessary but not sufficient to support the second. Learning at a workshop is dependent on a strong pedagogic design that attends to both the social and cognitive aspects of learning. Transfer of learning to practice is mediated by aspects of both the individual and the work environment. Although the Cutting Edge program did not have a direct impact on the work environment, all of the impacts reported support the ability of individuals to transfer learning to practice, as described below. In addition, many Cutting Edge workshops, particularly those that addressed career stage preparation (Early Career Faculty Workshop and Preparing for an Academic Career Workshop), addressed the ability of the individual to shape or influence the work environment.

Approximately half of faculty in the qualitative study indicated that their current teaching included activities and approaches that they learned from the Cutting Edge program. This finding suggests substantial transfer of workshop content into practice. Several features of the program design are aligned with practices that have been identified as important for supporting learning and transfer to the classroom. Cutting Edge workshop design (*11*, *14*) emphasizes abundant examples of application of theory to geoscience teaching, coupled with workshop activities that support participants in applying what they have learned to their own professional context (*16*). This practice is aligned with the development of metacognition and capacity for transfer through reflection (*19*) and capitalizes on the use of peers as trusted sources of information (*17*, *45*). It provides the "how to" information needed to transform ideas into practice (*33*, *45*, *46*). Further, it helps faculty to define a realistic path from where they are to where they would like to be, perhaps increasing the chance of persistent change, an ongoing problem in both K-12 and higher education (*29*, *33*, *45*, *46*). Sharing of resources is facilitated by the Cutting Edge website and supported through the website, which organizes access to all resources presented and discussed at the workshop by author and presentation time.

We interpret the remaining outcomes, namely, change in belief, renewed energy for change, and increased confidence and self-efficacy, to be related closely to the use of peer instruction in the workshop design. Cutting Edge workshops are designed to bring faculty together to learn from one another and from outside experts about a shared challenge (*11*, *14*). This productive learning situation counteracts, both pedagogically and emotionally, the tendency to pigeonhole faculty as good or bad teachers (*16*, *41*) and supports changes in teacher beliefs (*47*). Although Cutting Edge workshop experiences lasted only a few days, peer interactions supported many of the effects identified as central features of communities of practice (*48*): learning by problem solving, sharing information, and seeking expertise. Renewed energy for change was also identified as an outcome of peer interaction by Kezar and Gehrke (*20*) in their study of communities of transformation for undergraduate faculty.

The learning transfer system and associated Learning Transfer System Inventory developed by Holton *et al.* (*26*, *49*) provide a framework for understanding the relationship between peer instruction and transfer observed in our study. The Learning Transfer System Inventory identifies 16 constructs that affect transfer, including factors that are features of the individual and of the transfer environment. Four constructs appear related to peer instruction and interaction:

(1) Perceived content validity (the extent to which trainees judge training content to accurately reflect job requirements). This construct could be understood as the extent to which faculty believe that teaching strategies they are learning about will work in their own classrooms. The Cutting Edge workshop model features faculty sharing their own teaching successes with one another and includes opportunities to discuss implementation details.

(2) Performance-outcomes expectations (the expectation that changes in job performance will lead to valued outcomes). Participants in the Cutting Edge program are investing in good teaching for the purpose of increased student learning. Peer instruction allowed participants to learn from a trusted source about the relationship between change in teaching and change in student learning. This interpretation emphasizes alignment with the participants' own values, rather than those of the employer.

(3) Peer support (the extent to which peers reinforce and support use of learning on the job). The strong emphasis that interviewees placed on peer interactions indicates the value of finding peer reinforcement at the workshops. This effect is described in depth in the qualitative study of communities of transformation by Kezar and Gehrke (*20*).

(4) Performance self-efficacy (the general belief that individuals are able to change their performance when they want to). Participants self-reported this change as a result of peer interactions.

An area for further study is the relationship between Cutting Edge workshop participation, use of the Cutting Edge website, and development of a sense of belonging in a community that supports the ongoing interaction, leadership development, and exchange of ideas characteristic of communities of transformation (*50*). Our qualitative data show that, for many participants, the website provides information that supports the implementation of changes in teaching practices. The Cutting Edge program has fostered a culture and an ethic of sharing experiences and resources (*11*). Teaching activities and instructional resources are community-built and peer-reviewed, aggregating the practitioners' wisdom within the community. We do not know the extent to which faculty learn from the website nor do we know that their contributions to the website engender a sense of belonging to the Cutting Edge community of geoscience faculty.

Approximately one-half of survey responses from Cutting Edge participants and slightly more than one-third of interviews used in this study were obtained from individuals who had participated in more than one Cutting Edge workshop or webinar. Overall, approximately 25% of workshop participants have attended more than one workshop or webinar. We have not explored the role of repeat participation in this study, nor do we understand how the influence of participation in Cutting Edge, once or repeatedly, depends on the levels of other opportunities to learn about teaching within the geosciences, within STEM, or within the institution or higher-education community more broadly. For education-focused faculty, participation in the Cutting Edge program is a part of a larger picture of ongoing learning about teaching, resulting in strong teaching practices. This is an area where further study is warranted.

However, our data do not support an interpretation that relies on repeat participation, within the Cutting Edge program or participation in this program in conjunction with others, to explain the association of workshop participation with active learning teaching practices that support student engagement. This finding calls into question claims that professional development experiences must involve interactions over 1 month in duration to be effective (*9*) and that change in practice requires a program design that repeatedly brings participants together over a protracted interval of time (*17*).

Implications and next steps

This study adds to the growing body of evidence that investments that support faculty learning about teaching lead to improvements in teaching practice (*17*, *39*, *50–53*). Faculty who invest time in learning about teaching report the strongest teaching practices. Although more research is needed to understand the relationship between learning about teaching and improving teaching practice, it is clear that opportunities to learn have an impact.

