The ICAP Active Learning Framework Predicts the Learning Gains Observed in Intensely Active Classroom Experiences

Benjamin L. Wiggins
University of Washington

Sarah L. Eddy
Florida International University

Daniel Z. Grunspan
Arizona State University

Alison J. Crowe
University of Washington

STEM classrooms (science, technology, engineering, and mathematics) in postsecondary education are rapidly improved by the proper use of active learning techniques. These techniques occupy a descriptive spectrum that transcends passive teaching toward active, constructive, and, finally, interactive methods. While aspects of this framework have been examined, no large-scale or actual classroom-based data exist to inform postsecondary education STEM instructors about possible learning gains. We describe the results of a quasi-experimental study to test the apex of the ICAP framework (interactive, constructive, active, and passive) in this ecological classroom environment. Students in interactive classrooms demonstrate significantly improved learning outcomes relative to students in constructive classrooms. This improvement in learning is relatively subtle; similar experimental designs without repeated measures would be unlikely to have the power to observe this significance. We discuss the importance of seemingly small learning gains that might propagate throughout a course or departmental curriculum, as well as improvements with the necessity for faculty to develop and implement similar activities.

Keywords: *active learning, ICAP, interactive, constructive*

IMPROVEMENT efforts in postsecondary STEM education (science, technology, engineering, and mathematics) at the classroom level have largely focused on the opportunities provided by instructors for student engagement with course content (American Association for the Advancement of Science, 2009; Handelsman, Miller, & Pfund, 2007). While postsecondary classrooms traditionally employ passive delivery techniques, such as lecturing, a range of cognitive engagement activities is available to instructors. These diverse active learning approaches—including small group work, cooperative interactions, scaffolded practice, and developing metacognition—are implemented with wide variation. Although a meta-analysis of over 200 studies demonstrated the benefit of active learning over traditional lecturing in STEM environments (Freeman et al., 2014), few studies have compared the relative effectiveness of different types of active learning approaches. The traditional nature of lecturing in STEM classrooms is accentuated in large courses, especially those found at universities with large introductory course enrollments (President's Council of Advisors on Science and Technology, 2012). As postsecondary science classrooms move toward greater use of evidence-based active learning practices, experimental comparisons of these different approaches could provide guiding design principles for best practices in the active STEM classroom. This goes beyond accepting best practices in education to ask the more detailed question: In what ways should college STEM courses be redesigned for active learning?

Chi and Wylie (Chi, 2009; Chi & Wylie, 2014) developed a framework for categorizing different learning tasks in terms of what students are doing when they complete a task. While other frameworks exist to conceptualize active learning, the ICAP framework describes observable elements in classroom practice. This makes it particularly well suited for design, analysis, and research by instructors of large student populations. The ICAP framework outlines four levels of activity, ranging from most engaged to least engaged: interactive, constructive, active, and passive. As proposed by Chi and Wylie (2014), in a passive activity, students receive information but do not otherwise overtly engage with the learning material. Activities are defined as active if they require focused motor movement, such as underlining or

copying selected text passages. Constructive activities go beyond active, by requiring students to synthesize their own ideas and generate a novel output, such as a concept map. Finally, an interactive activity is one in which students engage in a substantive exchange of ideas leading to a new level of understanding. Starting with passive, each subsequent category involves an increased level of activity, with the highest amount of learning predicted to come from the more engaged activities (Menekse, Stump, Krause, & Chi, 2013). Based on this observable framework, the highest amount of learning is predicted to come from interactive activities in which students cocreate a tangible product that incorporates each student's ideas. Specifically, ICAP predicts that interactive activities will support more learning than constructive activities. These additional learning gains are hypothesized to come from increased levels of student engagement. The ICAP framework is a particularly well-suited framework from which to examine detailed active learning techniques because of the observable nature of the student activities at each level, but this does not diminish the quality or usefulness of other frameworks on teaching and learning (e.g., Chen et al., 2014; Oleson & Hora, 2014).

Support for the ICAP hypothesis comes indirectly from analysis of completed studies that compared the effectiveness of different modes of learning that could, in retrospect, be classified along the ICAP spectrum (Chi & Wylie, 2014; Linton, Farmer, & Peterson, 2014; Linton, Pangle, Wyatt, Powell, & Sherwood, 2014; Menekse et al., 2013). These retrospective analyses provide particularly strong support for the benefit of constructive activities, such as building concept maps and self-explaining, over either active or passive behaviors (Coleman, Brown, & Rivkin, 1997; Gobert & Clement, 1999). Fewer studies have compared the effectiveness of activities that could be classified as interactive, such as collaborative note taking or cooperatively building a concept map, and those that are primarily constructive. Studies testing constructive against interactive activities either were conducted in non–college classroom settings (Chi, Roy, & Hausmann, 2008; Kramarski & Dudai, 2009; Rummel, Spada, & Hauser, 2009) or did not detect differences between interactive and constructive activities (Kam et al., 2005; Menekse et al., 2013). Consistent with these findings, a recent study in an undergraduate biology classroom, which tested the benefit of coupling peer discussion with an independent writing component, did not find greater gains in student writing scores when compared with a writing-alone group, despite increased time on task (Linton, Farmer, & Peterson, 2014). Although this study was not explicitly designed to compare interactive and constructive activities, given the ICAP hypothesis, we would predict that adding the interactive peer discussion component would have increased student learning over the constructive activity of written self-explanations. However, a course-long comparison of the benefits of peer discussion in an undergraduate biology classroom did find that students who completed in-class activities in groups rather than individually performed better on exam questions requiring higher-order cognitive skills (Linton, Pangle, et al., 2014).

In addition to the retrospective studies described so far, the ICAP hypothesis has been directly tested by comparing activities explicitly designed to fall into one of the four ICAP categories (Menekse et al., 2013). In this direct test of the ICAP hypothesis performed in a college engineering classroom, constructive activities were found to be more beneficial for student learning than either active or passive activities. However, no added benefit was detected when students engaged in an interactive activity versus a constructive activity. This may have been due either to the small sample size ($n = 42$) or fidelity of implementation in a classroom setting, as a companion study under controlled laboratory conditions did show increased student learning with each subsequent level of activity, as predicted by the ICAP hypothesis. Thus, there is currently only limited support for the prediction that interactive activities promote higher learning gains than do constructive activities in the more ecologically relevant setting of a college classroom.

