Check for updates

Characteristics of interactive classrooms that first year students find helpful

Kristen Vroom[1*], Jessica Gehrtz[2], Naneh Apkarian[3], Tenchita Alzaga Elizondo[4], Brittney Ellis[4] and Jessica Hagman[5]

Abstract

Background: Implementing research-based teaching practices   has been repeatedly cited as an important factor for student success in university mathematics courses. Many research-based practices increase the amount of student–student and/or student–instructor interaction. However, some instructors are hesitant to implement such practices because they anticipate their students reacting negatively to experiencing an interactive classroom. As part of a larger project studying introductory undergraduate mathematics courses in the United States, we investigated students' perceptions of the helpfulness of various classroom characteristics, particularly those that require interaction.

Results: From analyzing quantitative student data, we found that students reported interactive classroom characteristics (e.g., group work) as less prevalent than other classroom characteristics (e.g., lecture). Moreover, the students tended to regard characteristics that they reported experiencing often as helpful for their learning. From analyzing qualitative data from student focus groups, we found that students considered several indicators when identifying if a characteristic was helpful for their learning. In particular, students suggested that they can identify a characteristic as helpful for their learning when it supported them in solving assigned problems and understanding why the procedures work, earning good grades, building on their knowledge or applying it in different contexts, and teaching others.

Conclusions: The key finding from our work is that students are likely to view classroom characteristics that they experience more often as more helpful for their learning and are less likely to view characteristics that they rarely experience as helpful for their learning. Students view the characteristics that they regularly experience as helping them to solve problems and understand why the procedures work, earn good grades, build on their knowledge or apply it in different contexts, and teach others. We discuss important implications for practice, policy, and research as it relates to both student and instructor buy-in for increasing interactions in class.

Keywords: Research-based teaching, Student interactions, Interactive classrooms, Student buy-in, Instructor buy-in

Introduction

Implementing research-based teaching practices has been repeatedly identified as a salient factor for student success in university mathematics courses (Ellis et al., 2014; Freeman et al., 2014; Seymour & Hewitt, 1997; Seymour & Hunter, 2019). We define "research-based teaching practices" as those practices that have been identified in science and mathematics education research as positively impacting student outcomes and success. For example, it is well-documented that incorporating time in class for students to have opportunities to work individually or in groups can increase students' conceptual learning and reduce achievement gaps based on gender, race and ethnicity, and income (e.g., Eddy & Hogan, 2014; Freeman et al., 2014; Kogan & Laursen, 2014; Laursen et al., 2014; Theobald et al., 2020). As a

*Correspondence: vroomk@oregonstate.edu

[1] Department of Mathematics, Oregon State University, Kidder Hall 368, Corvallis, OR 97331, USA
Full list of author information is available at the end of the article

result, there has been increased pressure to reform collegiate mathematics teaching in a way that aligns with research recommendations for more student-centered approaches (CBMS, 2016; NRC, 2013; PCAST, 2012; Saxe & Braddy, 2015). However, in undergraduate science, technology, engineering and mathematics (STEM) courses, traditional didactic lecture remains the predominant instructional approach (Rasmussen et al., 2019; Stains et al., 2018). Across STEM, education researchers have documented a wide range of reasons why individual instructors may choose not to incorporate research-based practices into their teaching or choose to stop using such practices after brief periods of usage. These reasons include particular beliefs about learning (Aragón et al., 2018; Sturtevant & Wheeler, 2019), a lack of personal experience with the methods (Andrews et al., 2015; Shadle et al., 2017), contradictory departmental or disciplinary contexts (Henderson et al., 2012; Reinholz & Apkarian, 2018), and even the physical barriers of the classroom (Apkarian et al., 2021; Foote et al., 2014; Knaub et al., 2016). Another prominent factor is that instructors fear their students will react negatively to the practices or report a poor experience on their student evaluations (Dancy & Henderson, 2012; Finelli et al., 2014; Froyd et al., 2013; Hayward et al., 2016; Shadle et al., 2017).

The relationship between research-based practices and student success paired with instructors' assumptions about students' reactions to classroom experiences that features these practices motivates the need for more research. Specifically, there is a need to better understand students' experiences with research-based practices, factors that contribute to students' perceptions of these experiences, and (if students do react negatively to particular experiences shaped by implementing research-based practices) ways to increase student buy-in regarding how these experiences can impact their success in the course. In this study, we investigated students' perceptions of a common feature of many research-based practices: increased interaction between the instructor and students and/or interaction among the students. In what follows, we elaborate on the notion of practices that necessitate interaction and discuss related literature to situate our study.

Teaching that promotes interaction in undergraduate STEM

Increased interaction in research-based teaching practices

As a response to research indicating the harm done to students through lecture and the benefit to students of increased interactions in class, there have been calls to change how undergraduate mathematics courses are taught, with particular calls to increase interactions between instructors and their students and students with their peers (e.g., CBMS, 2016; Saxe & Braddy, 2015). For instance, the Conference Board of Mathematical Sciences (CBMS, 2016) urged institutes of higher education to invest time and money to implement teaching practices that engage students in group problem-solving where they can get feedback from both experts and their peers.

Many scholars studying research-based teaching practices stress that students interacting with the instructor and each other are critical components of such teaching (e.g., Jacobs & Spangler, 2017; Lampert et al., 2010; Larsen et al., 2015; Laursen & Rasmussen, 2019; Stein et al., 2008). For instance, a defining characteristic of inquiry-based instruction is that instructors inquire into their students' thinking (Laursen & Rasmussen, 2019), which requires eliciting and building on students' ideas to drive the mathematical agenda of the class. That is, instructors must interact with students to meaningfully engage with their ideas in order to recognize their mathematical and pedagogical value (Kuster et al., 2018; Leatham et al., 2015; Speer & Wagner, 2009). Researchers have also emphasized the importance of instructors supporting students in engaging with each other's mathematical ideas. In order to facilitate a genuine mathematical discussion, instructors need to support students' engagement with each other rather than the instructor being the sole person engaging with student contributions (Jacobs & Spangler, 2017; Rasmussen et al., 2009; Staples, 2007; Stein et al., 2008).

It is clear that research on teaching urges practice to move away from teacher-centered approaches to student-centered approaches that support students in interacting with the instructor and each other. Thus, we argue that increasing student interactions with their instructor and peers is a central theme of research-based teaching practices. In order to study how students experience research-based teaching practices, we must understand how they experience such interactions in mathematics classrooms.

Value and usage of teaching that promotes interactions in undergraduate STEM

It has been well-demonstrated across contexts that increased student interactions with other students and instructors during class time has a positive impact on student learning, success, and persistence in STEM (Freeman et al., 2014; Hake, 1998). These results are not new, with research-based recommendations for increased interactivity in undergraduate STEM dating back to (at least) the late 1990s (Hake, 1998). However, in more recent years the preponderance of evidence has grown significantly (see Freeman et al., 2014 for a meta-analysis of such studies) and led to more formal, and stronger, recommendations for student-centered instructional practices from professional societies, including

Vroom *et al. International Journal of STEM Education* (2022) 9:38

Page 3 of 17

mathematics specifically (e.g., CBMS, 2016). In addition to research that links interactive learning approaches to student success within a particular course (e.g., Freeman et al., 2014; Larsen et al., 2013), studies have identified didactic lectures in undergraduate STEM courses as a reason why students leave STEM majors (Seymour & Hewitt, 1997; Seymour & Hunter, 2019). Moreover, of the students who leave STEM, women and students of color are the ones to least likely report experiencing courses with interactivity (Rainey et al., 2019). Further, when students report interactive instructional practices in STEM courses, the likelihood that they will switch out of calculus-intensive majors decreases (Ellis et al., 2014; Rasmussen & Ellis, 2013). While there is growing support for interactive learning that includes "communication and group problem-solving [and] receiving feedback on their work from both experts and peers" (CBMS, 2016), we note that more research is needed to understand practices that promote interactions from both teachers and students' perspectives (Johnson et al., 2020; Kogan & Laursen, 2014; Stains & Vickrey, 2017; Theobald et al., 2020).

We repeat a common refrain of undergraduate STEM education literature: "despite active learning being recognized as a superior method of instruction in the classroom [...] most college STEM instructors still choose traditional teaching methods" (Deslauriers et al., 2019, p. 1). The latter claim is supported by a major study of gateway undergraduate STEM courses using observation data (Stains et al., 2018), a national survey of chemistry, mathematics, and physics instructors (Apkarian et al., 2021), and earlier work from this project in university introductory mathematics (Apkarian & Kirin, 2017; Rasmussen et al., 2019).

In light of the evidence supporting the use of teaching practices that increase interaction and the predominant use of lecture by college STEM instructors, recent research has examined reasons why instructors choose to use (or not use) these practices and how to increase their usage. The results of this research suggest that there are many interrelated individual and systemic factors impacting instructors' pedagogical choices (Henderson & Dancy, 2007; Shadle et al., 2017; Sturtevant & Wheeler, 2019). For example, university science instructors who have beliefs consistent with a growth mindset are more likely to use research-based practices (Aragón et al., 2018); abstract algebra instructors who believe lecture is the best way to teach the requisite material in the given timeframe spend more class time lecturing (Johnson et al., 2018); STEM instructors implementing practices that increase interaction believe that collaboration and discussion are fundamental components of learning while 'chalk talkers' believe their role is to "model

problem solving through examples and demonstrations during class" (Ferrare, 2019, p. 12).