Given the evidence from this study and that of Condon *et al.* (*39*) that faculty engage in multiple professional development experiences, it appears that a productive line of investigation could focus on faculty members as independent, self-guided learners who seek opportunities and interactions that support their abilities to learn about teaching to improve teaching practice. In this frame, the following important questions are posed:

(1) What develops faculty motivation to improve their teaching through learning?

(2) What is the relationship between learning opportunities, development and implementation of higher-order thinking about teaching and learning, and capacity to improve instruction and adapt to changing instructional needs and contexts?

(3) How do individuals navigate and capitalize on the variety of learning opportunities on- and off-campus to maximize their learning and improvement in practice?

(4) What are the characteristics of a system of activities that support learning for the diverse faculty in higher education? How can this system be optimized?

Given how little we understand about this larger learning landscape, as well as the relatively weak understanding of the relationship between professional development and practice (*9*, *27*, *28*, *54*), it seems premature to make claims that all professional development should be of any particular type or duration. Our qualitative data suggest that even one-time participation in a workshop with peers can lead to improved teaching by supporting a combination of affective and cognitive learning outcomes. More information is needed about how this type of one-time, focused, peer-learning experience interacts with other kinds of learning on- and off-campus, in the short and long term, to support the development of a robust teaching capacity. This information is particularly needed at a time when overall access to learning opportunities and online resources appears to be increasing. It may be that different types of programs can serve different purposes for faculty who have developed more or less motivation to improve teaching through learning. For example, short activities might be used to interest faculty in making larger changes supported by longer, more intensive programming. Short activities supporting continuous improvement might complement intensive programming that supports rapid change in ways that maximize the overall capacity of the system to produce well-informed faculty and adaptive teaching practices.

However we seek to maximize our investments, undergraduate geoscience teaching improved in response to the learning opportunities available between 2002 and 2012, including those from the Cutting Edge program. Over this interval, geoscience faculty invested increasing amounts of time in learning about teaching, and learning made a difference in their teaching practices. These are encouraging findings in a time when geoscience learning is essential to our ability to live sustainably on Earth.

MATERIALS AND METHODS

We investigated to what extent faculty use methods other than traditional lecture in geoscience undergraduate instruction and the role that learning about teaching plays in supporting incorporation of active learning strategies into instruction. In addition, we examined whether participation in the Cutting Edge program was associated with engaged teaching practices and associated with the outcomes recognized by the participants. We tested our research questions through the analysis of three national surveys combined with purposive sampling of geoscience faculty using both observational and qualitative methods. Institutional review board guidelines were followed with human subjects.

Survey

The survey, administered in 2004, 2009, and 2012, was developed by Cutting Edge, a professional development program for geoscience faculty sponsored by the National Association of Geoscience Teachers (NAGT) and funded by grants from the National Science Foundation.

Instrument.

The survey instrument was developed in 2003 by the leadership of Cutting Edge in collaboration with project evaluator J. McLaughlin and the Statistical Research Group at the American Institute of Physics (AIP) (*21*). This group modified the instrument in 2009 on the basis of the results of the 2004 administration. In 2012, the survey was modified by the project leadership in consultation with Professional Data

Analysts Inc., who were contracted to complete data analysis of the 2009 survey, administer the 2012 survey, and analyze the results. Although some items of the survey underwent changes between each iteration, the general characteristics of the surveys remained essentially consistent. The survey items for all three administrations can be viewed from the Cutting Edge Evaluation Summary Web page: http://serc.carleton.edu/NAGTWorkshops/about/evaluation.html.

The items for the 2004 survey were tested for clarity in a pilot survey administered to 16 faculty, as well as through short interviews with 5 faculty at the American Geophysical Union Fall Meeting in 2002. For the 2009 survey, revised items were tested using a written survey combined with interview responses from 37 faculty at the American Geophysical Union Fall Meeting in 2007. For the 2012 survey, expert reviews and think-aloud administrations were conducted with four faculty.

The survey, which has three parts, follows a similar structure across the years.

(1) In the first section, faculty answer demographic questions about education and teaching experience, disciplinary focus, and faculty position and teaching responsibilities.

(2) In the second section, faculty answer questions about a specific introductory course or course for majors taught in the previous 2 years. Questions ask about the design of the course, teaching methods and strategies, content, and assessment. Faculty teaching only introductory courses or only courses for majors were asked to describe a course of that type. Faculty teaching both types of courses were randomly assigned to describe one or the other. In the first two survey administrations, faculty were allowed to repeat this section and describe as many courses as they chose.

(3) The third section asks questions focused on how faculty learn about content and teaching methods, how they put this learning to use in designing courses, and how they share information about teaching with colleagues.

Survey population.

We administered surveys to all identifiable geoscience faculty in the United States. We began with lists developed by the AGI that included full-time faculty, part-time and adjunct faculty, instructors, lecturers, research scientists, and other educators teaching undergraduate geoscience courses. The 2004 survey was emailed to approximately 5700 faculty and received 2207 responses, a response rate of 39% (*21*).