Common undergraduate STEM courses need improvement to meet national goals (President's Council of Advisors on Science and Technology, 2012), but curricular redesign requires considerable time and effort. It is important to assess the potential benefits of adopting different teaching strategies so that they can be weighed against the costs of design, especially in the most intensely active classroom designs. This is critical for STEM instructors, who often do not have training in education best practices and are most attuned to quantitative data to uptake new teaching strategies. Interactive activities require meaningful exchange between at least two different parties and thus are more challenging for even expert practitioners to implement in a large classroom setting. To outweigh this potential "cost," it is important to determine whether the theoretical gains predicted by the ICAP framework, especially at the higher levels of constructive and interactive, are reflected in the complex cultural environments of the college STEM classroom. In addition, no studies have explored whether the predicted benefits stemming from increased interaction are the same for different demographic populations. Given the diverse cultural funds of knowledge that exist in student populations, we expect differences in the engagement of these students with complicated social practices, such as learning in STEM classrooms (Azevedo, 2013). We attempt a more responsive understanding of student cultural practices, not with the intent to describe them in any appropriate detail, but rather to view this diversity as a rich source of potential solutions for historically intransigent gaps in outcomes for students from underrepresented groups in STEM fields (Gay, 2010). The observable nature of explicit

elements of the ICAP framework makes this investigation possible for real, large classroom environments.

Postsecondary STEM classrooms are a particularly good environment in which to study education outcomes—by the relative adherence to widespread traditional practices, the current increased focus on improving STEM graduation outcomes, the discrete answers on high-cognitive quizzes, and the large numbers of students enrolled in introductory STEM courses. While this quasi-experiment could have been performed in a suitable classroom outside of STEM, we used a STEM classroom because it provided a setting with these logistic needs while addressing growing national concerns over improved college STEM courses. In this study, we ask the following questions:

Research Question 1: Are interactive activities more effective than constructive activities at promoting student learning in a large undergraduate STEM classroom?

If so, then these data may help to determine whether additional learning benefits in interactive redesign are worth the resource costs for instructors.

Research Question 2: Do different demographic populations respond similarly to these interventions?

If so, then these data may help instructors take advantage of differences in responding to long-standing and problematic achievement gaps.

To answer these questions, we have designed paired sets of in-class activities that differ only in their mode of student engagement: constructive or interactive. The activities were implemented in two sections of a large college STEM classroom. In this course, each student engaged serially on different days in constructive and interactive activities on different topics, which enabled us to analyze student performance using a repeated measures approach. Demographic and student performance data were used to assess potential interactions between activity-type and student characteristics. Our results indicate that, in a pairwise comparison, interactive activities promote increased learning over constructive activities in an ecological setting and that all students benefit equally from this enhanced level of activity. We hope that this result will inform practitioners and investigators within and possibly beyond STEM higher education.

Methods

Setting

This research took place in an introductory biology course at a large research-focused (R1) public university. The class was taught in back-to-back sections of ~350 students, each for 50 minutes four times per week by the same instructor. Associated 2.5-hr labs were held each week of the

TABLE 1
Demographic Data for Course Sections

	Class section	
	1	2
Female:male, *n*	243:142	230:144
African American	15	7
American Indian	2	2
Asian	140	133
White	172	168
Hawaiian / Pacific Islander	5	11
Hispanic	22	19
First generation:non–first generation, *n*	185:200	168:206
Grade point average, *M* (*SD*)	3.23 (0.41)	3.24 (0.40)

Note. Demographic data provided by the university registrar.

10-week course. Students were evaluated through several assignments, with the majority of variation in grades coming from four noncumulative exams. Demographic and academic student information was collected from the university registrar.

As described by registrar statistics, this classroom comprised 61% female students, 6% community college transfer students, 56% non-Caucasian students, 46% first-generation students, and 18% underrepresented minorities. Students in the class were predominantly of sophomore and junior standing and declared a range of majors, typically in the natural sciences. The average SAT scores of this population were 549 for math and 515 for verbal. All classroom and student consent were protected and managed under Institutional Review Board protocols. No students opted out of the study population for use of their outcome and demographic data. Chi-square comparison of demographic data between the two sections did not reveal significant differences between the groups (Table 1).

The split-section environment allows for relatively controlled quasi-experimentation with teaching techniques or interventions. Within this framework, broad-spectrum variables are relatively well controlled, including topic, instructor, classroom environment, time on task, instruction wording, and motivating factors (Figure 1). Students self-selected into one of two large sections of the course; so, random assignment into treatment groups was not possible, and these methods technically constitute a quasi-experimental rather than a randomized controlled experiment. It is within this controlled system that instruction was manipulated between interactive and constructive student activities.

Four topics were chosen for use in constructive-versus-interactive quasi-experiments. These four commonly taught topics in cell biology were chosen for their ubiquity among introductory science coursework and applicability of designed activities to similar courses at other institutions.
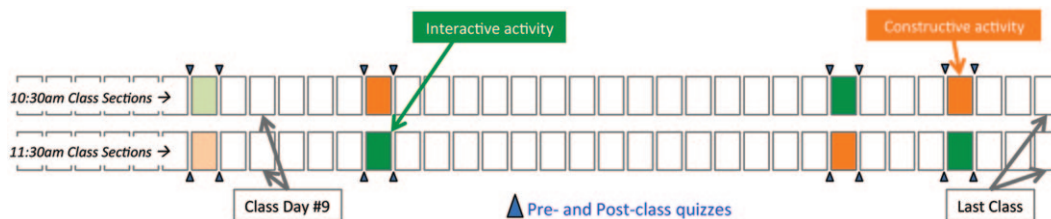
FIGURE 1. *Timing of quasi-experimental class sections and pre- and postquiz administration. This diagram shows the distribution of quasi-experiments throughout the course. For each of the four quasi-experimental days, one section of the class used an interactive classroom activity, while the other used a constructive activity. The choice of section was rotated to allow for a repeated-measures analysis. Prequizzes were administered as part of a daily reading quiz for each quasi-experimental day to students in both sections. Postquizzes were administered on the following reading quiz. Quizzes were taken between the previous class afternoon and the morning of the relevant class so that all learning-based outcome data were collected in the near term, within 24 hours of the quasi-experiment. Note that the first of four comparisons was removed in the final statistical analysis (explanation in Methods section).*

The four topics were protein translation, eukaryotic gene regulation, cell cycle and cancer, and polymerase chain reaction. One of the four quasi-experimental topic days was removed from data analysis as explained in Statistical Analysis of Student Performance on Posttests section.

Activity Design

Design of activities was conducted with principles described in the ICAP framework (Chi & Wylie, 2014). No real-world activity is a perfect fit into a single ICAP category, and the activities that we designed here are no exception; our goal was to create quasi-experimental treatments that were predominantly within a single category and for which the differences were otherwise minimal.

In designing constructive versions of activities, we focused on providing opportunities for students to generate outputs of their own understanding that went beyond the answers provided. Consistent with the criteria outlined by Chi and Wylie (2014), students were asked to integrate concepts across texts, to compare and contrast mechanisms, and to predict outcomes for new situations using conceptual understandings from relevant but distinct examples. These student actions were central to completing in-class tasks and were design elements that made constructive versions of activities more likely to support student engagement beyond the lower, "active" level of the ICAP framework. Throughout our constructive activities, students worked in small groups, but interaction with group members was not required to complete the activity.