Most relevant to this paper is the belief that students will react negatively to increased interaction during class time. College STEM instructors frequently reported students' negative reactions as a barrier to both starting and continuing to implement research-based teaching, often in conjunction with concerns about end-of-term student evaluations (Dancy & Henderson, 2012; Finelli et al., 2014; Froyd et al., 2013; Hayward et al., 2016; Henderson & Dancy, 2007; Shadle et al., 2017). Although end-of-term student evaluations have been deemed an ineffective and inequitable assessment tool (Boring, 2017; Chavéz & Mitchell, 2020; Fan et al., 2019), student judgements of instructors are still frequently used for instructor evaluation and thus this is a barrier that disproportionately affects certain populations of instructors like women, instructors of color, and instructors with non-English accents (Fan et al., 2019).

Students' perceptions of interactive classrooms

Much of the existing research on the benefits of research-based teaching practices for undergraduate STEM students does not incorporate students' perspectives. That is, researchers have designated particular markers for success (e.g., pass rates, attitudes toward math, persistence), established a link between those markers and the use of teaching methods, and have made recommendations based on these links (e.g., CBMS, 2016; Freeman et al., 2014; Seymour & Hunter, 2019). While these markers provide significant contributions for understanding the impact of various teaching practices, the students' perceptions of the practices can also lead to additional important insights. Students' perspectives can provide connections between their experiences and the teaching that some objective measure (or the instructor's report) of such practices may not detect. For instance, there is evidence that students in the same class report experiencing different teaching practices at different frequencies (including whole class discussion), and how often students reported experiencing these practices was related to their decision to stay in STEM (Borda et al., 2020; Ellis et al., 2014).

Consequently, it is important to further examine students' perceptions of classroom characteristics in which the teacher implements research-based teaching practices, such as those that increase interaction, since it can influence instructors' decisions to implement them. That is, it is particularly important to better understand students' perceptions of these characteristics because we see teaching practices as largely shaping the characteristics of the classroom. A teacher can implement particular teaching practices (e.g., request students to share their

first attempts at a task with a partner) to promote certain classroom characteristics (e.g., students talk with each other about their mathematical thinking during class). We see a need to further investigate students' perceptions of classroom characteristics (e.g., did students find it helpful to talk with other students about their mathematical thinking) to understand, and better support, teachers' instructional choices.

In what follows, we situate our work with the few relevant studies that we identified as investigating students' perceptions of interactive classroom characteristics (or the teaching that promotes such characteristics). While the research literature has begun to debunk the perspective that students will uniformly react poorly to research-based teaching practices (Andrews et al., 2020), there is some literature that provides grounding for this belief. For instance, Deslauriers et al. (2019) investigated undergraduate physics courses by comparing students' learning and their perceptions of their own learning. They randomly assigned approximately 150 students to two groups: one group attended two consecutive class meetings taught by passive instruction (e.g., lectures) and one group attended two consecutive class meetings that implemented active learning methods such as practices that increased interactions. They summarized their findings as:

> *Compared with students in traditional lectures, students in active classes perceived that they learned less, while in reality they learned more. Students rated the quality of instruction in passive lectures more highly, and they expressed a preference to have 'all of their physics classes taught this way', even though their scores on independent tests of learning were lower than those in actively taught classrooms (p. 1).*

These findings reflect similar sentiments from STEM students in other studies that have shown that students value lecture since they believe it better prepares them for exams (Borda et al., 2020) and view group work activities as having limited value and being unnecessarily difficult (Kressler & Kressler, 2020; Shekhar et al., 2020). Additionally, Bookman and Friedman's (1998) earlier work in the United States calculus reform movement of the 1990s documented that students (initially) reported disliking reformed teaching that included increased interactions via cooperative learning and group projects. At the end of the first term, many students reported on end-of-term course evaluations that "the course taught [them] very little" (p. 118) although after one- or two-years' time, "they reluctantly appreciated the course" and what they had learned therein (Bookman & Friedman, 1998, p. 121).

The literature also points to potential reasons why students might react negatively to interactive classrooms. Bookman and Friedman (1998) noted that aspects of reform-oriented calculus courses appeared to "violate students' deeply held beliefs about what mathematics is" (p. 121). Deslauriers et al. (2019) also suggested a mix of factors contributing to their results summarized above, including students' relative metacognitive naivete and not recognizing cognitive struggle as an aspect of learning. Relatedly, Sonnert et al. (2015) found that what they called "ambitious teaching"—or teaching "associated with pedagogical reform and novel approaches that aim at increasing the interactivity of the classroom experience and its relevance" (2015, p. 19)—was negatively related to college Calculus I students' reported views of their mathematical confidence, interest, and enjoyment. The authors pointed to potential reasons underlying these reports, including that students may not share their instructors' understandings of what is "good for them" (Sonnert et al., 2015, p. 385) or what might be helpful for their learning.

Taken together, these findings suggest that while students in their first year of university mathematics might initially react negatively to interactive classrooms and prefer more passive experiences, ultimately they can buy-in and see the value for their learning. However, this conjecture is based on an assortment of studies in different contexts, and the existing literature does not provide a comprehensive view of undergraduate students' perspectives of specific interactive classroom characteristics. There is a need to extend this work and test this conjecture in order to understand the larger impact on students' perceptions of a variety of classroom characteristics, particularly those that increase student interaction. In this study, we provide a larger-scale examination of about 5000 introductory undergraduate mathematics students' perceptions of the helpfulness of various classroom characteristics that require interaction for their learning. We compared students' reports of the helpfulness of these characteristics to others, and investigated how this might be accounted for by the extent to which students experience such characteristics. This is an important contribution since it has implications for understanding and developing student buy-in to teaching practices that are shown to support student success, and ultimately how to alleviate instructors' concerns about students' reactions when deciding whether to implement these practices in their classrooms. Further, it extends previous work that analyzed differences of students' perceptions of what happened in college Calculus I classes by (1) examining precalculus, Calculus I, and Calculus II students' perceptions of instruction and (2) their perceptions of how helpful classroom characteristics was for their learning.

Table 1 Classroom characteristic statements, categorized into Interactive Classroom Characteristics and Other Classroom Characteristics

	Item code	Student survey item
Interactive classroom characteristics	Peer support	Class is structured to encourage peer-to-peer support among students (e.g., ask a peer before you ask the instructor, having group roles, developing a group solution to share, etc.)
	Student talk	I talk with other students about course topics during class
	Criticize ideas	I constructively criticize other student's ideas during class
	Group work	I work with other students in small groups during class
	Immediate feedback	I receive immediate feedback on my work during class (e.g., student response systems such as clickers or voting systems; short quizzes; etc.)
	Questions	I am asked to respond to questions during class time
Other classroom characteristics	Instructor knows name	The instructor knows my name
	Connections	The class activities connect course content to my life and future work
	Wide participation	My instructor uses strategies to encourage participation from a wide range of students
	Individual work	I work on problems individually during class time
	Assignment feedback	I receive feedback from my instructor on homework, exams, quizzes, etc.
	Lecture	I listen as the instructor guides me through major topics

In particular, we investigate: To what extent do students regard various interactive classroom characteristics as helpful for their learning? How does the perceived helpfulness of these characteristics compare to the perceived helpfulness of other characteristics? How does this vary based on the extent to which those characteristics are present in introductory undergraduate mathematics courses?

Methods

Data for this analysis were gathered as part of Phase 2 of the Progress through Calculus study of university introductory undergraduate mathematics courses (i.e., Precalculus, Calculus 1, Calculus 2) and programs in the United States (funded by the National Science Foundation #1430540). Phase 1 of Progress through Calculus was a census survey that investigated how introductory mathematics courses are taught at universities across the United States, including aspects of the surrounding program and department. The results of Phase 1 indicated that the majority of introductory mathematics courses are taught primarily in a lecture format, and that they continue to function as gatekeepers (Apkarian & Kirin, 2017; Kirin et al., 2017; Rasmussen et al., 2019). Data from the census survey in Phase 1 informed Phase 2, which consisted of in-depth case studies of twelve university introductory mathematics programs which included surveys and focus groups, among other data collection methods, over a period of 2 years (2017–18 and 2018–19 academic years). These 12 sites were selected to represent a range of programs, both in terms of institutional characteristics (e.g., size of undergraduate population)

and program aspects (e.g., presence or absence of active learning, student success rates). For this study, we use a sequential explanatory design (Ivankova et al., 2006) to investigate a subset of the data collected from Phase 2 of the project. In particular, we draw on quantitative data (i.e., student survey responses to Likert-scale items) and then qualitative data (i.e., student focus group transcripts) to further interpret our findings.

Collection and analysis of student survey data

We analyzed student responses from surveys administered to all introductory mathematics students at the case study sites roughly 70% of the way through the Fall 2017 semester. The survey included 12 statements about characteristics of their class (e.g., "I am asked to respond to questions during class time") (see Table 1), which were adapted for students from Walter et al.'s (2016) Postsecondary Instructional Practice Survey (see Apkarian et al., 2019 for the full survey instrument and details of its development). With each statement, the students were asked to use a 5-point Likert scale to rate how descriptive the statement was of their specific class (Descriptiveness Item; 5 = very descriptive, 4 = mostly descriptive, 3 = somewhat descriptive, 2 = minimally descriptive, 1 = does not occur). For instance, we take a student's response who selected "minimally descriptive" for the Peer Support item (see Table 1) to mean that the student viewed Peer Support as minimally descriptive of their classroom experience. The students were then prompted to indicate whether each statement that they reported experiencing in their class was helpful for their learning (Helpfulness Item). That is, for each item in which

Vroom *et al. International Journal of STEM Education* (2022) 9:38

Page 6 of 17

a student responded with 2 or higher on the descriptiveness scale, they were asked to report how much that aspect of the course helped their learning (3 = very helpful, 2 = somewhat helpful, and 1 = not helpful). The project team also developed an instructor survey that featured parallel Descriptiveness Items. See Apkarian et al. (2019) for more details about the full survey suite, and Street et al. (2021) for a compilation of descriptive statistics of responses to the survey suite.

