RESEARCH

Evaluating the impact of malleable factors on percent time lecturing in gateway chemistry, mathematics, and physics courses

Brandon J. Yik[1], Jeffrey R. Raker[1*], Naneh Apkarian[2], Marilyne Stains[3], Charles Henderson[4], Melissa H. Dancy[5] and Estrella Johnson[6]

Abstract

Background: Active learning used in science, technology, engineering, and mathematics (STEM) courses has been shown to improve student outcomes. Nevertheless, traditional lecture-orientated approaches endure in these courses. The implementation of teaching practices is a result of many interrelated factors including disciplinary norms, classroom context, and beliefs about learning. Although factors influencing uptake of active learning are known, no study to date has had the statistical power to empirically test the relative association of these factors with active learning when considered collectively. Prior studies have been limited to a single or small number of evaluated factors; in addition, such studies did not capture the nested nature of institutional contexts. We present the results of a multi-institution, large-scale ($N = 2382$ instructors; $N = 1405$ departments; $N = 749$ institutions) survey-based study in the United States to evaluate 17 malleable factors (i.e., influenceable and changeable) that are associated with the amount of time an instructor spends lecturing, a proxy for implementation of active learning strategies, in introductory postsecondary chemistry, mathematics, and physics courses.

Results: Regression analyses, using multilevel modeling to account for the nested nature of the data, indicate several evaluated contextual factors, personal factors, and teacher thinking factors were significantly associated with percent of class time lecturing when controlling for other factors used in this study. Quantitative results corroborate prior research in indicating that large class sizes are associated with increased percent time lecturing. Other contextual factors (e.g., classroom setup for small group work) and personal contexts (e.g., participation in scholarship of teaching and learning activities) are associated with a decrease in percent time lecturing.

Conclusions: Given the malleable nature of the factors, we offer tangible implications for instructors and administrators to influence the adoption of more active learning strategies in introductory STEM courses.

Keywords: Contextual factors, Personal factors, Beliefs about teaching, Institutional change, Undergraduate education, Teaching practices, Evidence-based instructional practices, Chemistry, Mathematics, Physics

Introduction

It is established that using active learning instructional approaches (i.e., less time spent lecturing) are associated with higher conceptual understanding and persistence in

postsecondary (i.e., undergraduate) STEM courses (Freeman et al., 2014; Lorenzo et al., 2006; Ruiz-Primo et al., 2011; Springer et al., 1999; Theobald et al., 2020). This result holds across a variety of class sizes, disciplines, and levels (Freeman et al., 2014). Importantly, many studies on the use of active learning have also demonstrated a reduction in achievement gaps for minoritized populations (i.e., low-income students or underrepresented minorities) in STEM (Kogan & Laursen, 2014; Lorenzo

*Correspondence: jraker@usf.edu
[1] Department of Chemistry, University of South Florida, Tampa, FL 33620-5250, USA
Full list of author information is available at the end of the article

et al., 2006; Theobald et al., 2020). In particular, active learning has been shown to reduce achievement gaps in exam scores by a third between minoritized groups and non-minoritized groups, narrow gaps in passing rates by nearly half (Theobald et al., 2020), decrease failure rates (as defined by the percentage of students receiving a D or F grade, or withdrawing from the course), and increase achievement across all STEM disciplines when compared to traditional lecture courses (Freeman et al., 2014). Despite evidence of the benefit of active learning strategies (Ballen et al., 2017; Freeman et al., 2014; Haak et al., 2011; Harris et al., 2020; Styers et al., 2018; Theobald et al., 2020), observation-based studies have confirmed that lecture-oriented pedagogical approaches remain as significant components in most STEM courses (Stains et al., 2018), which may be due in part to the lack of departmental norms for using research-based instructional methods (Henderson & Dancy, 2007; Shadle et al., 2017), faculty reward structures (Brownell & Tanner, 2012; Michael, 2007; Shadle et al., 2017), and student resistance (Henderson & Dancy, 2007; Michael, 2007; Shadle et al., 2017).

Prior work in understanding the adoption of active learning strategies has strived towards identifying factors or has considered a single or small number of associated factors. For example, studies have used faculty discussions (Shadle et al., 2017), interviews (e.g., Henderson & Dancy, 2007; Oleson & Hora, 2014), and survey methodologies (e.g., Gibbons et al., 2018; Johnson et al., 2019; Lund & Stains, 2015) to identify factors related to the influence of pedagogical decisions; large class sizes, fixed-seat classroom layouts, lack of pedagogical knowledge in research-based instructional practices, and insufficient faculty assessment methods and processes have all been reported as barriers to uptake of active learning strategies. Lund and Stains (2015) explored departmental influences (e.g., perceived norms towards teaching) and classroom influences (e.g., class size and layout) on adopted pedagogies also finding that departments place high expectations on research output and such classroom elements further restrict instructor teaching. Studies by Gibbons et al. (2018) and Popova et al. (2020) explored the link between postsecondary instructors' thinking and enacted instructional practices, showing a connection between the two. Later, Popova et al. (2021) found evidence for the interconnectedness of instructors' beliefs about teaching and learning with personal (e.g., nature and extent of instructors' preparation and learning efforts) and contextual factors (e.g., course, department, and broader cultural contexts).