In designing interactive versions of activities, interactions among students to cogenerate new understanding were prioritized. Typically, this was done through adaptation of a "jigsaw" model (Aronson, 2002; Johnson, Maruyama, Johnson, Nelson, & Skon, 1981) in which students were first given one of three possible worksheets (Figure 2). Students first learned one of the three subtopics and then reorganized to peer-teach their subtopics to peers who had initially learned about other

subtopics. While peer-teaching, or jigsaw, strategies are not inherently interactive in the ICAP framework, these adapted jigsaws required groups to solve high-level cognitive tasks requiring information from each student and thinking significantly beyond what any one student was given. Unlike the constructive activities, student interactions in the groups were structured through guiding prompts to maximize substantive exchange of ideas, a key element of interactive activities (Chi & Wylie, 2014). These prompts explicitly structured group participation (Bell, 2004; Weinberger, Ertl, Fischer, & Mandl, 2005) by promoting dialogue among students by (1) building "turn taking" into the activity to limit the possibility that one student would dominate the conversation and (2) guiding students to act as facilitators, not lecturers, when teaching the concepts that they learned to their new group to minimize the potential for students to resort to passively lecturing one another. Without prompting, students may fail to determine the key concepts that they have learned prior to trying to answer a review question, or they may jump to assembling each person's pieces of information without first critically analyzing each piece. Our aim was to promote "interactive" engagement by scripting what students should be doing at each step during the activity and structuring how students shared their learning through a "social script" (Kollar, Fischer, & Hesse, 2006; Weinberger et al., 2005) that specifies the way that the learners interact.

To keep the two types of activities as similar as possible, we focused only on scaffolding the interactive activity with social scripts that structured the interaction among group members rather than providing additional conceptual support (Kollar et al., 2006; Weinberger et al., 2005). An example of a social script is

> When the instructor tells you to, form a new group with a Data Set 3 and Data Set 4 expert. Group members will each have 4 min to briefly summarize their gene regulation data and to test their groupmates' understanding with a "check for understanding" question.
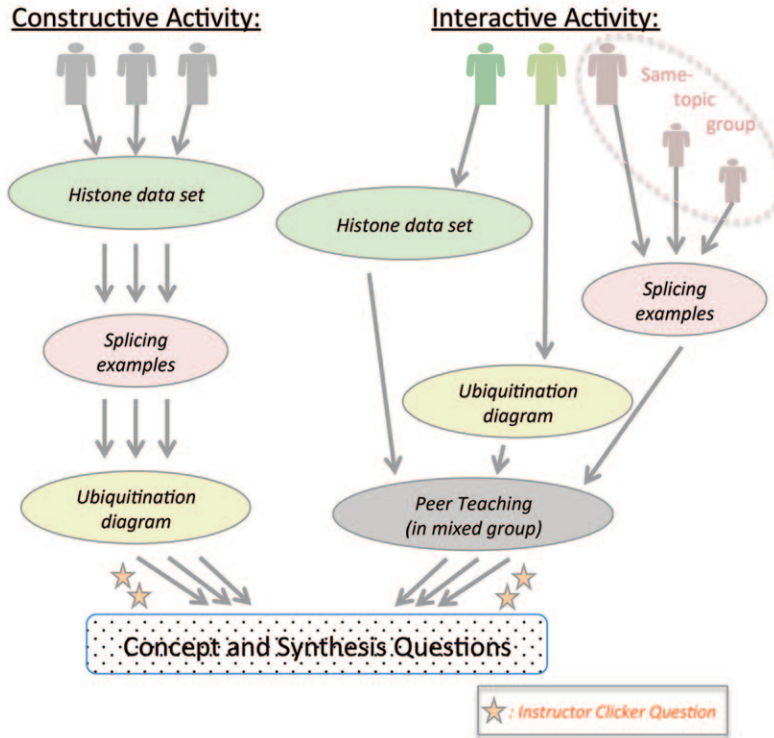
An example of an epistemic script is

FIGURE 2. *Example of the difference between interactive and constructive strategies for a single learning goal. This diagram gives an example of the differences in design between interactive and constructive activities. In the constructive strategy, students collaboratively work through three mechanisms. They then use this conceptual understanding to build new knowledge in the synthesis questions that goes beyond the initial three mechanisms. In the interactive strategy, each student becomes a "micro expert" in one of the mechanisms. Concept questions provide opportunities for students to engage with increasingly higher levels of cognitive difficult on the topic material. The synthesis questions can be successfully completed only by students who interact through debate and justification to parse a correct answer.*

To plan your summary, answer the question below. As you answer the question, write down what elements from your data set/supporting information were critical for you to successfully answer the question.

An example of an epistemic script included in both interactive and constructive activities is

List the key concepts and/or main points that you have learned from the data and the supporting information. In your list include the following information: What proteins are involved in DNA packing and histone acetylation? How does this differ between prokaryotes and eukaryotes?

Sample scripts are included in the online supplementary materials.

Importantly, the two versions of the activities contained identical content. The only difference was in the presence or absence of guiding scripts and the way that the group

5

interaction was structured in class. Group work was facilitated in both versions of activities, even though it would have been possible to create constructive versions without student interaction. Reflection on early versions of similar quasi-experiments convinced us that comparing a group activity with a personal activity would introduce confounds that would have made interpretation of outcome data problematic. In other words, experience demonstrated that it was more important to keep data interpretation simple than to increase the likelihood of observing differences between versions.

Redesign of classroom activities was conducted iteratively across several prior quarters of the same course. Student feedback, including online written evaluations and directed focus groups, was used to develop clearer and more focused activities. Redesign tasks within our research team and with input from colleagues helped to incrementally center each activity into predominantly constructive and predominantly interactive domains. During this process, activities were optimized for clarity, and low cognitive-level questions were removed. In their place were questions that created better links for all activities among elements of the classroom materials. All versions of the activities underwent three to six rounds of such editing. Investigators and instructors held discussions and editing meetings to go through several iterations of each activity, and editing goals were explicitly focused on the multiple audiences of research, on student use, and on scientific accuracy.

Implementation of research activities was overseen by investigators in addition to course instructors. For each quasi-experimental day, at least two investigators were on hand to assist and observe procedural flow. Postclass research meetings involved careful discussions comparing the activity implementation in both sections as well as across quasi-experimental topics. This included comparisons of student concerns and problem areas (as directly observed and reported through teaching assistants working with the course) as well as comparisons of classroom timelines with plans and across sections. Irregularities were noted and discussed before data analysis for comparison.

Measures

Pre- and posttests. To assess student understanding of the key concepts introduced in the in-class activities, we designed multiple-choice tests aligned with each activity's learning goals. Each test contained 8 items focused primarily on assessing higher-order cognitive skills (Bloom, Krathwohl, & Masia, 1956; Crowe, Dirks, & Wenderoth, 2008). For example, a question about protein translation asked,

> You are tasked with genetically engineering a new ribosome that can still do the same job as a normal ribosome but has less mass. You can eliminate the A-site, the P-site, or the E-site. Which one do you eliminate to make a smaller ribosome that would still function in translation?