To gain a better understanding of students' perceptions of the specific classroom characteristics, we sorted the statements that were descriptive of interactive classrooms from those that did not. We categorized statements as describing an interactive classroom when it required (a) more than one student exchanging ideas or questions with each other during class time (i.e., student–student interaction), or (b) a student and a teacher exchanging ideas or questions with one another during class time (i.e., student–instructor interaction). The first four statements listed under Interactive Classroom Characteristics in Table 1 satisfied the student–student interaction requirement and the last two statements under Interactive Classroom Characteristics satisfied the student–instructor interaction requirement. Table 1 also includes six survey items that did not satisfy either criterion. We included these "Other Classroom Characteristics" in our analysis because we interpreted some statements as related to classrooms that required student interactions. For example, we view increased Lecture as likely decreasing interactions. Additionally, we conjectured that comparing responses of the Interactive Classroom Characteristics to Other Classroom Characteristics would enable richer interpretations of the data by allowing for contrasting cases. This enabled us to better understand whether students' perceptions could be a product of the classroom characteristic rather than some other factor (e.g., the prevalence of a classroom characteristic).

While we were primarily focused on the students' perceptions of classroom characterizations, we also evaluated the consistency of the students' responses on the student survey to their instructor responses on the instructor survey. Following Ellis et al. (2014), we first reduced the data set to responses from courses with an instructor response and at least five student responses. These restrictions resulted in a total of 4904 student responses from 171 introductory mathematics classes (1864 student responses from 57 Precalculus classes, 1807 student responses from 74 Calculus 1 classes, and 1233 student responses from 40 Calculus 2 classes). Then, we conducted paired samples *t*-tests to compare the mean scores for each of the 12 student–instructor Descriptiveness Items (see "Appendix"). We found that the difference in student and instructor ratings were

significant for all items except Peer Support. The significance was negligible for all but three items because the effect size was small[1] (Cohen's $d \leq 0.23$). The three items for which the differences were not negligible were the following items: Questions ($d = 0.73$), Group Work ($d = 0.35$), and Lecture ($d = 0.36$). The large difference for Questions could be explained by wording differences of the paired items. The instructor version read, "I ask students to respond to questions during class time", while the student version read, "I am asked to respond to questions during class time." The instructor average is likely to be greater than the student average since (a) instructors can pose questions without directing them to all of their students and (b) an individual student may not interpret a question being directed to them when an instructor poses a question to the whole class. Overall, these findings are consistent with Ellis et al.'s (2014) findings that Calculus I instructors tended to report lower use of traditional practices (such as lecture) and report higher use of innovative practices (such as group work) when compared to their students. These results allow us to treat students' Descriptiveness Items reports as a reliable measure for how descriptive the corresponding characteristics were of the classes, noting that the instructor's reports of Questions and Group Work are likely to be slightly higher than their students and the instructor's reports of Lecture slightly lower than their students. We present these results as part of the overall methodology in order to provide the reader with additional context about the classroom environment, as well as to suggest that instructors' self-reflection on their practice can be used (in most cases) as a reasonable proxy for classroom activities as experienced by students.

In the Results section, we present descriptive statistics for the student Descriptiveness Items and Helpfulness Items. In addition, we used 2×2 contingency Chi-square to compare students' perceptions of the helpfulness of classroom characteristics, grouping students based on their reports of how descriptive the characteristic was of their course. We report the significance level as well as the effect size because it is more indicative of differences given our large sample size.

Collection and analysis of data from student focus groups

To gain insight into how students from introductory mathematics courses might interpret the Helpfulness items, we conducted student focus groups at four of the case study sites during Spring 2019. In total, there were

---

[1] Here and throughout the report, we interpret the effect sizes based on Cohen's (1992) recommendations and use d = 0.2/φ = 0.1, d = 0.5/φ = 0.3, and d = 0.8/φ = 0.5 as guidelines for small, medium, and large effect sizes, respectively.

16 focus groups with 124 students that volunteered from 11 different introductory mathematics courses. The participating students were recruited from the same population of those who were surveyed, in that they were from the same institutions and same surveyed courses taught by the same pool of instructors, but were not necessarily survey participants. The semi-structured focus groups were audio recorded and lasted approximately one hour each. The focus groups served several purposes, one of which was to investigate introductory mathematics students' interpretations of the Helpfulness Items. For this study, we analyzed the focus group discussions about aspects of the course that the students considered helpful for their learning. This discussion was initiated by a question that focused students' attention on practices that supported their learning (e.g., what things happen in your class that are most helpful for your learning?). Then, the interviewer prompted students to describe what they had in mind when considering a teaching practice to be helpful for their learning (e.g., what does it mean for it to be helpful for your learning?, how do you know if something is helpful for your learning?). This portion of the discussion typically lasted no more than 10 min.

The first and second authors used thematic analysis (Braun & Clarke, 2006) to identify indicators that an aspect of the course was helpful for the students' learning from their perspective. We began by independently open coding the focus group transcripts, tagging segments that illuminated student interpretations of learning and then taking notes of the expressed ideas. The second author generated closed form codes and descriptions of the codes based on the notes. We then returned to the transcript data and independently tagged the data with the codes, allowing excerpts to be double coded when multiple ideas were represented. We met to discuss our independent coding and, in order to reach consensus on coding decisions, we discussed the interpretations of the codes and resolved any disagreements. Then, we refined the definitions of codes based on interpretation differences and grouped the codes into larger themes. Four themes emerged from this process: students considered aspects of the course as helpful for their learning when they (1) could understand how to solve assigned problems and understand why the procedures work; (2) earn good grades; (3) build on their knowledge or apply it in different contexts, and (4) teach others. We describe these themes in more detail in the results section.

Results and discussion

In what follows, we answer our research questions by first investigating the extent to which students reported the specific classroom characteristics listed on the survey in their class. In the second section, we investigate characteristics that students deemed helpful for their learning, both in the aggregate as well as grouped by how often students reported experiencing such characteristics. In order to add nuance to the quantitative findings, we then present our findings regarding what students likely had in mind when indicating that a characteristic was helpful (or not) for their learning. Throughout the report, we will refer to the students who rated the items mostly or very descriptive of their class as students who *often experienced* the particular classroom characteristic and students who indicated that the items were minimally or somewhat descriptive of their class as students who *rarely experienced* the characteristic.

Interactive classroom characteristics are less prevalent than other classroom characteristics

We first consider the student responses to the Descriptiveness Items on the student survey to understand the extent to which Interactive Classroom Characteristics and Other Classroom Characteristics were present in the classes. In what follows, we present descriptive statistics of the students' responses which indicates that students did not report experiencing Interactive Classroom Characteristics very often.

Figures 1 and 2 depict the percent of students that reported each item as 'does not occur', 'minimally descriptive', 'somewhat descriptive', 'mostly descriptive' or 'very descriptive'. For the Interactive Classroom Characteristics (Fig. 1), less than half of the students reported that they often experienced each of the six Interactive Classroom Characteristics: Student Talk (39.06%), Peer Support (36.96%), Questions (34.32%), Group Work (29.56%), Immediate Feedback (29.32%), Criticize Ideas (13.52%). Additionally, a relatively large percentage of students reported that the Interactive Classroom Characteristics did not occur at all in their classes [Criticize Ideas (47.78%), Immediate Feedback (33.73%), Group Work (31.44%), Questions (19.37%), Peer Support (18.74%), Student Talk (17.75%)]. Most of the students in our study did not report experiencing the Interactive Classroom Characteristics very often, if at all.

Figure 2 offers more insight regarding what goes on in the introductory mathematics courses. More than half of the students indicated they often experienced the following characteristics: Lecture (83.76%), Instructor Knows Name (55.56%) and Assignment Feedback (54.82%). Moreover, the distribution of the Lecture responses is distinct from the Other Classroom Characteristics as well as the Interactive Classroom Characteristics in that the distribution appears to be highly left-skewed (the skewness value was equal to -1.34), with only 1.37% of students said that Lecture does not occur and an additional 14.86% reported that they rarely experienced it.
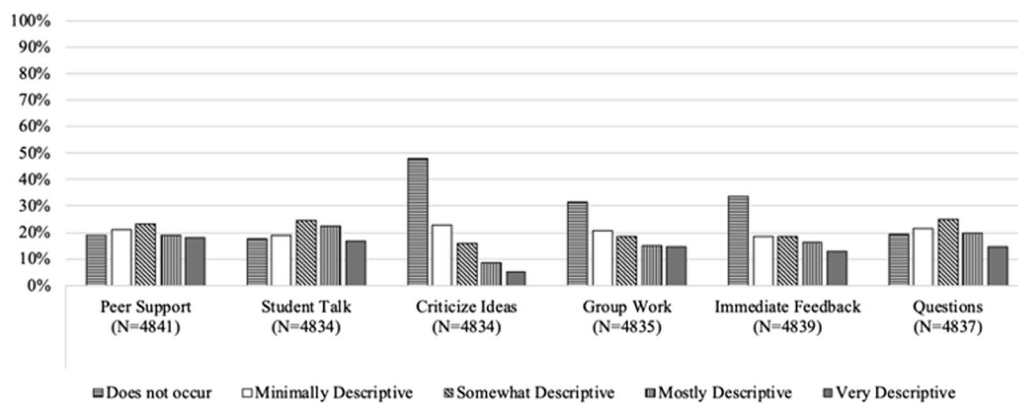
Fig. 1 Summary of descriptiveness of Interactive Classroom Characteristics
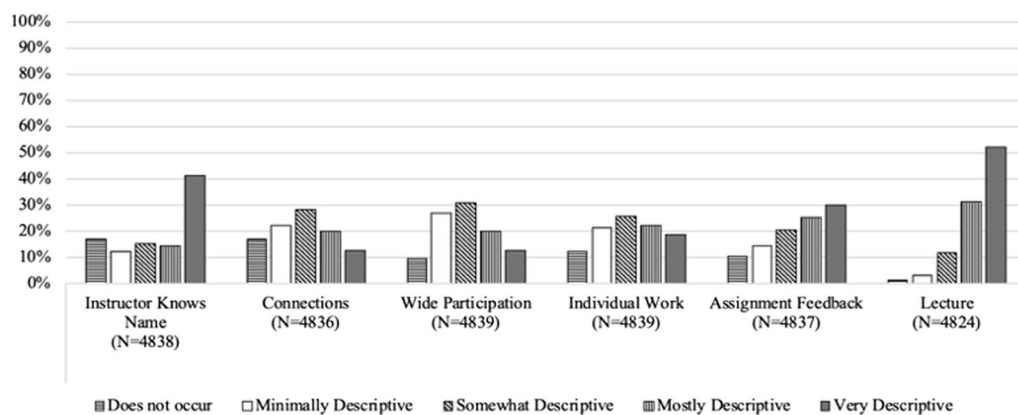