We electronically sent the 2009 survey in March 2009 to 5107 geoscience faculty and received 2537 completed responses (50% response rate). For the 2009 survey, we discovered that at least 810 geoscientists who had been past participants in Cutting Edge workshops were not on the original survey invitation list compiled from AGI data. We sent a second wave of invitations to these 810 geoscientists in September 2009 and received 337 responses. This brought the 2009 survey combined total to 2874 responses from 5917 emails (49% response rate). In comparing the original survey sample with the supplementary sample, individuals in the supplementary group were far more likely to have a master's degree as their highest degree, were far more likely to have earned degrees within the previous 5 years, and were less likely to be teaching. For questions about teaching practice, the analysis between the two pools of survey respondents demonstrated no statistically significant differences based on the populations we examined (workshop plus website, website only, and neither). In short, the supplemental group appeared to be younger and less likely to be teaching undergraduates than the original group, but their behavior in the classroom seemed the same. We combined these two groups of response in the analysis of the 2009 data.

We sent the 2012 survey to a list created from records from four sources: (i) the AGI list, (ii) an email list of 2-year college geoscience faculty compiled from institutional data sources and augmented by other 2-year college instructors who signed up to be included on such a list, (iii) a list of atmospheric science or meteorology faculty compiled by institutional data sources based on schools that offer degree programs in meteorology or atmospheric science (as listed on the American Meteorological Society website) that were not otherwise included in the AGI list, and (iv) the list of Cutting Edge participants. A listing of oceanography faculty was not available for the survey, although many oceanography programs are included in the AGI directory. We sent the survey to 7813 faculty and received 2466 responses, of which 2157 completed either the introductory or majors course set of questions (a response rate of 32%).

Administration.

In each administration, the survey was presented online with a skip logic structure. After responding to a question asking about teaching responsibilities, respondents were sent either a set of questions asking for a description of an introductory course or a course for majors. Faculty not teaching undergraduate students were sent to the final sections of survey, which included questions related to professional development and contributions to the field.

For the 2004 and 2009 survey, AIP sent email invitations that appeared to be from a Cutting Edge principal investigator. Several reminder emails were sent in this manner as well. For the 2012 survey, participants received an initial invitation signed by Cutting Edge leadership and endorsed by NAGT, indicating that an email from the survey contractor, Professional Data Analysts Inc., would follow. This email was followed by two reminders over a 3-week interval.

Direct observations of teaching practice

In-person classroom observations were made between March 2011 and June 2014 as part of the Cutting Edge Classroom Observation Project using the RTOP (*25*). These observations resulted in 205 sets of quantitative data (RTOP scores), for which there are 174 sets of observer comments made by trained observers during RTOP scoring, and 203 instructor surveys completed by participating faculty. Observers used a modified version of the RTOP rubric described by Budd *et al.* (*55*) that clarifies criteria for scores (0 to 4) for each rubric item, includes comments that help observers understand subtle differences between scores, and places those criteria in a more specific geoscience context.

We selected faculty for observation using a group characteristic sampling frame designed to collect sufficient numbers of observations (greater than 30 per dimension) that demonstrated variation on the dimensions of interest (gender, appointment type, institution types, class size, introductory and majors courses, and Cutting Edge participants versus non–Cutting Edge participants). The group characteristic sampling allows an investigation to document diversity and identify common themes (*56*).

Four cohorts of observers participated in training. The first cohort (*12*) observed classrooms and watched recorded classroom sessions to test and adapt the use of the RTOP rubric of Budd *et al.* (*55*) and to adjust scoring criteria and rubric comments as needed. Following each observation, the group reported scores, discussed scoring, and came to a consensus. After approximately eight observations and adjustments to the scoring criteria, observation scores of each observer were reproducible, and a revised version of the Budd *et al.* rubric was created.

The mean scores of this original cohort were henceforward considered to be the "standard" scores for subsequent training. Three additional cohorts of observers were trained from 2012 to 2014 using a three-stage process of observing videos, scoring videos, and discussing scores with a trained observer. At each stage, if a trainee's scores were within 1 SD of the standard, the trainee would advance to the next stage of training, which included more videos, discussions, and score comparisons. The final stage in the training process required participants to score final calibration videos. Intraclass correlation coefficients (ICCs) were used to report interrater reliability for the 22 trained observers, which yielded an ICC of 0.93. In addition, Cronbach's α was calculated for each calibration video as a measure of reliability for the 25-item RTOP rubric. The rubric was found to be reliable with $\alpha = 0.84$ for the first calibration video ($n = 24$) and $\alpha = 0.81$ ($n = 22$) for the second calibration video. This exceeds the acceptable threshold for interrater reliability, $\alpha > 0.7$ (*57*).

The trained observers assigned quantitative scores for 205 individual class sessions using the revised version of the Budd *et al.* (*55*) RTOP rubric. For these observations, 203 could be grouped by Cutting Edge participation (table S16). The remaining two scores in the data set belonged to workshop participants who self-reported that they did not use the Cutting Edge website. Workshop-only participants were not included in the reported data because of the small sample size of these type of participants.

Qualitative studies

We conducted a retrospective study of a subset of previous interview studies. The original interviews were conducted between the onset of the project and just before the 2009 survey administrations. Each study centered on investigating the role of Cutting Edge in influencing faculty (*58*). However, each set of interviews used a different sampling frame and was conducted by different sets of researchers with different protocols. Although each of the original interview studies had different goals, all of the semistructured protocols included a set of questions designed to elicit a description of the teaching strategies used by a faculty member, plus a set of questions to investigate the type of impacts of the Cutting Edge program on participants. Original interview protocols are provided in the Supplementary Materials.

In collaboration with the external researcher at Rockman et al, we used a causal pathway case sampling approach (*56*) to intentionally select cases from the original interviews for the analysis that would provide insight into causal mechanisms (*59*) of the Cutting Edge program. We first limited the case selection to only those who had participated in a Cutting Edge workshop. We then limited the analysis to only cases where participants had completed a survey. We ran a Structured Query Language (SQL) matching query using a participant-identifying variable across the three retrospective interview study lists and the three survey administration lists. This yielded a sample ($n = 71$) across the three studies (Table 2). From the remaining list, we ran descriptive statistics in SPSS to identify the faculty type and teaching style of each faculty member.