Correctly answering this novel and imaginative problem requires an understanding of process details and big-picture synthesis of topics from class. Questions were designed to be at exam-level difficulty and were iteratively improved over multiple courses based on student answers and feedback to eliminate confusing elements or grammar. As a measure of content validity, each question was reviewed by at least four experts in cell and molecular biology. Experts were in consensus regarding the scientific accuracy of 88% of the questions. For the remaining 12% of the items, up to two reviewers identified possible exceptions to the expected answers. After reviewing these latter items, we determined that students would have needed expert-level understanding to be aware of the rare exceptions or alternative interpretations that were identified by the expert reviewers. As we are measuring understanding of introductory biology students, we made the decision to include student data from all the questions in our analysis. To test the assumption that each question was measuring the same construct as the rest of the items on the test, we assessed item fit (Bond & Fox, 2001) using the eRM package in R (Mair & Hatzinger, 2007). Based on earlier analyses, three questions were revised to more closely align with the learning goals of the activities. Revised item fit for questions indicated that all items fit within the ranges intended for multiple-choice questions based on INFIT and OUTFIT statistics (Gustafsson, 1980). Rasch analysis of posttest results confirmed that there was a range of item difficulty on each of the tests allowing discrimination among students. For each activity, students completed the 8-item test online on the night prior to the activity as part of a daily reading quiz and then repeated the same 8-item test the night of the activity as part of the daily reading quiz for the subsequent day's lecture.

In-class observations. To document the overall level of student behavioral engagement during the two treatments, we monitored the number of students interacting during the class sessions. Two experienced observers from the university's Center for Teaching and Learning attended each of the six class sessions in which the constructive and interactive activities were implemented. Prior to observation of these class sessions, observers met to discuss the observation protocol and then conducted a trial observation using the proposed protocol. The protocol was refined to ensure that all observers were using consistent criteria for interactive behaviors. To determine whether students were interacting, observers looked for the following behaviors: talking and/or listening to another student, leaning toward and looking at another student, and sharing or looking at the worksheet with another student. The observers then followed the refined observation protocol for each class session. Observers recorded two data points for each row: the total number of students in the row and, of those students, the total number of students visibly engaged in discussion. These two data

points were recorded at two time points during the class session: (1) during the first small group activity as students were working through the activity themselves (constructive) or peer-teaching the others in their group (interactive) and (2) during the second small group activity as students were completing the "synthesis questions" (used for both constructive and interactive treatments). Each row was observed for approximately 1 min during the first activity and 45 s during the second activity. All observation data were recorded on a dedicated observation chart. After the observations were completed, all data were entered into a spreadsheet, and counts of the numbers of interactive students were averaged for each row of the lecture hall, including a total class average for each activity. We assumed that the total number of students per row remained constant within a single class session. The instructor kept each class to a strict time schedule and allotted equivalent time on task for each activity in the constructive and interactive sessions.

Observation data were pooled across topics for constructive and interactive activity days. A similar number of students were observed for each treatment (constructive, *n* = 660; interactive, *n* = 647). Only students in the first 10 rows out of the 15-row classroom were included in the analysis, due to incomplete observation data for the last five rows for two activities. Interrater reliability calculated with the intraclass correlation ranged from moderate to strong depending on the observer pair (0.62–0.85; Landis & Koch, 1977; Portney & Watkins, 2000). One of the observers was present at all six class sessions and performed the count of total number of students per row at the beginning of each class session. We chose to use that observer's data for maximum consistency across treatment and topic.

Statistical Analysis of Student Performance on Posttests

Data were analyzed with a generalized mixed effects model with ordinal regression. Our overall process was to (1) curate a final data set, (2) build and test a final model, and (3) expand on the model to investigate interactions between student demographics and activity type.

Data from this study included repeated measures of students' performance in two contexts: constructive versus interactive activities. Thus, statistical analyses had to account for the nonindependence of the posttest scores (posttest scores of the same student are more likely to be similar than are posttest scores of different students). All pre- and postitems used are available in the online supplementary materials. In addition, posttest scores on any individual activity were not normally distributed; instead, they were left skewed and tightly bounded. Both properties made typical linear regression analysis inappropriate. Instead, we employed a generalized mixed effects model with ordinal regression using the ordinal package in R (Christensen, 2010). Mixed effect models include a random effect term that can account for hierarchical structure in the data (in this case,

multiple posttest scores per student). Ordinal regression treats the posttest score as if it were an ordered categorical measure, which is a reasonable approach in this case because the possible scores on the posttest are tightly bounded (ranging 0–8) and partial credit was not possible. Ordinal regressions model the odds of getting at least one additional question correct on the posttest with an increase in an explanatory variable (e.g., as student grade point average [GPA] increases, the odds that a student will get at least one additional question correct on the posttest increases). Because the study design was quasi-random (students self-selected into two sections), we included cumulative college GPA (as of their participation in the course studied) in the model to control for potential differences in student ability between the two classes. This measure has been shown in prior studies at this institution to strongly predict student performance in the introductory biology series (Eddy, Converse, & Wenderoth, 2015; Freeman, Haak, & Wenderoth, 2011). The model also included measures of demographics, including gender and race (including ethnicity) as described by the university registrar. The university's binary designation category for Education Opportunities Program was used to approximate socioeconomic status, as is customary in this institution.

Data from one of the four quasi-experiments was removed from the statistical analysis. On the day of the first ~750-student experiment, researchers identified procedural errors that confused participants and confounded student engagement in the in-class activities. Students were confused about activity instructions to the point that they were unlikely to productively engage with the tasks given. These observations were corroborated by responses to a survey item asking students whether they agreed that expectations for the assignment were clear. Students in both treatments of this first experiment were far more likely to indicate confusion with the goals of that particular class when compared with all other treatments and experiments. We therefore did not include data from the first experimental day in the final statistical analysis, because these data were drawn from incomparable implementations and would be misleading. Instead, we focus on the remaining three activities that were successfully implemented.

In a preliminary analysis, we explored whether the treatment effect was consistent across the three activities or whether it varied by activity topic (a Treatment × Activity interaction term). We did not find support for this interaction term ($p > .28$ for each comparison) and did not include it in the final model. Thus, our final model was as follows:

$$\text{posttest score} \sim \text{cumulative GPA} + \text{pretest score} + \text{treatment} + (1 \,|\, \text{student ID}).$$

We expanded this model to test for interactions between activity type and several demographic variables, including socioeconomic status (a binary variable indicating whether a student was eligible for the Education Opportunities Program), gender (represented as a binary, as we did not

TABLE 2
Levels of Student Talk During Different Activity Types

	Version	
Observational period	Constructive	Interactive
First		
No. of students talking within their group	560	589
No. of students not talking within their group	100	58
p value	.0007	
Second		
No. of students talking within their group	566	585
No. of students not talking within their group	94	62
p value	.0104	

Note. At two times during the class session, observers noted whether students in a sample of working groups were talking to each other. The data were pooled across three class sessions for each version of the activity. More students were talking with their groupmates during interactive versions of the activities. *P* values based on Fisher's exact test (2-tailed).

have the sample size to test the impact of activities on students who did not identify as male or female), and race as defined by federally recognized underrepresented minority populations (Office of Minority Affairs and Diversity, 2016). Likelihood ratio tests were used to test the goodness of fit between the base model and alternative models including these demographic variables and interaction terms.