Fig. 2 Summary of descriptiveness of Other Classroom Characteristics

This suggests that despite the mounting evidence that lecture is ineffective compared to more active experiences it is still prevalent in many introductory mathematics courses.

Student reports of the descriptiveness of Assignment Feedback is slightly left-skewed (the skewness value was equal to − 0.46); however, nearly a tenth of the students (10.52%) said that it did not occur and an additional 34.65% reported it rarely occurring. This coupled with the finding that a large percent of students indicated that Immediate Feedback did not occur, suggests that a considerable number of students are rarely receiving feedback if at all.

Students regard classroom characteristics that they reported experiencing often as helpful for their learning
Recall that when students indicated that the items were at all descriptive of their class (minimally, somewhat, mostly, or very), they were prompted to rate how helpful the characteristic was for their learning. In what

follows, we present the descriptive summaries of the student responses to these items. First, we present a summary of the Helpfulness Item and then we compare the students' reports after grouping students based on their reports of how descriptive the characteristic was of their course.

Figures 3 and 4 depict the percent of students that rated the Interactive Classroom Characteristics and Other Classroom Characteristics, respectively, as not helpful, somewhat, or very helpful. For instance, of the students who said that their class was structured to encourage peer-to-peer support among students, 3,907 reported how helpful they thought it was for their learning and 14.64% of the 3,907 said it was not helpful for their learning, 45.45% said it was somewhat helpful, and 39.90% said it was very helpful. We found that more than half of the students that responded to the items said that Lecture (73.53%), Assignment Feedback (64.27%), Immediate Feedback (51.57%), and Instructor Knows Name (51.07%) were very helpful for their learning.
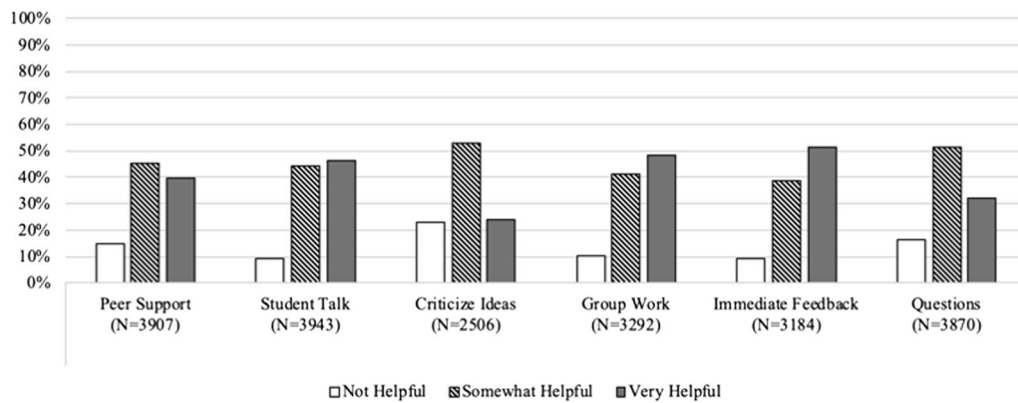
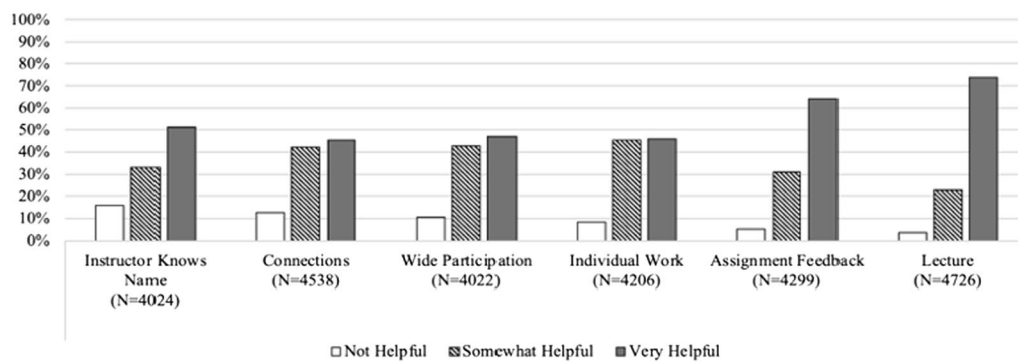**Fig. 3** Summary of helpfulness of Interactive Classroom Characteristics

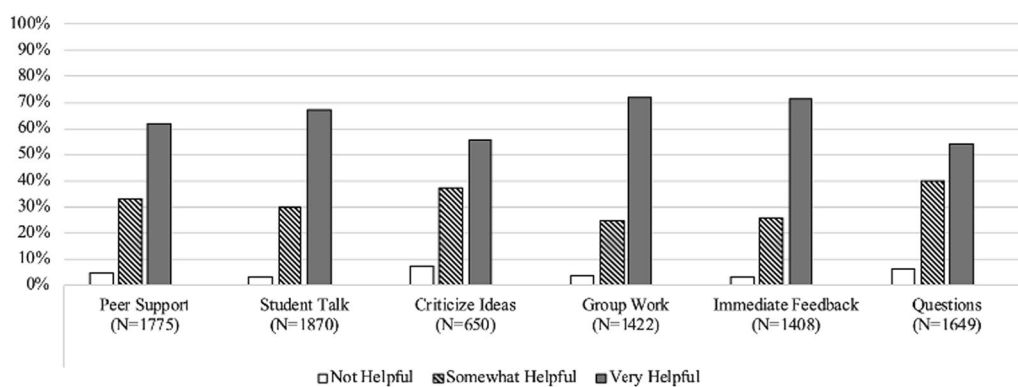**Fig. 4** Summary of helpfulness of Other Classroom Characteristics

Fig. 5 Summary of helpfulness of very or mostly descriptive Interactive Classroom Characteristics

Next, we considered the helpfulness ratings from students who reported often experiencing the classroom characteristics (Figs. 5 and 7) as well as from students who reported rarely experiencing classroom characteristics (Figs. 6 and 8).

From Fig. 5, notice that students who often experienced the Interactive Classroom Characteristics tended to deem them as helpful for their learning. In fact, more than half of these students marked each of the Interactive Classroom Characteristics as very helpful for their
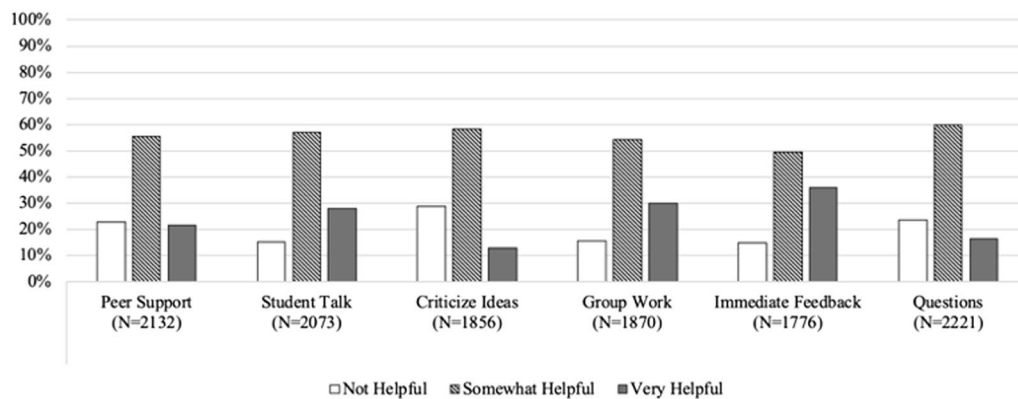
Fig. 6 Summary of helpfulness of minimally or somewhat descriptive Interactive Classroom Characteristics
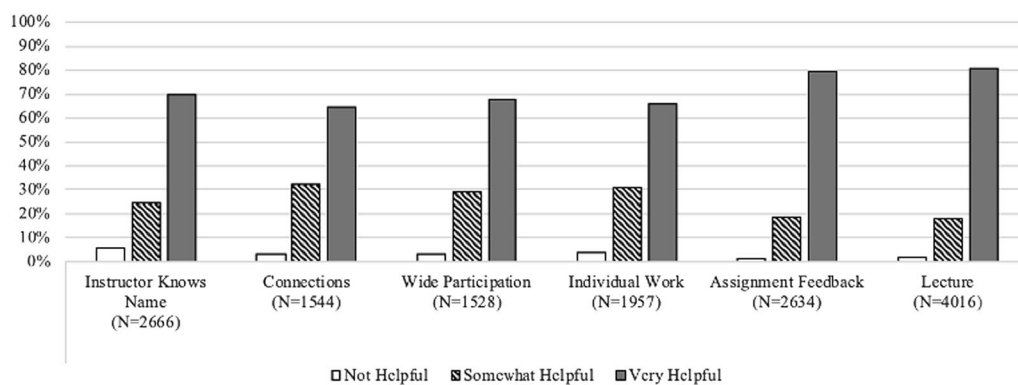

Fig. 7 Summary of helpfulness of very or mostly descriptive Other Classroom Characteristics

learning and more than 90% of the students said that the characteristics were very or mostly helpful for their learning. This suggests that students think Interactive Classroom Characteristics are helpful for their learning when they experience them often.