For the independent analysis, two researchers at Rockman et al, in consultation with Cutting Edge researchers, developed a preliminary coding system based on the patterns drawn from the survey analysis. The Rockman et al researchers analyzed the interview cases using these codes related to the faculty type and teaching style. For example, the initial sets of codes for teaching style included subcodes: lecture that encourages class discussion, lecture with demonstrations, and usage of small-group discussions or think-pair-share. These are all changes in teaching directly attributed to Cutting Edge workshops. Other initial codes included features of student-engaged teaching activities and responses related to workshop program feedback. The Features of Strong Teaching Activities codes were based on the criteria established by the Cutting Edge review process (*60*) and incorporate the features of teaching that promote active learning.

Each of the Rockman et al researchers initially coded the same 10 interviews, and the percentage of agreement for codes was 75%. Coders subsequently discussed the conflicting codes and came to a consensus before coding the remaining interviews. The coding guide is presented in the Supplementary Materials. These codes were entered into the software package NVivo, and all of the transcripts were imported into this program.

The researchers analyzed the transcribed interviews using a pattern-matching strategy related to the initial set of codes. They used a coding approach that used a process of constant comparative analysis (*61*). Rather than attempting to identify more subtle themes at the start, the researchers initially used larger "buckets" for these codes. This approach allowed the researchers to conduct multiple iterations. Through this iterative process, additional patterns emerged from the data, and additional codes were used.

Statistical analysis

We conducted exploratory factor analyses to examine differences in geoscience faculty teaching practices from 2004 to 2012. We conducted a cluster analysis to examine whether there were subpopulations of geoscience faculty who exhibited differences in their teaching practices and, if so, what characterized these differences. We conducted regression analyses to investigate whether Cutting Edge workshop participants and Cutting Edge website users were more likely to report interactive class time and active learning teaching strategies after controlling for other differences in the participants.

The questions in the survey evolved from one administration to another in response to the utility of data from the previous surveys and the evolving focus of both Cutting Edge and interests in geoscience and higher-education research. Thus, the survey was not identical from one administration to another and varied in length, content, and order. However, the root question used in the factor analysis to determine teaching strategy remained unchanged across administrations of the survey.

Factor analysis: Teaching strategy.

Seven survey items from the 2004 and 2009 surveys were reduced to construct the teaching strategy variables (traditional lecture, active lecture, and active learning) by applying results from a factor analysis. Survey responders who taught introductory and major courses were asked, "In the lecture portion, please indicate how frequently you used the following teaching strategies." A core set of responses was available across surveys:

(1) Traditional lecture

(2) Lecture with demonstration

(3) Lecture in which questions posed by the instructor are answered by individual students (for example, a professor calls on individual students)

(4) Lecture in which questions posed by the instructor are answered simultaneously by the entire class (for example, students vote using cards or electronic response systems)

(5) Small-group discussion or think-pair-share

(6) Whole-class discussion

(7) In-class exercises

(8) Classroom debates or role-playing (2004 survey only)

(9) Fieldwork (2004 survey only)

Table 2. Summary of qualitative interview cases sampled for the retrospective study.

Years	Project	Total number of interviews	Number of interviews included in the retrospective study
2005	Phone interviews of 2002–2004 workshop participants about workshop impact	54	34
2005–2009	Phone interviews of participants and nonparticipants about website	29	10
2007	Face-to-face interviews of both Cutting Edge participants and nonparticipants	37	27
	Total	120	71

For each item, respondents were given a five-point rating scale to indicate frequency: never, once, several times, weekly, and every class. The response choices were held nearly constant:

(1) Never
(2) Once (2004–2009)/once or twice (2012)
(3) Several times
(4) Weekly
(5) Every class (2004–2009)/nearly every class (2012)

For the factor analyses, ratings for the seven strategies included in all three surveys were recoded into frequent (weekly and every class) and infrequent (never, once, and several times). The seven binary items were included in a factor analysis that used principal axis factoring with varimax rotation. The number of rotations ultimately extracted was based on an iterative process guided by several principles: (i) those with eigenvalues greater than 1, (ii) scree plots of eigenvalues by factor number, and (iii) the interpretability of factors.

There were missing data in 586 responses in 2004, 865 responses in 2009, and 725 responses in 2012. The surveys had missing data for several reasons: (i) respondents either were not presented with or did not complete the survey sections asking about an introductory or major course, (ii) respondents skipped the items in the question, or (iii) respondents did not do any of the items "frequently." We minimized the effects of missing values by using pairwise deletion where all valid data are used in the analyses.

The factor analysis performed on 2004–2009 data suggests three factors:

(1) Factor 1 (active lecture): lecture with demonstration, lecture in which questions posed by instructor are answered by individual students, and lecture in which questions posed by instructor are answered simultaneously by the entire class
(2) Factor 2 (active learning): small-group discussion, whole-class discussions, and in-class exercises
(3) Factor 3 (traditional lecture): traditional lecture

A teaching strategy variable was constructed using the three factors from the 2004 and 2009 data. The same variable was created in the 2012 data by applying the factors from the early surveys.

Cluster analysis: Faculty type.

The faculty type variable (education-focused faculty, geoscience research–focused faculty, and teaching faculty) was constructed from a cluster analysis using survey items asking about participation and engagement in the geoscience community. We ran the cluster analysis on the 2004 and 2009 data, and it was verified in the 2012 data. Demographic questions used in the cluster analysis had only minor variations across administrations. The seven questions used in the cluster analysis included the following:

(1) At how many meetings have you presented your scientific research within the past 2 years?
(2) How many articles about your research have you published in the past 2 years?
(3) How many articles have you published about educational topics within the past 2 years? (In 2004, the question was not limited to the past 2 years.)