Results

Did Interactive Activities Lead to More Student Interaction Than Constructive Activities?

Consistent with the increased structure and role assignments built into the interactive activities, we found significantly more students talking to one another during the interactive activities than the constructive activities (Fisher's exact test, $p < .001$, $p = .01$, respectively, for the two observation time points) as shown in Table 2.

Did Interactive Activities Lead to Higher Learning Gains Than Constructive Activities?

Results for pre- and postscores for the three quasi-experimental activities are shown in Figure 3.

Model estimates from the ordinal regression analyses (shown in Table 3) show a significant effect of interactive activity on student learning gains. On an eight-item content quiz in a pre- and postformat, a student taught with an interactive strategy was 25% more likely to answer at least one additional question correctly on the posttest than that same

student taught with a constructive strategy. This change is similar in magnitude to the difference that we would expect on the posttest for a student who has a cumulative GPA that is a quarter point higher than another student's. This supports the prediction made within the ICAP framework (Chi & Wylie, 2014).

Did the Impact of Interactive Activities Differ Among Groups of Students?

By adding demographic variables to the model, we could test whether the impact of interactive activities was greater for students of different backgrounds. Likelihood ratio test results (shown in Table 4) reveal that adding a main effect for socioeconomic class, gender, or race or an interaction term between treatment and that particular identity did not increase the fit of the model to the data ($p = .9975$, .3375, and .9789, respectively). Thus, we find no significant evidence that the impact of completing a constructive versus interactive activity varied among student groups.

Discussion

Given infinite time and resources, all classrooms are likely to be best taught with an ambitious blend of multiple strategies that have been honed over time with similarly diverse students. In reality, college STEM classrooms are often taught in whatever method is most logistically expedient. For those who would produce positive change in these learning environments and, thus, on their outcomes for graduates and future scientists, deeper investigation into costs and benefits is crucial. Practical limitations put pressure on the STEM education workforce to make well-informed decisions about the small number of focused changes that can be implemented.

We have reported quasi-experimental results indicating a significant benefit for student learning from interactive activities versus constructive activities. This not only supports a prediction in previous theoretical work (Chi & Wylie, 2014), but also informs instructional design choices that are being made in STEM classrooms as instructors implement active learning. While other parts of this framework have conceptual or lab-based support, our work uses rigorous quasi-experimentation in a "field test" of active learning in a large-enrollment STEM classroom environment. When controlling for the learning environment, student characteristics and specific aspects of the activities (e.g., the total time on task or amount of group work), we see a small but significant benefit from interactive learning as compared with even very active learning strategies categorized as constructive. This can inform implementation of best practices in and perhaps beyond STEM classrooms.

In this study, we explored the relative impact of interactive activities on student learning in a college classroom. An
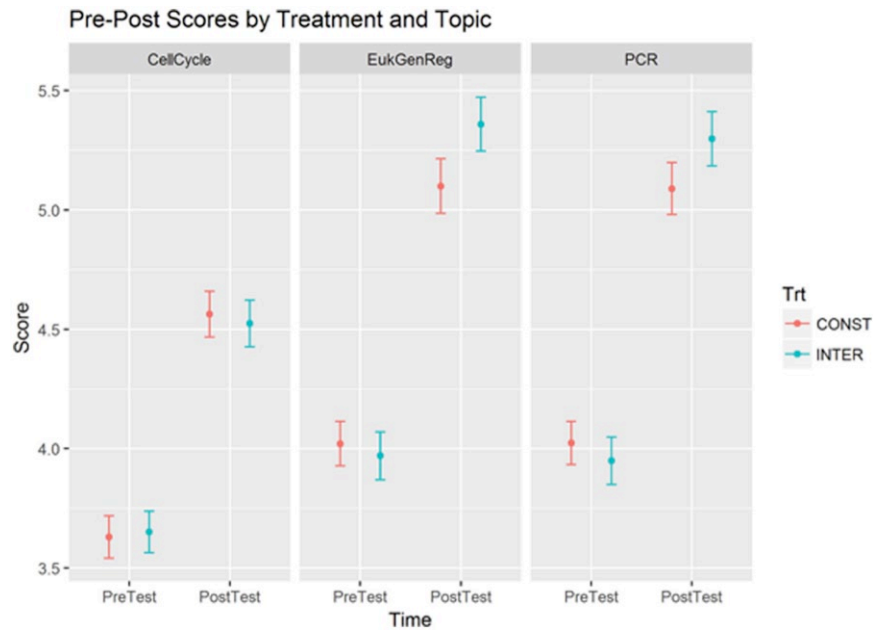
FIGURE 3. *Student pre- and postscores across activities. Aggregated student results are shown here from activities on pre- and postquizzes. Classes based on a constructive approach are shown in red, while classes based on an interactive approach are shown in teal. All quizzes had 8 equally valuable questions, and the same questions were given for each pre- and postquiz pair and comparison. This figure shows that prescores are similar between groups and that the overall pattern of improved student gains in interactive classrooms is not due to a single activity.*

TABLE 3
Interactive Activities Benefit Students

Model terms	Regression coefficient (log odds) ± *SE*	*p* value	Odds ratio
Treatment: interactive (reference: constructive)	0.219 ± 0.0877	.0123	1.24
Cumulative college GPA (mean centered)	0.978 ± 0.1432	<.00001	2.66
Pretest score (mean centered)	0.628 ± 0.0308	<.00001	1.87

Note. Across the three activity days, students given interactive activity were more likely to answer at least one additional question correct on the posttest relative to students given a constructive version of the same activity, after controlling for a student's cumulative college GPA and his or her score on the pretest for that activity. Results are based on the following ordinal regression model: posttest score ~ treatment + mean-centered cumulative college GPA + mean-centered pretest score + (1|student ID). Only the fixed effects are shown in the table. These model results demonstrate a positive effect on learning observed in the interactive classrooms sessions and gives context for this effect by comparison with other predictors in the model. GPA = grade point average.