However, we did not observe the same pattern when we considered the helpfulness ratings from students who rarely experienced the Interactive Classroom Characteristics (Fig. 6). These students were most likely to say that the Interactive Classroom Characteristics were somewhat helpful for their learning, and a considerable proportion of students reported that they were not helpful for their learning. It is also interesting to note the difference in the distributions of the graphs given in Fig. 6. For example, consider the responses for Student Talk and Criticize Ideas. Most of the students who rarely experienced Student Talk reported that it was somewhat or very helpful for their learning. Whereas most students who rarely experienced Criticize Ideas claimed that it was somewhat or not helpful for their learning. Similarly, students who rarely experienced Immediate Feedback and

Group Work were most likely to say the characteristics were somewhat or very helpful while students who rarely experienced Peer Support and Questions were likely to report the characteristics as somewhat or not helpful.

We found a similar trend when we considered the helpfulness ratings for Other Classroom Characteristics. Students who reported often experiencing Other Classroom Characteristics tended to rate them as very helpful for their learning (see Fig. 7); whereas when students reported rarely experiencing Other Classroom Characteristics they were likely to claim that it was somewhat helpful for their learning (see Fig. 8). Students who reported rarely experiencing Wide Participation, Individual Work, Assignment Feedback, and Lecture were most likely to report the characteristics were somewhat or very helpful. Students who reported rarely experiencing Instructor Knows Name and Connections were likely to report the characteristics as somewhat or not helpful.

To further investigate the difference that we observed in students' helpfulness ratings, we conducted 2×2 Chi-square tests to determine whether there was a significant
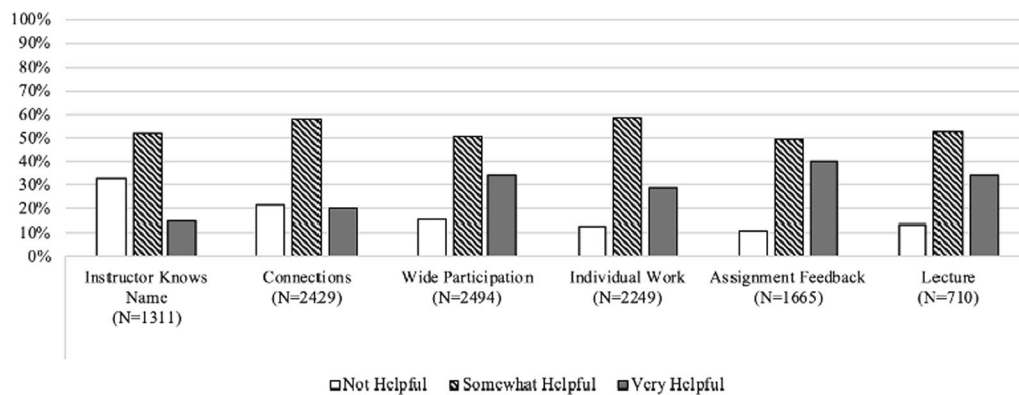
Fig. 8 Summary of helpfulness of minimally or somewhat descriptive Other Classroom Characteristics

Table 2 2×2 Chi-square tests

	Item code	P_{rarely}	P_{often}	$\chi^2(1)$	ϕ
Interactive classroom characteristics	Peer support*	77.11% (1644/2132)	95.27% (1691/1775)	255.52	0.26
	Student talk*	84.95% (1761/2073)	96.90% (1812/1870)	165.08	0.20
	Immediate feedback*	85.30% (1515/1776)	97.02% (1366/1408)	125.13	0.20
	Criticize ideas*	71.28% (1321/1853)	92.77% (603/650)	124.86	0.22
	Group work*	84.33% (1577/1870)	96.48% (1372/1422)	127.80	0.20
	Questions*	76.23% (1693/2221)	93.75% (1546/1649)	213.03	0.23
Other classroom characteristics	Instructor knows name*	66.90% (877/1311)	94.37% (2516/2666)	529.66	0.36
	Connections*	78.30% (1902/2429)	96.82% (1495/1544)	261.26	0.26
	Wide participation*	84.64% (2111/2494)	96.99% (1482/1528)	151.57	0.19
	Individual work*	87.06% (1958/2249)	96.47% (1888/1957)	118.47	0.17
	Assignment feedback*	89.40% (1488/1665)	98.48% (2594/2634)	176.72	0.20
	Lecture*	86.76% (616/710)	98.36% (3950/4016)	248.03	0.23

*Indicates that the difference between the proportions were significantly different using $p < 0.0001$

difference between students who reported often experiencing the different classroom characteristics and students who reported rarely experiencing the characteristics in their perception of how helpful the characteristic was for their learning (grouped by not helpful and very/somewhat helpful). Table 2 offers the percent of students who reported rarely experiencing each characteristic that said it was helpful for their learning (P_{rarely}), the percent of students who reported often experiencing each characteristic that said it was helpful for their learning (P_{often}), Pearson Chi-square value [$\chi^2(1)$], and the effect size (ϕ).

Notice from Table 2 that the differences in proportions were all significant, and the phi coefficients suggested that the differences were medium for most of the characteristics. For instance, 95.27% of students who reported often experiencing Peer Support reported it was helpful for their learning while only 77.11% of students who reported rarely experiencing it reported it as helpful for their learning. This

difference was significant, $\chi^2(1) = 255.52, p < 0.0001$, and the phi coefficient, $\phi = 0.26$, suggested a medium effect. It is interesting to note that the largest difference in the student reports on the items was for Instructor Knows Name ($\phi = 0.36$, medium effect). In particular, 94.37% of students who reported often experiencing their instructor knowing their name reported it as helpful for their learning while only 66.90% of students who reported rarely experiencing it reported it as helpful for their learning. This might be attributed to additional factors; for instance, it is possible that instructors who often refer to students by name foster a sense of community within the class which, from the students' perspective, supports their learning. The differences between students who reported often experiencing the different characteristics and students who reported rarely experiencing the characteristics in their perception of how helpful the characteristic was for their learning suggests that students who experience

Table 3 Frequency of focus groups that discussed themes

Theme	Number of focus groups that discussed theme (*N* = 16)
Solve assigned problems and understand why the procedures work	15
Earn good grades	8
Build on their knowledge or apply it in different contexts	4
Teach others	3

classroom characteristics (interactive or otherwise) more often are more likely to find them helpful for their learning.

Students considered several indicators when identifying if a classroom characteristic was helpful for their learning

The qualitative analysis of the focus group interviews gave insight into how students might have interpreted the Helpfulness survey items, adding context to our quantitative findings. This section includes a description of the four themes that focus group participants expressed as indicators that an aspect of a course was helpful for their learning. Specifically, students suggested that they can identify a characteristic as helpful for their learning when they are able to (1) solve assigned problems and understand why the procedures work; (2) earn good grades; (3) build on their knowledge or apply it in different contexts, and (4) teach others. In Table 3, we offer how many of the 16 focus groups held discussions related to each theme.

The first theme was the most common among our data, and arose as part of the student discussion in 15 of the 16 focus groups. In particular, students most commonly considered a classroom characteristic to be helpful for their learning when they were able to solve problems (familiar or novel) that were assigned for class, including homework questions or exam items, and were able to understand why the procedures worked. This theme included instances when students suggested that they were equipped to solve problems without needing to reference course resources (e.g., notes, textbook) and could self-assess or evaluate the correctness of their work. For instance, a student explained:

I think, for me, it's being able to quickly ... right when you're explained how to do it, it makes sense of how to do the math. You're not struggling 10 minutes after class, and you're like, I need to learn this because I didn't really quite get how the instructor taught it or because I had to use a textbook with no professor ... You have to figure it out yourself. So it's really nice when that's not happening. So I guess that

thing and then just being able to go on the test and you're like, I understand it so well that when I get to a problem that I forgot how to solve, I actually know that. I'm not walking away being like, I'm not sure what my test score is going to be. I know that I either solved it correctly, or I'm iffy on it, or I didn't [solve it correctly], and just being able to be accurate in that.

This student viewed solving homework problems independently without needing resources, such as the instructor or textbook, as an indicator of their learning in addition to solving problems on a test and being able to assess accuracy.

Students often talked about solving problems and understanding why as indicators that they were learning the material. One student said, "Just having things click and make sense, and ... you can work through a problem and kind of know sort of why that works", suggesting that understanding how and why procedures worked aided them in solving the problem. Several students also discussed understanding or seeing the "bigger picture", saying things like: "I know it's helpful when I get the idea behind the answer rather than just getting the answer—because I know what I don't understand."