(The previous three questions were open-ended in 2004 and had the following response options in the 2009 and 2012 administrations: none, 1, 2, 3, 4, 5, 6, and 7 or more.)

(1) How often did you talk or correspond with your colleagues about course content over the past 2 years? (In 2012, course content was underlined.) The possible responses were as follows: (i) never, (ii) once or twice per term, (iii) several times per term, (iv) weekly, and (v) nearly every day.
(2) Approximately how many talks on teaching methods, other topics related to science education, or geoscience education have you attended in the past 2 years at professional meetings, on campus, or at other venues? (This question was open-ended in 2004 and had the following response options in 2009 and 2012: none, 1 or 2, 3 or 4, 5 or 6, 7 or 8, 9 or 10, and 11 or more.)
(3) How many workshops related to improving your teaching did you attend in the past 2 years? (This question was open-ended in 2004 and had the following response options in 2009 and 2012: none, 1, 2, 3, 4, and 5 or more.)
(4) Have you presented research on teaching methods or student learning at meetings within the past 2 years? Response choices were yes and no.

We ran a k-means clustering algorithm on the 2004 and 2009 data in SPSS Quick Cluster (version 18) to determine whether respondents could be grouped on the basis of responses to these items. The analysis examined two-, three-, and four-cluster solutions, and the three-cluster solution was found to be the most interpretable and was therefore retained. The three-cluster solution was verified in the 2012 survey data and fit the data well. There were missing data for 575 cases in 2004, 388 cases in 2009, and 319 cases in 2012. These cases were not grouped into a faculty type.

Logistic regression analyses.

Logistic regression was used for the dichotomous variable interactive class time (more than 20% versus 20% or less), and three separate logistic regression models were used to assess associations with the tri-categorical variable teaching strategy: active learning versus traditional lecture, active lecture versus traditional lecture, and active learning versus active lecture. Analyses were performed with listwise deletion of missing data. All models contained Cutting Edge participation as the main predictor of interest. Models also included a number of background covariates (that is, control variables), which allows analyses to isolate the effect of Cutting Edge participation after controlling for the effects of other covariates thought to be related to the outcome variables.

Control variables consisted of background characteristics, course characteristics, and faculty type. A forward step variable selection was used in all regression models: beginning with the intercept-only model and adding one covariate at a time, at each step choosing the control covariate whose inclusion improves the model the most. This procedure was repeated until there were no remaining covariates whose inclusion significantly improved the model. Finally, Cutting Edge participation was added to the model containing all significant covariates.

For logistic regression models, the odds ratio (OR) is typically reported along with the coefficient estimate of each covariate. However, a more interpretable coefficient is the relative risk (RR), which is produced using the OR and the reference category probability (RCP; obtained from cross-tabulation of the covariates and the response) (Eqs. 1 to 3).

$$\text{Comparison category probability (CCP)} = \frac{(\text{RCP*OR})}{(1-\text{RCP}) + (\text{RCP*OR})} \quad (1)$$

$$\text{Predicted probability differential (PPD)} = \text{CCP} - \text{RCP} \quad (2)$$

$$\text{RR} = \frac{\text{CCP}}{\text{RCP}} \quad (3)$$

We report the significance, OR, RR, and PPD in logistic regression summaries (tables S14 and S15).

Spearman's rank-order correlations.

Spearman's rank-order correlations in SPSS were used to compare the relationship between teaching style (active learning, active lecture, and traditional lecture) and responses to items related to features of engaged teaching. This was run on 2012 survey responses.

Chi-squares.

Chi-square tests in SPSS were used to compare for each engaged pedagogy item the number of faculty who reported how prominently students exhibited a particular behavior related to engaged pedagogy. Each item was compared by faculty group. Comparisons between faculty type and the 11 engaged pedagogies were assessed using chi-square tests with a Bonferroni-adjusted P value of 0.005.

Intraclass correlation coefficients.

A "two-way mixed" model was run in SPSS and reported for the single measure on the basis of the trained observers' total scores on the two videos. This model was run on 22 of the 24 observers who completed the training, because two observers (each contributed five scores to the total data set) did not score both calibration videos.

Cronbach's α.

Cronbach's α was calculated to determine the rubric internal consistency (interitem correlations) using scores from two calibration videos for trained RTOP observers whose observations were part of the total data set. The calibration videos were scored using the RTOP scoring rubric at the end of the observers' training. Using scores from the 25-item rubric and an Excel formula, Cronbach's α was calculated for each video (*62*). Scores from all 24 observers were available for the first video, but 2 observers were missing scores for the second video. Therefore, Cronbach's α was run for the 22 available scores.

Analysis of variance

A one-way ANOVA in SPSS was used to compare the total RTOP scores. Comparisons were made between observed faculty who had no Cutting Edge participation, observed faculty who had only used the Cutting Edge website, and observed faculty who had used the website and attended a Cutting Edge workshop. Bonferroni post hoc tests were conducted to confirm which Cutting Edge groups were different from each other on mean RTOP score.

SQL queries

SQL queries were used in Microsoft Access to identify repeat participation of Cutting Edge events for the Cutting Edge population as a whole, Cutting Edge survey respondents, and interview respondents. Each query compared the identifiers of the particular population to the database of workshop seats by participant ID. A frequency of unique participant identifiers where count was greater than 1 was computed for each population. To identify only repeat participation related to the particular method (before interview or survey response), the database of workshop seats for the query was limited to the years of interest for the particular population.