TABLE 4
All Students Benefit from Interactive Activities

Term added	AIC for model		*p* value of likelihood ratio test
	Without bold term	With bold term	
+ Gender	6520.9	6522.0	.338
Gender + **Gender × Treatment**	6522.0	6523.5	.4962
SES	6520.9	6522.9	.998
SES + **SES × Treatment**	6522.9	6524.1	.3712
+ Race	6408.0	6413.8	.9780
+ Race + **Race × Treatment**	6413.8	6419.2	.9133

Note. Student demographics did not increase the fit of the models to the data, so they were not included in the final model. Using forward model selection, we stepwise added a main effect term and compared this expanded model with the base model using a likelihood ratio test. We then compared a model with only the main effect of the demographic variable with a model containing an interaction between the term and the treatment. The base model was as follows: posttest score ~ cumulative GPA + pretest score + treatment. According to the table data, student demographic data were not important predictors of differences in student learning gains. AIC = Akaike information criterion; SES = socioeconomic status; GPA = grade point average.

intrinsic limitation to this approach is the inability to ensure that students participating in the two types of activities (interactive and constructive) were indeed engaging in intended

behaviors. Classroom observations indicate increased inter-action among students completing the interactive activities relative to the constructive activities. However, students completing the constructive activity may not have been engaged in the interactive behavior of co-constructing knowl-edge. Similarly, students participating in the interactive activi-ties may have been engaged only in the constructive process of building their own knowledge. This is true of any activity in real classrooms: Even students engaged in the seemingly identical process of highlighting a text passage may be oper-ating at a simply active level of engagement or instead be internally summarizing ideas and making connections among different ideas as they highlight. We can likely assume (1) that a certain percentage of the students completing the con-structive activity were actually engaged in co-construction and thus working at an interactive level and (2) that a certain percentage of the students participating in the interactive activity were not building off of one another's ideas but instead operating at a constructive level. In this case, the observed small difference in student learning between the two modes of instruction may be relevant to real classrooms while remaining a conservative estimate of the relative ben-efit that an interactive activity might demonstrate in an ideal environment. However, since these activities will be imple-mented in large classrooms where students are unlikely to follow guidelines explicitly and there will likely be other fidelity-of-implementation issues, it is important to recognize that the theoretical gains of an interactive activity may not be fully realized but can still result in a small and valuable increase in learning.

The 24% gain observed is encapsulated in the possible correct choice of at least one multiple-choice question on each quiz. STEM students answer many thousands of multi-ple-choice questions within their overall education, so these four posttests represent a tiny part of their total output and are only a partial representation of their actual learning. Prior success in university courses (measured by GPA) and prequiz scores are better predictors of posttest outcomes, indicating that this effect is subtle. How much stock should we put in this observed learning benefit? There are good rea-sons to be skeptical. Multiple-choice questions may be answered correctly for a variety of reasons unrelated to deep conceptual knowledge (Darling-Hammond & Adamson, 2014; Stanger-Hall, 2012). Furthermore, no single interven-tion demonstrated a statistically significant increase in class population performance on a pre- and postquiz. However, the subtlety of the treatment makes the presence of a signifi-cant signal noteworthy. Indeed, several previous studies in college classrooms that compared predominantly interactive activities with activities that could be classified as construc-tive did not find support for the predicted benefit of interac-tive activities (Kam et al., 2005; Linton, Pangle, et al., 2014; Menekse et al., 2013). There are several possible explana-tions for the lack of significant differences in these studies,

as compared with studies performed in controlled laboratory settings. First, many of the studies had small sample sizes ($n < 50$) and so may not have had the statistical power neces-sary to detect small differences in learning gains. Second, fidelity of implementation is difficult to achieve in any eco-logical classroom setting. Third, optimal experimental design is often not compatible with an existing course syl-labus or classroom logistics. Results may be confounded by other factors, such as the order in which students complete different types of activities. For example, if students com-plete an interactive activity prior to a constructive activity on the same topic, student performance on the constructive activity may be inflated due to an increased understanding gained from prior engagement during the interactive activity (Menekse et al., 2013).

In the study reported here, the large student population and use of a repeated measures study design may have been necessary to detect a signal resulting from these subtle changes in the mode of instruction. While the effects are subtle, they are still on a comparable scale to two intrinsic factors (GPA and prior knowledge) that reflect years of stu-dent learning. At this scale, it is impressive to see a result based on a slight shift in activity design (1.24×) that is almost two thirds of the magnitude of all prior student prior learning that influenced the prequiz score (1.87×). Last, it is worth noting that pre- and postexams earned credit for students based on nongraded participation. For this reason, the observed increase in relative score may have been somewhat conservative as a proxy for student learning.

The learning outcomes observed here may be additive when interactive education is used repeatedly. Indeed, a course-long study comparing students who completed in-class activities in groups and individually showed that stu-dents who worked interactively performed better on exam questions requiring higher-order cognitive skills (Linton, Farmer, & Peterson, 2014). Given the complicated and social nature of human learning (Lave & Wenger, 1991), it may be that any small intervention with an observed statisti-cally significant improvement in outcome may hint at the rest of the iceberg of possible benefits related to slight instructional shift in practice. Increases in learning gains may benefit students in learning later topics more quickly (due to their improved and more receptive background) or in learning related topics more deeply (due to a smaller accu-mulation of misconceptions).

Intensely active strategies at the top of the ICAP spec-trum are likely to bring intrinsic benefits not captured within our result. Constructive and interactive methods give stu-dents opportunities to practice collaborative work, which will ideally lead to improved social skills and training for collaborative work in science (Fine & Harrington, 2004; Rosenberg, Lorenzo, & Mazur, 2006). Both methods have the capacity to position students in a growth mind-set as apprenticing experts instead of passive receivers as in

traditional lecture classrooms (Nasir & Hand, 2006; Yeager & Dweck, 2012). Both activities are likely to result in improved learning outcomes similar to those observed in a variety of STEM teaching environments from even the partial use of active learning strategies (Dauer & Long, 2015; Freeman et al., 2014). Indeed, more inclusive and more effective classrooms in other fields and types of institutions may already tend toward these methods (Pascarella & Blaich, 2013).

In addition to the shared benefits of constructive and interactive activities, there are unique benefits that come from engaging in an interactive activity. For example, interactive methods have the special characteristic of placing students as "micro experts," which gives them opportunities to improve attitudinal outcomes, such as confidence and grit (Duckworth, Peterson, Matthews, & Kelly, 2007). Also, there is an increased level of accountability built into the interactive activity since all members of the group must contribute, which may help causally explain the learning gains seen here. Last, the increased structure built into the interactive activity may create more equitable group dynamics by explicitly tasking each group member with an active role. While these likely gains are important for the development of next-generation scientists, all are extra benefits unrelated with the positive results described here.

We found no evidence for disparities in learning gains between the two different instructional strategies in reference to different demographic groups of students. We cannot rule out that there may be greater diversity within registrar-delineated groups than between these groups or that there may be a lack of power in our analytic procedures and statistical analysis. Further research would clearly need to be done to demonstrate equity among intensely active instructional strategies. These studies should include inspection of finer-grain analyses of student intersectionality/identity, as well as a better understanding of the diverse relationships within working groups (Eddy et al., 2015). For now, it is at least encouraging, given commonly observed gaps in student outcomes, that our close analysis does not demonstrate obvious learning gaps along ethnic, racial, gender, or first-generation-status groups. This suggests that highly active learning strategies may be a partial solution in light of historically intransigent gaps in STEM education outcomes for underrepresented groups and that these highly active learning strategies may represent a more equitable mode of instruction. More work is needed.