Many students expressed that they knew classroom characteristics were helpful for their learning when they earned good grades on an exam or in the course. This idea surfaced in eight of the focus groups, with students saying that they knew something was helpful for their learning when they received "an A on the test" or when they were "not worried … that [they were] not gonna pass the test." Some students emphasized that both understanding material *and* doing well on exams were part of knowing if something was helpful for their learning. One student explained, "I think you can gauge if it was helpful or not if you understand the material enough to do okay on the test." However, there were three instances (from three different focus groups) when students acknowledged that getting a grade did not always align with their understanding of the material. For example, one student stated:

I wish being successful in the class meant understanding the material, but it seems like there's more of an emphasis on— you're successful in this if you can make it through all these homework assignments and get a good grade on the test and that means you have a good GPA. And like in reality those are all just numbers, but they play such a big role—... it's on your resume... and that decides your future.

This quote points to the tension or disconnect students might feel between their understanding and what is indicated by a grade. Further, because of this disconnect,

students might report that a classroom characteristic is helpful for their learning when it supports them in getting a good grade, even though it might not be helpful for developing their understanding of mathematics (which we view as a critical component of learning).

Other students communicated that something was helpful for their learning when they were able to use or build on their knowledge. This theme arose in four focus groups, and coded segments captured instances of students explaining that they could build on their understanding to learn more within the same class as well as instances of students explaining that they were able to recognize and apply their understandings across different contexts (e.g., Chemistry class, everyday situations). For example, one student considered their experience learning new material within their mathematics class, saying, "Being able to add new material onto stuff we have already learned, and it be easier than the first time, then I think you're actually learning." Another student reflected on being able to apply what they were learning in precalculus to their economics (econ) class, saying:

> *I find it helpful when... I'm able to apply it on my own in another class... Even if they don't explain it. In my Econ class we do a lot with graphs. A lot of times I am like, oh that's a constant rate of change and all that stuff, and this is what this means and this is why we have this formula. Because in Econ they will just give you formulas, but now I am able to apply it to this class so now I know it... I one hundred percent understand this.*

Finally, we identified that some students expressed being able to teach someone else the material as something helpful for their own learning. For example, this theme was evidenced by students saying things like, "I get it now, I actually understand it and I could explain it to somebody else." Although such instances were less common overall (occurring in only three of the 16 focus groups), we found this to be a useful theme to shed light on how students might have interpreted when something was helpful for their learning.

Conclusions

The key finding from our work is that students are likely to view classroom characteristics that they experience more often as more helpful for their learning and are less likely to view characteristics that they rarely experience as helpful for their learning. Students view the characteristics that they regularly experience as helping them to solve problems and understand why the procedures work, earn good grades, build on their knowledge or apply it in different contexts, and teach others. In what

follows, we highlight two important interpretations and implications of this key finding.

First, there is a feedback loop regarding students' views of a teaching approach and how often they experience it. Consistent with national postsecondary STEM teaching trends (e.g., Stains et al., 2018), students in our study reported experiencing lectures more often than any of the other characteristics that were included in the survey. Most of the students in our sample (84%) indicated that they often listened as their instructor guided them through major topics and nearly all of these students (98%) reported this to be somewhat or very helpful for their learning. The more a student experiences a lecture, the more they expect that lecture is how they are supposed to be taught and how they will learn best, despite research indicating otherwise (Ellis et al., 2014; Freeman et al., 2014; Kogan & Laursen, 2014; Murphy et al., 2016; Petrillo, 2016). This feedback loop legitimizes instructors' hesitancy to implement teaching practices that foster interactive classrooms because they fear that students may react negatively to these practices because they are counter to students' expectations (Bookman & Friedman, 1998; Dancy & Henderson, 2012; Finelli et al., 2014; Froyd et al., 2013; Hayward et al., 2016; Shadle et al., 2017).

Importantly, this feedback loop goes in the other direction as well. In our study, students who regularly experienced more interactive classrooms rated these characteristics as more helpful for their learning. That is, they are likely to think that the characteristics will help them to solve problems and understand why the procedures work, earn good grades, build on their knowledge or apply it in different contexts, and teach others. However, our study also shows that students may not consistently have this view of interactive classrooms on their learning when they are not consistently interacting with others in class; sprinkling in some interaction can be perceived as less helpful than courses that consistently include interactions during class.

Since teaching practices that foster interactions are shown to influence student success in a STEM major (e.g., Seymour & Hewitt, 1997), one implication for practice and policy is to support instructors in *consistently* implementing practices that increase student interactions. This support can come in the form of professional development aimed at supporting instructors to overcome barriers while implementing interactive approaches, shifting a department or university culture to expect all courses to include more interactive approaches, and not (solely) relying on Students Evaluations of Teaching (Kreitzer & Sweet-Cushman, 2022) to assess the quality of instruction.

Research can aid these supports in numerous ways. For example, Tharayil et al. (2018) and Nguyen et al. (2021)

have studied STEM instructors' strategies to mitigate students' resistance to interactive teaching practices. This body of research points to specific ways instructors can use explanation and facilitation strategies to improve students' perception of interactive classroom experiences. Future research is needed to connect these strategies to classes where students experience varying levels of interactive characteristics. An additional area for research is to unpack the process of departmental and university culture change and how it is related to teaching expectations that include more interactive approaches. Future work might also explore shifts in student reports of the helpfulness of classroom characteristics if they take a class that is primarily lecture-oriented (say for precalculus) followed by a class that includes more interactions (for Calculus 1), and vice versa. Future work could also investigate students' mathematical confidence, interest, and enjoyment with respect to how often they experience interactive classroom characteristics, both within one specific class and across their undergraduate experience. Such studies could provide a more nuanced understanding to Sonnert and Sadler's (2015) and Sonnert et al.'s (2015) finding that research-based teaching practices like those that increase interactions were negatively related to students' reports of their attitude towards mathematics. Further, such studies could illuminate other ways to assess teaching systematically beyond Student Evaluations of Teaching.

Our second interpretation of our key finding is that assessment is a notable factor in students' perceptions of classroom characteristics. Students frequently discussed earning good grades and being able to solve problems as a way to know that a classroom characteristic was helpful for their learning. We note that if students are assessed on homework and exams to simply solve routine problems, then it is understandable that lectures presenting clear steps to solving procedural problems are viewed as most helpful for learning because such exposure may seemingly support them in earning a good grade. Recall that some students also highlighted that they felt there was a disconnect between the grade they received on an assessment and their understanding of the material, indicating that the assessment may not have been designed to assess understanding beyond working through procedures. Tallman and colleagues (2016) found that the vast majority of college Calculus I final exams assessed highly procedural and skill-based ideas rather than conceptual problems that require students to explain their thinking. A clear implication for practice is to align assessment strategies with instructional approaches. We believe this should start with the instructor deciding what they want students to be able to accomplish

coming out of their class, and aligning assessment approaches with these goals, and then instructional approaches with these assessments. This approach aligns with Wiggins and McTighe's (2005) backwards design principles which have been successful in a variety of contexts.

For teachers who want their students to memorize procedures and skills, traditional assessments and lectures focused on how to do these procedures appear successful. But for teachers who want their students to learn problem-solving skills, and how to apply key ideas in novel settings, can be better supported through interactive instructional approaches paired with assessments tied to these goals. If instructors want to move towards more interactive classrooms, their assessment approaches need to change as well. However, many introductory STEM courses are taught in a coordinated way with assessments that are consistent across sections (Apkarian & Kirin, 2017; Martinez et al., 2021; Rasmussen & Ellis, 2015), and so individual instructors may not be able to change the assessments although they can change their instructional approaches. Thus, an implication for policy is for departments to support more innovative assessment approaches (e.g., oral assessments, see Nelson, 2010; specifications grading, see Nilson, 2014) tied to more interactive instructional approaches on a broader scale, including in coordinated sections. Lastly, this main finding points to a direction for future research to explore the connection between student perceptions of various classroom characteristics in relation to the assessment approaches of those classes. We conjecture that assessment changes are a critical way to support a shift in what students deem as helpful for their learning.

Our study not only contributes to a richer understanding of the student experience in introductory mathematics classes, but it also sheds light on the development of student buy-in for teaching that increases student interactions. Future work should investigate patterns of student experience in relation to various intersecting social identity markers (such as race, gender, and sexual orientation), including examining how these markers interact with the degree to which students found certain characteristics helpful. We view this work as an important step toward removing barriers to the implementation of research-based teaching practices and supporting a better experience for students and instructors related to research-based teaching.

Appendix
Paired samples *t*-tests for each of the student–instructor descriptiveness items

	Item code	N	Instructor mean/ standard deviation	Student mean/ standard deviation	Cohen's *d*
Interactive classroom characteristics	Peer support	4847	2.97/1.35	2.96/1.37	0.00
	Student talk*	4843	3.17/1.31	3.01/1.34	0.09
	Immediate feedback*	4847	2.88/1.48	2.56/1.43	0.17
	Criticize ideas*	4843	2.28/1.25	1.99/1.19	0.18
	Group work*	4843	3.17/1.41	2.60/1.43	0.35
	Questions*	4845	4.15/1.30	2.89/1.33	0.73
Other classroom characteristics	Instructor knows name*	4844	3.58/1.44	3.51/1.53	0.05
	Connections*	4845	2.74/1.20	2.88/1.26	0.08
	Wide participation*	4845	3.43/1.17	3.25/1.29	0.12
	Individual work*	4847	2.80/1.30	3.14/1.29	0.21
	Assignment feedback*	4844	3.85/1.24	3.49/1.33	0.23
	Lecture*	4849	3.81/1.09	4.31/0.88	0.36

*Indicates that the difference between the instructor and student reports were significantly different using $p < 0.01$

Author contributions
KV, JG, NA, and JEH contributed to the survey instrument development and all authors contributed to data collection. The quantitative analysis was led by KV, and the qualitative analysis was led by KV and JG. All authors contributed to the interpretation of the data analysis with KV leading the efforts to identify the main results. All authors contributed to the writing of the manuscript, and approved the final manuscript. All authors read and approved the final manuscript.