SUPPLEMENTARY MATERIALS

Supplementary material for this article is available at http://advances.sciencemag.org/cgi/content/full/3/2/e1600193/DC1

fig. S1. Venn diagram showing survey year and number of participant responses for each survey.
fig. S2. Frequency by survey year of percentage spent on interactive class time.
table S1. Frequency by survey year of percentage spent on interactive class time.
table S2. Percentage of survey respondents reporting that more than 20% of class time is interactive by year.
table S3. Percentage of faculty group reporting that more than 20% of class time is interactive by year.
table S4. Percentage of responses reporting that more than 20% of class time is interactive for each teaching strategy by year.
table S5. Percentage of responses reporting that more than 20% of class time is interactive for each course level by year.
table S6. Spearman correlation between teaching strategy and other measures of engaged teaching.
table S7. Faculty types by survey year.
table S8. Cluster variables by faculty type: mean value of cluster variable (95% confidence interval).
table S9. Reporting of teaching strategy and interactive class time by faculty type.
table S10. Reporting of other measures of engaged teaching by faculty type.
table S11. Background and teaching characteristics by faculty types from the 2012 survey.
table S12. Relationship between faculty type and institution type in 2012 survey responses determined using cross-tabulation analysis in SPSS 22.
table S13. Engagement in learning about pedagogy characteristics by faculty type from the 2009 and 2012 surveys.
table S14. Logistic regression model for predicting more than 20% of class time on student activities, questions, and discussion.
table S15. Logistic regression results for predicting active learning (versus active lecture).
table S16. RTOP score ranges, quartiles, mean, and medians as reported for 203 observations.
table S17. Cutting Edge participation by faculty type.
table S18. Growth in Cutting Edge participants in U.S. geoscience faculty population and in survey sample.
Interview protocols
Coding book for qualitative analysis