Several important factors are beyond our observation. Prior student experience with well-facilitated active learning may influence the benefits that students receive from these classroom strategies. The classrooms used in the quasi-experiments described here employed intensely active learning strategies within an environment in which active learning was a daily norm. The benefits of interactive activities might be predicted to be higher in classrooms where students were already acculturated to their use, or they might be predicted to be lower in situations where the novelty and increased classroom energy would wane after repeated usage. While we have no instrument with which to measure "shyness" in students, this might be a predictor for relatively poor performance in more conversationally based teaching modes. However, it is also possible that students may have had previous negative experiences with active learning and were therefore less willing to fully engage. While it is likely that a diversity of instructional methods is best, the extent to which prior student experience with active learning mediates these gains will require further research.

Implications

Improved student outcomes suggest that interactive teaching strategies are likely to be the superior mode of intensely active instruction. These results provide an initial suggestion that interactive instruction may be the best option for STEM classrooms in which instructors are well resourced. To adjust curricula and implement more interactive student activities, undergraduate instructors could engage in a combination of professional development, skills-based practice, mentored instructional design, and observation of peers already utilizing these strategies effectively. These investments will help to guide iterative improvement of courses and sessions even around cutting-edge topics in science.

Achieving sustainable and faithful adoption of research-based teaching practices has proven to be extremely difficult (Henderson & Dancy, 2007). It is therefore very important to weigh the relative costs and benefits of different learning strategies prior to investing resources in curricular redesign. Interactive strategies may require more from instructors both in development time and class management skills. In our experience, implementing the interactive activities in a large lecture hall required a higher level of organization and preplanning to ensure that each student initially received a different activity packet and then was able to easily identify nearby students who had worked on the other packets for the "sharing" portion of the activity. During debriefing after the activity, instructors also indicated that the interactive activities required more effort in class management than the constructive activities due to the need to ensure that all students had sufficient time to "teach" their sections to the group. It may therefore be best to limit the use of interactive methods to those topics most easily adapted for student use in this format (e.g., subjects for which the learning goals require conceptual understanding of multiple mechanisms). Greater overall improvement of student learning might be provided, not by moving activities from constructive to interactive, but rather by targeting those resources to shifts from passive to active instruction or from active to constructive. More research is needed to better understand the potential gains to students and instructors from interactive or constructive

methods. This will help to better guide instructional development decisions at many levels. Given the lack of uniformity of quality active learning in postsecondary education, developing interactive activities might not be the "low-hanging fruit" for change that would best serve students.

Student learning is not a simple scale; the benefits of instructional choices will be mediated by complex characteristics of individual and group learning within a dynamic cultural environment (Bang & Medin, 2010). For active learning, these mediations remain incompletely understood. Active classrooms may have some benefit for groups traditionally underserved by more didactic instruction (Freeman et al., 2014). The extent to which this benefit reaches all students will require deeper cultural research with implementation of intensely active experiences. Our data suggest no overt link among race, gender, and added benefit from interactive activities when compared with constructive. This may indicate that we do not have the power or the breadth to understand these links yet, that students benefit equitably from increased activity, or that predictors for differential benefit are more complex and/or social than those to which we have access. This research will be most useful when conducted locally by practitioners to best inform instructional choices

Conclusion

Through a repeated measures quasi-experimental design, these data support the prediction of the highest contrast in the ICAP framework that had not previously been examined in ecological classroom studies. Specifically, student outcomes were improved in a STEM classroom when taught in an interactive manner versus a constructive manner. No differences were observed in the relative impact on students from different underrepresented groups. Teachers, even with postsecondary audiences, must make instructional decisions that influence student social interactions. These instructional decisions are likely to be more important as postsecondary courses continue to transition from passive lecture to more involved and socially engaged models of instruction especially in large college classrooms even beyond STEM. To best understand the use of intensely active teaching strategies, the professional development and discipline-based practices of instructors must be better understood. Future research into the predictions of frameworks around the use of active learning strategies must necessarily engage directly in the dialogue within and between students and instructors.

Acknowledgments

References

American Association for the Advancement of Science. (2009). *Vision and change in undergraduate biology education*. Washington, DC: Author.

Aronson, E. (2002). Building empathy, compassion, and achievement in the jigsaw classroom. *Educational Psychology*, *10*, 209–225.

Azevedo, F. S. (2013). The tailored practice of hobbies and its implication for the design of interest-driven learning environments. *Journal of the Learning Sciences*, *22*(3), 462–510.

Bang, M., & Medin, D. (2010). Cultural processes in science education: Supporting the navigation of multiple epistemologies. *Science Education*, *94*(6), 1008–1026. doi:10.1002/Sce.20392

Bell, P. (2004). Promoting students' argument construction and collaborative debate in the science classroom. In *Internet environments for science education* (pp. 115–143). Mahwah, NJ: Erlbaum.

Bloom, B. S., Krathwohl, D. R., & Masia, B. B. (1956). *Taxonomy of educational objectives: The classification of educational goals*. New York, NY: D. McKay.

Bond, T. G., & Fox, C. M. (2001). *Applying the Rasch model: Fundamental measurement in the human sciences* (2nd ed.). Mahwah, NJ: Erlbaum.

Chen, R. F., Eisenkraft, A., Fortus, D., Krajcik, J., Neumann, K., Nordine, J., & Scheff, A. (2014). *Teaching and learning of energy in K–12 education*. New York, NY: Springer.

Chi, M. T. (2009). Active-constructive-interactive: A conceptual framework for differentiating learning activities. *Topics in Cognitive Science*, *1*(1), 73–105. doi:10.1111/J.1756-8765.2008.01005.X

Chi, M. T., Roy, M., & Hausmann, R. G. (2008). Observing tutorial dialogues collaboratively: Insights about human tutoring effectiveness from vicarious learning. *Cognitive Science*, *32*(2), 301–341.

Chi, M. T., & Wylie, R. (2014). The ICAP framework: Linking cognitive engagement to active learning outcomes. *Educational Psychologist*, *49*(4), 219–243. doi:10.1080/00461520.2014.965823

Christensen, R. H. B. (2010). Ordinal—regression models for ordinal data. *R Package* (version 2012.09-11) [Computer software]. Retrieved from https://cran.r-project.org/web/packages/ordinal/

Coleman, E. B., Brown, A. L., & Rivkin, I. D. (1997). The effect of instructional explanations on learning from scientific texts. *Journal of the Learning Sciences*, *6*(4), 347–365. doi:10.1207/s15327809jls0604_1

Crowe, A., Dirks, C., & Wenderoth, M. P. (2008). Biology in Bloom: Implementing Bloom's taxonomy to enhance student learning in biology. *CBE—Life Sciences Education*, 7, 368–381.

Darling-Hammond, L., & Adamson, F. (2014). *Beyond the bubble test: How performance assessments support 21st century learning*. New York, NY: Wiley.

Dauer, J. T., & Long, T. M. (2015). Long-term conceptual retrieval by college biology majors following model-based instruction. *Journal of Research in Science Teaching*, *52*(8), 1188–1206.