Availability of data and materials
The datasets generated during and analyzed during the current study are not publicly available due to it containing identifying information. De-identifiable data may be available from the corresponding author on reasonable request.

Declarations

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

Author details
[1]Department of Mathematics, Oregon State University, Kidder Hall 368, Corvallis, OR 97331, USA. [2]Department of Mathematics, University of Texas at San Antonio, Ransom Rd, San Antonio, TX 78249, USA. [3]School of Mathematical and Statistical Sciences, Arizona State University, P.O. Box 871804, Tempe, AZ 85287, USA. [4]Portland State University, Portland, USA. [5]Colorado State University, Fort Collins, USA.

References
Andrews, M. E., Graham, M., Prince, M., Borrego, M., Finelli, C. J., & Husman, J. (2020). Student resistance to active learning: do instructors (mostly) get it wrong? *Australasian Journal of Engineering Education, 25*, 142.

Andrews, T. C., & Lemons, P. P. (2015). It's personal: Biology instructors prioritize personal over empirical evidence in teaching decisions. *CBE Life Sciences Education, 14*(1), 1–18. https://doi.org/10.1187/cbe.14-05-0084

Apkarian, N., Henderson, C., Stains, M., Raker, J., Johnson, E., & Dancy, M. (2021). What really impacts the use of active learning in undergraduate STEM education? Results from a national survey of chemistry, mathematics, and physics instructors. *PLoS ONE, 16*(2), 1–15.

Apkarian, N., & Kirin, D. (2017). *Progress through calculus: Census survey technical report.* Retrieved from Mathematical Association of America website: http://bit.ly/PtC_Reporting.

Apkarian, N., Smith, W. M., Vroom, K., Voigt, M., Gehrtz, J., PtC Project Team, & SEMINAL. Project Team. (2019). *X-PIPS-M Survey Suite.* Available: https://www.maa.org/sites/default/files/XPIPSM%20Summary%20Document.pdf.

Aragón, O. R., Eddy, S. L., & Graham, M. J. (2018). Faculty beliefs about intelligence are related to the adoption of active-learning practices. *CBE Life Sciences Education, 17*(3), 1–9. https://doi.org/10.1187/cbe.17-05-0084

Bathgate, M. E., Aragón, O. R., Cavanagh, A. J., Waterhouse, J. K., Frederick, J., & Graham, M. J. (2019). Perceived supports and evidence-based teaching in college STEM. *International Journal of STEM Education, 6*(1), 1–14.

Bookman, J., & Friedman, C. P. (1998). Student attitudes and calculus reform. *School Science and Mathematics, 98*(3), 117–122. https://doi.org/10.1111/j.1949-8594.1998.tb17404.x

Borda, E., Schumacher, E., Hanley, D., Geary, E., Warren, S., Ipsen, C., & Stredicke, L. (2020). Initial implementation of active learning strategies in large, lecture STEM courses: Lessons learned from a multi-institutional, interdisciplinary STEM faculty development program. *International Journal of STEM Education, 7*(1), 1–18. https://doi.org/10.1186/s40594-020-0203-2

Boring, A. (2017). Gender biases in student evaluations of teaching. *Journal of Public Economics, 145*, 27–41.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77–101.

CBMS. (2016). *Active learning in post-secondary mathematics education.* Conference Board of the Mathematical Sciences. http://www.cbmsweb.org/Statements/Active_Learning_Statement.pdf.

Chávez, K., & Mitchell, K. M. (2020). Exploring bias in student evaluations: Gender, race, and ethnicity. *PS Political Science & Politics, 53*(2), 270–274.

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112*(1), 155.

Dancy, M. H., & Henderson, C. (2012). Experiences of new faculty implementing research-based instructional strategies. In N. S. Rebello, P. V. Engelhardt, & C. Singh (Eds.), *AIP Conference Proceedings* (Vol. 1413, pp. 163–166). American Institute of Physics.

Deslauriers, L., McCarty, L. S., Miller, K., Callaghan, K., & Kestin, G. (2019). Measuring actual learning versus feeling of learning in response to being actively engaged in the classroom. *Proceedings of the National Academy of Sciences, 116*(39), 19251–19257. https://doi.org/10.1073/pnas.1821936116

Eddy, S. L., & Hogan, K. A. (2014). Getting under the hood: How and for whom does increasing course structure work? *CBE Life Sciences Education, 13*(3), 453–468.

Vroom *et al. International Journal of STEM Education*   *(2022) 9:38*

Page 16 of 17

Ellis, J., Kelton, M. L., & Rasmussen, C. (2014). Student perceptions of pedagogy and associated persistence in calculus. *ZDM Mathematics Education, 46*(4), 661–673.

Fan, Y., Shepherd, L. J., Slavich, E., Waters, D., Stone, M., Abel, R., & Johnston, E. L. (2019). Gender and cultural bias in student evaluations: Why representation matters. *PLoS ONE, 14*(2), 1–16.

Ferrare, J. J. (2019). A multi-institutional analysis of instructional beliefs and practices in Gateway Courses to the Sciences. *CBE Life Sciences Education, 18*(2), 1–16. https://doi.org/10.1187/cbe.17-12-0257

Finelli, C. J., Daly, S. R., & Richardson, K. M. (2014). Bridging the research-to-practice gap: Designing an institutional change plan using local evidence. *Journal of Engineering Education, 103*(2), 331–361.

Foote, K. T., Neumeyer, X., Henderson, C., Dancy, M. H., & Beichner, R. J. (2014). Diffusion of research-based instructional strategies: The case of SCALE-UP. *International Journal of STEM Education, 1*(1), 1–18. https://doi.org/10.1186/s40594-014-0010-8

Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., & Wenderoth, M. P. (2014). Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the National Academy of Sciences, 111*(23), 8410–8415.

Froyd, J. E., Borrego, M., Cutler, S., Henderson, C., & Prince, M. J. (2013). Estimates of use of research-based instructional strategies in core electrical or computer engineering courses. *IEEE Transactions on Education, 56*(4), 393–399. https://doi.org/10.1109/TE.2013.2244602

Hake, R. R. (1998). Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses. *American Journal of Physics, 66*(1), 64–74. https://doi.org/10.1119/1.18809

Hayward, C. N., Kogan, M., & Laursen, S. L. (2016). Facilitating instructor adoption of inquiry-based learning in college mathematics. *International Journal of Research in Undergraduate Mathematics Education, 2*(1), 59–82.

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), 1–14. https://doi.org/10.1103/PhysRevSTPER.3.020102

Henderson, C., Dancy, M., & Niewiadomska-Bugaj, M. (2012). Use of research-based instructional strategies in introductory physics: Where do faculty leave the innovation-decision process? *Physical Review Special Topics Physics Education Research, 8*(2), 1–15. https://doi.org/10.1103/PhysRevSTPER.8.020104

Ivankova, N. V., Creswell, J. W., & Stick, S. L. (2006). Using mixed-methods sequential explanatory design: From theory to practice. *Field Methods, 18*(1), 3–20.

Jacobs, V., & Spangler, D. (2017). Research on core practices in K-12 mathematics teaching. In J. Cai (Ed.), *Compendium for research in mathematics education* (pp. 766–792). National Council of Teachers of Mathematics.

Johnson, E., Andrews-Larson, C., Keene, K., Melhuish, K., Keller, R., & Fortune, N. (2020). Inquiry and gender inequity in the undergraduate mathematics classroom. *Journal for Research in Mathematics Education, 51*(4), 504. https://doi.org/10.5951/jresematheduc-2020-0043

Johnson, E., Keller, R., & Fukawa-Connelly, T. (2018). Results from a survey of abstract algebra instructors across the United States: Understanding the choice to (not) lecture. *International Journal of Research in Undergraduate Mathematics Education, 4*(2), 254–285. https://doi.org/10.1007/s40753-017-0058-1

Kirin, D., Vroom, K., Larsen, S., Apkarian, N., Progress through Calculus Team. (2017). Instruction in precalculus and single-variable calculus in the United States: A bird's eye view. In A. Weinberg, C. Rasmussen, J. Rabin, M. Wawro, & S. Brown (Eds.), *Proceedings of the 20th Annual Conference on Research in Undergraduate Mathematics education* (pp. 1267–1272). San Diego.

Knaub, A. V., Foote, K. T., Henderson, C., Dancy, M., & Beichner, R. J. (2016). Get a room: the role of classroom space in sustained implementation of studio style instruction. *International Journal of STEM Education, 3*(1), 1–22. https://doi.org/10.1186/s40594-016-0042-3

Kogan, M., & Laursen, S. L. (2014). Assessing long-term effects of inquiry-based learning: A case study from college mathematics. *Innovative Higher Education, 39*(3), 183–199. https://doi.org/10.1007/s10755-013-9269-9

Kreitzer, R. J., & Sweet-Cushman, J. (2022). Evaluating Student Evaluations of Teaching: A Review of Measurement and Equity Bias in SETs and Recommendations for Ethical Reform. *Journal of Academic Ethics*, *20*(1), 73–84. https://doi.org/10.1007/s10805-021-09400-w.

Kressler, B., & Kressler, J. (2020). Diverse student perceptions of active learning in a large enrollment STEM course. *Journal of the Scholarship of Teaching and Learning, 20*(1), 40–64. https://doi.org/10.14434/josotl.v20i1.24688

Kuster, G., Johnson, E., Keene, K., & Andrews-Larson, C. (2018). Inquiry-oriented instruction: A conceptualization of the instructional principles. *Primus, 28*(1), 13–30.