REFERENCES AND NOTES

1. President's Council of Advisors on Science and Technology, "Engage to excel: Producing one million additional college graduates with degrees in science, technology, engineering, and mathematics" (Executive Office of the President, 2012); http://files.eric.ed.gov/fulltext/ED541511.pdf.
2. National Science Foundation (NSF), "Shaping the future: New expectations for undergraduate education in science, mathematics, engineering, and technology" (nsf96139, NSF, 1996); http://nsf.gov/publications/pub_summ.jsp?ods_key=nsf96139.
3. S. Freeman, S. L. Eddy, M. McDonough, M. K. Smith, N. Okoroafor, H. Jordt, M. P. Wenderoth, Active learning increases student performance in science, engineering, and mathematics. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 8410–8415 (2014).
4. S. Singer, K. A. Smith, Discipline-based education research: Understanding and improving learning in undergraduate science and engineering. *J. Eng. Educ.* **102**, 468–471 (2013).
5. S. L. Eddy, K. A. Hogan, Getting under the hood: How and for whom does increasing course structure work? *CBE Life Sci. Educ.* **13**, 453–468 (2014).
6. D. C. Haak, J. HilleRisLambers, E. Pitre, S. Freeman, Increased structure and active learning reduce the achievement gap in introductory biology. *Science* **332**, 1213–1216 (2011).
7. D. Ebert-May, T. L. Derting, J. Hodder, J. L. Momsen, T. M. Long, S. E. Jardeleza, What we say is not what we do: Effective evaluation of faculty professional development programs. *Bioscience* **61**, 550–558 (2011).
8. National Research Council, *Discipline-Based Education Research: Understanding and Improving Learning in Undergraduate Science and Engineering* (The National Academies Press, 2012).
9. C. Henderson, A. Beach, N. Finkelstein, Facilitating change in undergraduate STEM instructional practices: An analytic review of the literature. *J. Res. Sci. Teach.* **48**, 952–984 (2011).
10. K. Eagan, "Becoming more student centered? An examination of teaching practices across stem and non-stem disciplines between 2004 and 2014" (2016): https://sloan.org/storage/app/media/files/STEM_Higher_Ed/STEM_Faculty_Teaching_Practices.pdf.
11. C. A. Manduca, D. W. Mogk, B. Tewksbury, R. H. Macdonald, S. P. Fox, E. R. Iverson, K. Kirk, J. McDaris, C. Ormand, M. Bruckner, On the Cutting Edge: Teaching help for geoscience faculty. *Science* **327**, 1095–1096 (2010).
12. Science Education Resource Center (SERC), "On the Cutting Edge" (SERC, 2016); http://serc.carleton.edu/NAGTWorkshops/index.html.
13. C. Wilson, "Status of the Geoscience Workforce 2014" (American Geosciences Institute, 2014).
14. SERC, "Overall Philosophy of Cutting Edge Workshop Design" (SERC, 2016); http://serc.carleton.edu/NAGTWorkshops/about/design.html.
15. C. Amundsen, M. Wilson, Are we asking the right questions? A conceptual review of the educational development literature in higher education. *Rev. Educ. Res.* **82**, 90–126 (2012).
16. C. D'Avanzo, Post–vision and change: Do we know how to change? *CBE Life Sci. Educ.* **12**, 373–382 (2013).
17. D. Ebert-May, T. L. Derting, T. P. Henkel, J. M. Maher, J. L. Momsen, B. Arnold, H. A. Passmore, Breaking the cycle: Future faculty begin teaching with learner-centered strategies after professional development. *CBE Life Sci. Educ.* **14**, ar22 (2015).
18. C. A. Manduca, D. W. Mogk, N. Stillings, "Bringing research on learning to the geosciences" (SERC, 2004); http://serc.carleton.edu/files/research_on_learning/ROL0304_2004.pdf.
19. D. C. Edelson, Learning-for-use: A framework for the design of technology-supported inquiry activities. *J. Res. Sci. Teach.* **38**, 355–385 (2001).
20. A. Kezar, S. Gehrke, "Communities of transformation and their work in achieving scale of STEM reform" (Pullias Center for Higher Education, 2015); www.uscrossier.org/pullias/wp-content/uploads/2016/01/communities-of-trans.pdf.
21. R. H. Macdonald, C. A. Manduca, D. W. Mogk, B. J. Tewksbury, Teaching methods in undergraduate geoscience courses: Results of the 2004 On the Cutting Edge survey of U.S. faculty. *J. Geosci. Educ.* **53**, 237–252 (2005).
22. J. A. McLaughlin, E. Iverson, R. Kirkendall, C. Manduca, M. Bruckner, "On the Cutting Edge 2010 evaluation report" (SERC, 2010); https://serc.carleton.edu/files/NAGTWorkshops/2009_cutting_edge_evaluation_1265409435.pdf.
23. Rockman et al, "On The Cutting Edge Project: Year 3 independent evaluation report" (SERC, 2013); http://serc.carleton.edu/files/NAGTWorkshops/about/2013_external_evaluation_repor.pdf.
24. D. Sawada, M. Piburn, E. Judson, J. Turley, K. Falconer, R. Benford, I. Bloom, Measuring reform practices in science and mathematics classrooms: The Reformed Teaching Observation Protocol. *Sch. Sci. Math.* **102**, 245–253 (2002).
25. SERC, "Classroom observation project: Understanding and improving our teaching" (SERC, 2016); http://serc.carleton.edu/NAGTWorkshops/certop/index.html.
26. E. F. Holton, Holton's evaluation model: New evidence and construct elaborations. *Adv. Develop. Hum. Resour.* **7**, 37–54 (2005).
27. E. W. L. Cheng, I. Hampson, Transfer of training: A review and new insights. *Int. J. Manag. Rev.* **10**, 327–341 (2008).
28. J. Gitonga, "Transfer of learning in continuing medical education (CME): A conceptual model", paper presented at the International Research Conference in the Americas of the Academy of Human Resource Development, Indianapolis, IN, 28 February to 4 March 2007.
29. S. M. Wilson, Professional development for science teachers. *Science* **340**, 310–313 (2013).
30. L. M. Desimone, Improving impact studies of teachers' professional development: Toward better conceptualizations and measures. *Educ. Res.* **38**, 181–199 (2009).
31. M. S. Garet, A. C. Porter, L. Desimone, B. F. Birman, K. S. Yoon, What makes professional development effective? Results from a national sample of teachers. *Am. Educ. Res. J.* **38**, 915–945 (2001).
32. M. Dancy, C. Henderson, Pedagogical practices and instructional change of physics faculty. *Am. J. Phys.* **78**, 1056–1063 (2010).
33. J. E. Froyd, M. Borrego, S. Cutler, C. R. Henderson, M. J. Prince, Estimates of use of research-based instructional strategies in core electrical or computer engineering courses. *IEEE Trans. Edu.* **56**, 393–399 (2013).
34. E. F. Redish, *Teaching Physics with the Physics Suite* (John Wiley and Sons, 2003).
35. E. Mazur, *Peer Instruction: A User's Manual* (Prentice Hall, 1997).
36. T. J. Lund, M. Pilarz, J. B. Velasco, D. Chakraverty, K. Rosploch, M. Undersander, M. Stains, 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 Sci. Educ.* **14**, ar18 (2015).
37. American Association for the Advancement of Science (AAAS), "Describing and measuring undergraduate STEM teaching practice: A report from a national meeting on the measurement of undergraduate science, technology, engineering and mathematics (STEM) teaching" (AAAS, 2013); http://ccliconference.org/files/2013/11/Measuring-STEM-Teaching-Practices.pdf.
38. K. Eagan, E. B. Stolzenberg, J. B. Lozano, M. C. Aragon, M. R. Suchard, S. Hurtado, *Undergraduate Teaching Faculty: The 2013-2014 HERI Faculty Survey* (University of California, 2014).
39. W. Condon, E. R. Iverson, C. A. Manduca, C. Rutz, G. Willett, *Faculty Development and Student Learning: Assessing the Connections* (Scholarship of Teaching and Learning, Indiana Univ. Press, 2015).
40. D. L. Kirkpatrick, Evaluation of training, in *Training and Development Handbook*, R. L. Craig, L. R. Bittel, Eds. (McGraw-Hill, 1967), pp. 87–112.
41. S. Naquin, E. Holton, Motivation to improve work through learning in human resource development. *Hum. Res. Dev. Int.* **6**, 355–370 (2003).
42. M. K. Smith, E. L. Vinson, J. A. Smith, J. D. Lewin, M. R. Stetzer, A campus-wide study of STEM courses: New perspectives on teaching practices and perceptions. *CBE Life Sci. Educ.* **13**, 624–635 (2014).
43. J. Fairweather, "Linking evidence and promising practices in science, technology, engineering, and mathematics (STEM) undergraduate education" (The National Academies National Research Council Board of Science Education, 2009).
44. T. T. Baldwin, J. K. Ford, Transfer of training: A review and directions for future research. *Person. Psychol.* **41**, 63–105 (1988).
45. M. J. Prince, M. Borego, S. Cutler, C. Henderson, J. E. Froyd, Use of research-based instructional strategies in core chemical engineering courses. *Chem. Eng. Educ.* **47**, 27–37 (2013).
46. C. Henderson, M. Dancy, Barriers to the use of research-based instructional strategies: The influence of both individual and situational characteristics. *Phys. Rev. S.T. Phys. Educ. Res.* **3**, 020102-1–020102-14 (2007).
47. S. E. Brownell, K. D. Tanner, Barriers to faculty pedagogical change: Lack of training, time, incentives, and… tensions with professional identity? *CBE Life Sci. Educ.* **11**, 339–346 (2012).
48. E. Wenger-Trayner, B. Wenger-Trayner, "Communities of practice: A brief introduction" (Wenger-Trayner, 2016); http://wenger-trayner.com/introduction-to-communities-of-practice.
49. E. F. Holton, R. A. Bates, W. E. A. Ruona, Development of a generalized learning transfer system inventory. *Hum. Resource Dev. Q.* **11**, 333–360 (2000).
50. S. Gehrke, A. Kezar, STEM reform outcomes through communities of transformation. *Change* **48**, 30–38 (2016).
51. K. L. Kenyon, M. E. Onorato, A. J. Gottesman, J. Hoque, S. G. Hoskins, Testing CREATE at community colleges: An examination of faculty perspectives and diverse student gains. *CBE Life Sci. Educ.* **15**, 1–19 (2016).
52. M. A. Pelch, D. A. McConnell, Challenging instructors to change: A mixed methods investigation on the effects of material development on the pedagogical beliefs of geoscience instructors. *Int. J. STEM Educ.* **3**, 1–18 (2016).
53. C. Rutz, J. Lauer-Glebov, Assessment and innovation: One darn thing leads to another. *Assessing Writing* **10**, 80–99 (2005).
54. D. Russ-Eft, H. Preskill, Evaluating learning, performance, and change initiatives, in Evaluation in Organizations: A Systematic Approach to Enhancing Learning, Performance, and Change (Basic Books, ed. 2, 2009), chap. 3.