Duckworth, A. L., Peterson, C., Matthews, M. D., & Kelly, D. R. (2007). Grit: Perseverance and passion for long-term goals. *Journal of Personality and Social Psychology*, *92*(6), 1087.

Eddy, S. L., Converse, M., & Wenderoth, M. P. (2015). PORTAAL: A classroom observation tool assessing evidence-based teaching practices for active learning in large science, technology, engineering, and mathematics classes. *CBE–Life Sciences Education*, *14*(2), ar23. doi:10.1187/cbe.14-06-0095

Fine, G. A., & Harrington, B. (2004). Tiny publics: Small groups and civil society. *Sociological Theory*, *22*(3), 341–356.

Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., & Wenderoth, M. P. (2014). Active learning increases student performance across STEM disciplines. *Proceedings of the National Academy of Sciences of the United States of America*, *11*(23), 8410–8415.

Freeman, S., Haak, D., & Wenderoth, M. P. (2011). Increased course structure improves performance in introductory biology. *CBE–Life Sciences Education*, *10*(2), 175–186. doi:10.1187/cbe.10-08-0105

Gay, G. (2010). *Culturally responsive teaching: Theory, research, and practice*. New York, NY: Teachers College Press.

Gobert, J. D., & Clement, J. J. (1999). Effects of student-generated diagrams versus student-generated summaries on conceptual understanding of causal and dynamic knowledge in plate tectonics. *Journal of Research in Science Teaching*, *36*(1), 39–53.

Gustafsson, J. E. (1980). Testing and obtaining fit of data to the Rasch model. *British Journal of Mathematical and Statistical Psychology*, *33*(2), 205–233.

Handelsman, J., Miller, S., & Pfund, C. (2007). *Scientific teaching*. New York, NY: Freeman.

Henderson, C., & Dancy, M. H. (2007). Barriers to the use of research-based instructional strategies: The influence of both individual and situational characteristics. *Physical Review Special Topics: Physics Education Research*, *3*(2), 020102.

Johnson, D. W., Maruyama, G., Johnson, R., Nelson, D., & Skon, L. (1981). Effects of cooperative, competitive, and individualistic goal structures on achievement: A meta-analysis. *Psychological Bulletin*, *89*(1), 47.

Kam, M., Wang, J., Iles, A., Tse, E., Chiu, J., Glaser, D., . . . Canny, J. (2005). Livenotes: A system for cooperative and augmented note-taking in lectures. In *Proceedings of the SIGCHI Conference on Human factors in computing systems* (pp. 531–540). New York, NY: *ACM*.

Kollar, I., Fischer, F., & Hesse, F. W. (2006). Collaboration scripts—A conceptual analysis. *Educational Psychology Review*, *18*(2), 159–185.

Kramarski, B., & Dudai, V. (2009). Group-metacognitive support for online inquiry in mathematics with differential self-questioning. *Journal of Educational Computing Research*, *40*(4), 377–404.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*(1), 159–174. doi:10.2307/2529310

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. New York, NY: Cambridge University Press.

Linton, D. L., Farmer, J. K., & Peterson, E. (2014). Is peer interaction necessary for optimal active learning? *CBE–Life Sciences Education*, *13*(2), 243–252. doi:10.1187/cbe.13-10-0201

Linton, D. L., Pangle, W. M., Wyatt, K. H., Powell, K. N., & Sherwood, R. E. (2014). Identifying key features of effective active learning: The effects of writing and peer discussion. *CBE–Life Sciences Education*, *13*(3), 469–477. doi:10.1187/cbe.13-12-0242

Mair, P., & Hatzinger, R. (2007). Extended Rasch modeling: The eRm package for the application of IRT models in R. *Journal of Statistical Software*, *20*(i09).

Menekse, M., Stump, G. S., Krause, S., & Chi, M. T. (2013). Differentiated overt learning activities for effective instruction in engineering classrooms. *Journal of Engineering Education*, *102*(3), 346–374. doi:10.1002/jee.20021

Nasir, N. S., & Hand, V. M. (2006). Exploring sociocultural perspectives on race, culture, and learning. *Review of Educational Research*, *76*(4), 449–475. doi:10.3102/00346543076004449

Office of Minority Affairs and Diversity. (2016). *2015–2016 fact sheet*. Seattle, WA: University of Washington.

Oleson, A., & Hora, M. T. (2014). Teaching the way they were taught? Revisiting the sources of teaching knowledge and the role of prior experience in shaping faculty teaching practices. *Higher Education*, *68*(1), 29–45. doi:10.1007/s10734-013-9678-9

Pascarella, E. T., & Blaich, C. (2013). Lessons from the Wabash National Study of Liberal Arts Educaton. *Change: The Magazine of Higher Learning*, *45*(2), 6–15.

Portney, L., & Watkins, M. (2000). *Foundations of clinical research applications to practice*. Hillsdale, NJ: Prentice Hall.

President's Council of Advisors on Science and Technology. (2012). *Engage to excel: Producing one million additional college graduates with degrees in science, technology, engineering, and mathematics*. Arlington, VA: Networking and Information Technology Research and Development.

Rosenberg, J. L., Lorenzo, M., & Mazur, E. (2006). Peer instruction: Making science engaging. In J. J. Mintzes, & W. H. Leonard (Eds.), *Handbook of college science teaching* (pp. 77–85). Arlington, VA: National Science Teachers Association.

Rummel, N., Spada, H., & Hauser, S. (2009). Learning to collaborate while being scripted or by observing a model. *International Journal of Computer-Supported Collaborative Learning*, *4*(1), 69–92.

Stanger-Hall, K. F. (2012). Multiple-choice exams: An obstacle for higher-level thinking in introductory science classes. *CBE–Life Sciences Education*, *11*(3), 294–306. doi:10.1187/Cbe.11-11-0100

Weinberger, A., Ertl, B., Fischer, F., & Mandl, H. (2005). Epistemic and social scripts in computer-supported collaborative learning. *Instructional Science*, *33*(1), 1–30.

Yeager, D. S., & Dweck, C. S. (2012). Mindsets that promote resilience: When students believe that personal characteristics can be developed. *Educational Psychologist*, *47*(4), 302–314.

Authors

BENJAMIN L. WIGGINS is a lecturer and the manager of instruction at the University of Washington biology department. Ben's work focuses on implementation of active learning in STEM classrooms as well as tactical improvement for instructors at multiple levels of higher education.

SARAH L. EDDY is an assistant professor in the Department of Biology and in the STEM Transformation Institute at Florida International University. Her research focuses on identifying factors influencing student persistence in undergraduate science majors for the purpose of increasing retention, especially of students from historically underserved populations.

DANIEL Z. GRUNSPAN is a postdoctoral scholar at the Center for Evolution and Medicine at Arizona State University. His research interests include the formation and importance of under-graduate networks, equity in STEM classrooms, and evolutionary medicine education.

ALISON J. CROWE is a principal lecturer in the Department of Biology at University of Washington. Her research interests include developing programmatic level assessments to facilitate curricular reform and designing active learning exercises that promote equity in the classroom.