Lampert, M., Beasley, H., Ghousseini, H., Kazemi, E., & Franke, M. (2010). Using designed instructional activities to enable novices to manage ambitious mathematics teaching. In M. S. Stein & L. Kucan (Eds.), *Instructional explanations in the disciplines* (pp. 129–141). Springer.

Larsen, S., Glover, E., & Melhuish, K. (2015). Beyond good teaching: The benefits and challenges of implementing ambitious teaching*. In D. Bressoud, V. Mesa, & C. Rasmussen (Eds.), *Insights and recommendations from the MAA national study of college calculus* (pp. 93–106). MAA Press.

Larsen, S., Johnson, E., & Bartlo, J. (2013). Designing and scaling up an innovation in abstract algebra. *The Journal of Mathematical Behavior, 32*(4), 693–711.

Laursen, S. L., Hassi, M. L., Kogan, M., & Weston, T. J. (2014). Benefits for women and men of inquiry-based learning in college mathematics: A multi-institution study. *Journal for Research in Mathematics Education, 45*(4), 406–418.

Laursen, S. L., & Rasmussen, C. (2019). I on the prize: Inquiry approaches in undergraduate mathematics. *International Journal of Research in Undergraduate Mathematics Education, 5*(1), 129–146.

Leatham, K. R., Peterson, B. E., Stockero, S. L., & Zoest, L. R. V. (2015). Conceptualizing mathematically significant pedagogical opportunities to build on student thinking. *Journal for Research in Mathematics Education, 46*(1), 88–124.

Martinez, A. E., Gehrtz, J., Rasmussen, C., LaTona-Tequida, T., & Vroom, K. (2021). Course coordinator orientations toward their work and opportunities for professional development. *Innovative Higher Education*, 1–20

Murphy, J., Chang, J. M., & Suaray, K. (2016). Student performance and attitudes in a collaborative and flipped linear algebra course. *International Journal of Mathematical Education in Science and Technology, 47*(5), 653–673.

Nelson, M. A. (2010). Oral assessments: Improving retention, grades, and understanding. *Primus, 21*(1), 47–61.

Nguyen, K. A., Borrego, M., Finelli, C. J., DeMonbrun, M., Crockett, C., Tharayil, S., Shekhar, P., Waters, C., & Rosenberg, R. (2021). Instructor strategies to aid implementation of active learning: a systematic literature review. *International Journal of STEM Education, 8*(1), 1–18. https://doi.org/10.1186/s40594-021-00270-7

Nilson, L. (2014). *Specifications grading: Restoring rigor, motivating students, and saving faculty time*. Stylus Publishing.

NRC. (2013). *The mathematical sciences in 2025*. National Academies Press. https://doi.org/10.17226/15269

PCAST. (2012). *Engage to excel: Producing one million additional college graduates with degrees in science, technology, engineering, and mathematics*. President's Council of Advisors on Science and Technology; Office of Science and Technology Policy

Petrillo, J. (2016). On flipping first-semester calculus: A case study. *International Journal of Mathematical Education in Science and Technology, 47*(4), 573–582.

Rainey, K., Dancy, M., Mickelson, R., Stearns, E., & Moller, S. (2019). A descriptive study of race and gender differences in how instructional style and perceived professor care influence decisions to major in STEM. *International Journal of STEM Education, 6*(1), 1–13. https://doi.org/10.1186/s40594-019-0159-2

Rasmussen, C., Apkarian, N., Hagman, J. E., Johnson, E., Larsen, S., Bressoud, D., Progress through Calculus Team. (2019). Characteristics of Precalculus through Calculus 2 programs: Insights from a national census survey. *Journal for Research in Mathematics Education, 50*(1), 98–112.

Rasmussen, C., & Ellis, J. (2013). Who is switching out of calculus and why. In A. M. Lindmeier & A. Heinze (Eds.), *Proceedings of the 37th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 4, pp. 73–80). PME.

Rasmussen, C., & Ellis, J. (2015). Calculus coordination at PhD-granting Universities: More than just using the same syllabus, textbook, and final exam. In D. Bressoud, V. Mesa, & C. Rasmussen (Eds.), *Insights and recommendations from the MAA national study of college calculus* (pp. 107–115). MAA Press.

Rasmussen, C., Zandieh, M., & Wawro, M. (2009). How do you know which way the arrows go? The emergence and brokering of a classroom mathematics practice. In W. M. Roth (Ed.), *Mathematical Representation at the Interface of Body and Culture* (pp. 171–218). Information Age Publishing.

Reinholz, D. L., & Apkarian, N. (2018). Four frames for systemic change in STEM departments. *International Journal of STEM Education, 5*(1), 1–10. https://doi.org/10.1186/s40594-018-0103-x

Saxe, K., & Braddy, L. (2015). *A common vision for undergraduate mathematical sciences programs in 2025*. MAA Press.

Seymour, E., & Hewitt, N. M. (1997). *Talking about leaving: Why undergraduates leave the sciences*. Westview Press.

Seymour, E., & Hunter, A.-B. (2019). Talking about leaving revisited: Persistence, relocation, and loss in undergraduate STEM education. *Springer*. https://doi.org/10.1007/978-3-030-25304-2

Shadle, S. E., Marker, A., & Earl, B. (2017). Faculty drivers and barriers: Laying the groundwork for undergraduate STEM education reform in academic departments. *International Journal of STEM Education, 4*(1), 1–13.

Shekhar, P., Borrego, M., DeMonbrun, M., Finelli, C., Crockett, C., & Nguyen, K. (2020). Negative student response to active learning in STEM classrooms. *Journal of College Science Teaching, 49*(6), 45–54.

Sonnert, G., & Sadler, P. (2015). The impact of instructor and institutional factors on students' attitudes. In D. Bressoud, V. Mesa, & C. Rasmussen (Eds.), *Insights and recommendations from the MAA national study of college calculus* (pp. 17–29). MAA Press.

Sonnert, G., Sadler, P. M., Sadler, S. M., & Bressoud, D. M. (2015). The impact of instructor pedagogy on college calculus students' attitude toward mathematics. *International Journal of Mathematical Education in Science and Technology, 46*(3), 370–387. https://doi.org/10.1080/0020739X.2014.979898

Speer, N. M., & Wagner, J. F. (2009). Knowledge needed by a teacher to provide analytic scaffolding during undergraduate mathematics classroom discussions. *Journal for Research in Mathematics Education, 40*(5), 530–562.

Stains, M., Harshman, J., Barker, M. K., Chasteen, S. V., Cole, R., DeChenne-Peters, S. E., Eagan, M. K., Esson, J. M., Knight, J. K., Laski, F. A., Levis-Fitzgerald, M., Lee, C. J., Lo, S. M., McDonnell, L. M., McKay, T. A., Michelotti, N., Musgrove, A., Palmer, M. S., Plank, K. M., … Young, A. M. (2018). Anatomy of STEM teaching in North American universities. *Science, 359*(6383), 1468–1470. https://doi.org/10.1126/science.aap8892

Stains, M., & Vickrey, T. (2017). Fidelity of Implementation: An overlooked yet critical construct to establish effectiveness of evidence-based instructional practices. *CBE Life Sciences Education, 16*(1), 1–11. https://doi.org/10.1187/cbe.16-03-0113

Staples, M. (2007). Supporting whole-class collaborative inquiry in a secondary mathematics classroom. *Cognition and Instruction, 25*(2–3), 161–217.

Stein, M. K., Engle, R. A., Smith, M. S., & Hughes, E. K. (2008). Orchestrating productive mathematical discussions: Five practices for helping teachers move beyond show and tell. *Mathematical Thinking and Learning, 10*(4), 313–340.

Street, C., Apkarian, N., Gehrtz, J., Tremaine, R., Barron, V., Voigt, M., & Hagman, J. E. (2021). *X-PIPS-M Data Summary*. Available: arXiv:2111.01795.

Sturtevant, H., & Wheeler, L. (2019). The STEM Faculty Instructional Barriers and Identity Survey (FIBIS): Development and exploratory results. *International Journal of STEM Education, 6*(1), 1–22. https://doi.org/10.1186/s40594-019-0185-0

Tallman, M. A., Carlson, M. P., Bressoud, D. M., & Pearson, M. (2016). A characterization of calculus I final exams in US colleges and universities. *International Journal of Research in Undergraduate Mathematics Education, 2*(1), 105–133.

Tharayil, S., Borrego, M., Prince, M., Nguyen, K. A., Shekhar, P., Finelli, C. J., & Waters, C. (2018). Strategies to mitigate student resistance to active learning. *International Journal of STEM Education, 5*(1), 1–16. https://doi.org/10.1186/s40594-018-0102-y

Theobald, E. J., Hill, M. J., Tran, E., Agrawal, S., Arroyo, E. N., Behling, S., Chambwe, N., Cintrón, D. L., Cooper, J. D., Dunster, G., Grummer, J. A., Hennessey, K., Hsiao, J., Iranon, N., Jones, L., Jordt, H., Keller, M., Lacey, M. E., Littlefield, C. E., … Freeman, S. (2020). Active learning narrows achievement gaps for underrepresented students in undergraduate science, technology, engineering, and math. *Proceedings of the National Academy of Sciences*, *117*(12), 6476–6483. https://doi.org/10.1073/pnas.1916903117

Walter, E. M., Henderson, C. R., Beach, A. L., & Williams, C. T. (2016). Introducing the Postsecondary Instructional Practices Survey (PIPS): A concise, interdisciplinary, and easy-to-score survey. *CBE Life Sciences Education, 15*(4), 1–11.

Wiggins, G. P., & McTighe, J. (2005). *Understanding by design*. Association for Supervision & Curriculum Development.

Publisher's Note