55. D. A. Budd, K. Kraft, D. A. McConnell, T. Vislova, Characterizing teaching in introductory geology courses: Measuring classroom practices. *J. Geosci. Educ.* **61**, 461–475 (2013).
56. M. Q. Patton, *Qualitative Research & Evaluation Methods: Integrating Theory and Practice* (SAGE, 2015).
57. C. Henderson, R. Cole, J. Froyd, D. Gilbuena, R. K. Friedrichsen, C. Stanford, *Designing Educational Innovations for Sustained Adoption: A How-to Guide for Education Developers Who Want to Increase the Impact of Their Work* (Increase the Impact, 2015).
58. SERC, "On the Cutting Edge evaluation" (SERC, 2015); http://serc.carleton.edu/NAGTWorkshops/about/evaluation.html.
59. J. Gerring, *Case Study Research: Principles and Practices* (Cambridge Univ. Press, 2007).
60. SERC, "Activity design: Questions to consider when designing or reviewing an activity" (SERC, 2012); http://serc.carleton.edu/NAGTWorkshops/servicelearning/workshop10/activity_design/index.html.
61. B. Glaser, A. Strauss, *Discovery of Grounded Theory: Strategies for Qualitative Research* (Aldine, 1967).
62. D. Siegle, "Educational research basics: Excel spreadsheet to calculate instrument reliability estimates," http://researchbasics.education.uconn.edu/excel-spreadsheet-to-calculate-instrument-reliability-estimates.

Acknowledgments: We thank the individuals involved in the study, including R. Czujko and the Statistical Research Center at the AIP, who conducted the 2004 and 2009 surveys and the initial data analysis; M. Bruckner, R. McFadden, and T. Kjerland who assisted in the data management at Science Education Resource Center; Professional Data Analysts Inc. staff E. Graalum and M. Thao who assisted with database management and contributed to data analysis and interpretation; and L. Greenseid who facilitated survey design and implementation and contributed to interpretation of data analysis, together with B. Lien. RTOP data were collected by the On the Cutting Edge Classroom Observation Project under the leadership of J. Bartley, D. Farthing, R. Teasdale, and K. Viskupic and supported by M. Bruckner. The success of the Cutting Edge project is the result of leadership of project principal investigators including R. Beane, K. Wiese, and M. Wysession in addition to the authors, as well as to the engagement of community members as workshop leaders, website authors, workshop participants, and website users. The anonymous reviewers contributed substantially to the improvement of the paper. **Funding:** The Cutting Edge program and its evaluation were supported by the NSF Division of Undergraduate Education grants DUE-0127310, DUE-0127141, DUE-0127257, DUE-0127018, DUE-0618482, DUE-0618725, DUE-0618533, DUE-1022680, DUE-1022776, DUE-1022844, and DUE-1022910. Any opinions, findings, conclusions, or recommendations expressed in this article are those of the authors and do not necessarily reflect the views of the NSF. **Author contributions:** C.A.M. was responsible for the overall survey project; R.H.M., D.W.M., and B.J.T. contributed to the intellectual design of the survey; M.L. was responsible for survey administration and statistical analysis; E.R.I. was responsible for the qualitative studies; and D.A.M. was responsible for the RTOP study. **Competing interests:** C.A.M. is Executive Director of the NAGT, which now manages the On the Cutting Edge program. All other authors remain invested in managing the ongoing program but declare that they have no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. Additional data related to this paper may be requested from the authors.

Submitted 1 February 2016
Accepted 2 January 2017
Published 15 February 2017
10.1126/sciadv.1600193

Citation: C. A. Manduca, E. R. Iverson, M. Luxenberg, R. H. Macdonald, D. A. McConnell, D. W. Mogk, B. J. Tewksbury, Improving undergraduate STEM education: The efficacy of discipline-based professional development. *Sci. Adv.* **3**, e1600193 (2017).