However, due to small sample sizes and lack of statistical power, these studies could not empirically test the relative association of these factors when considered all together with active learning. In addition, these studies did not account for the nested nature of their institutional contexts (i.e., instructors within departments within institutions). Therefore, a large-scale, multidisciplinary study of malleable factors (i.e., things that can be changed and altered) related to adoption of such active learning pedagogies in postsecondary STEM courses is needed to complement the research literature and provide further opportunity for actionable changes at instructor, department, and institution levels.

The study reported herein focuses specifically on instructors of introductory chemistry, mathematics, and physics and the malleable factors that influence their uptake of active learning practices as measured by a proxy, i.e., percent time not lecturing. Specifically, we use multilevel modeling to account for the nested nature of our data by discipline and institution, and evaluate 17 factors situated within the three categories (i.e., contextual factors, personal factors, and teacher thinking) as to their relationship with reported percent time lecturing. The teacher-centered systemic reform model (Gess-Newsome et al., 2003; Woodbury & Gess-Newsome, 2002) suggests that factors within these categories are related to enacted teaching practices. In the next section, we describe this conceptual framework and detail the literature that report these 17 factors related to the adoption of active learning.

Conceptual framework

Research on pedagogy adoption and colloquial anecdotes about why instructors choose to enact active learning pedagogies informed the selection of modeled factors in our study. The teacher-centered systemic reform model was developed from an exhaustive review of the literature as a mechanism to understand the evolution of classroom practices as a result of reform initiatives (Woodbury & Gess-Newsome, 2002) and was later modified to better reflect the nature of teaching in a university context (Gess-Newsome et al., 2003). This framework is useful as it considers the situational teaching contexts along with an instructor's educational influences and their beliefs about teaching and learning all within a complex educational system.

Fundamentally, the TCSR framework is focused on teacher change as the source of grander changes within the larger institutional system. An instructor is the ultimate authority of the enacted practices that occur in a classroom and is aligned with the TCSR model's theoretical underpinnings that instructors' beliefs influence their practices as embedded with a larger system, such as in classrooms, departments, institutions, and disciplines (Woodbury & Gess-Newsome, 2002). This study aims to quantify the extent that malleable factors have on

the uptake of active learning and the TCSR framework is best suited, since its focus is on the instructor and their personal and teaching contexts which are the most malleable along with nested contexts (i.e., instructors within departments that are within institutions) in which teaching reform occurs.

According to Gess-Newsome et al. (2003), the TCSR model for a university context has three broad categories: contextual factors, personal factors, and teacher thinking factors. Situated within contextual factors are the broader cultural context (e.g., teacher development and teaching materials), school context (e.g., institution type, physical space, and technology), department and subject area content (e.g., department and cultural norms and teacher's class load), and classroom context (e.g., class size and physical organization of the room). Situated within personal factors are instructors' demographic profile, types and years of teaching experience, and nature and extent of teacher preparation and continued learning efforts. Situated within teacher thinking factors are instructors' sense of dissatisfaction with current practices; and knowledge and beliefs about teachers and teachers' roles, students and learning, schooling and schools, and content being taught.

Extensive literature used to develop the TCSR model aims at capturing as many of the intricacies that can be situated under the broad factors that comprise the general context of reform. Nevertheless, the framework does not fully capture all of the complexities of higher education institutions when the model was used to frame the study reported herein. As a consequence of the literature review on malleable factors that affect pedagogical change reported herein, our conceptualization of the TCSR model necessitated further modifications. For example, factors were found that necessitated inclusion in our study as distinctly different from department contexts (e.g., discipline and cultural norms), such as tenure status (e.g., non-tenure-track lecturers, tenure-track faculty, and tenured faculty), teaching load, and instructors' teaching evaluation. It is crucial to delineate *department appointment expectations* as a subcategory under contextual factors apart from *department contextual factors* to aid in accounting for the different department and institutional policies that differ with the distinct instructional positions (e.g., lecturers versus tenure-track professors). Higher education systems undergo gradual change and theoretical frameworks on change theories need to be explored and reevaluated in light of new research. Findings since the original conceptualization necessitated modification of the TCSR model with the ones we present in this study.

Our study intends to evaluate the effects of malleable factors related to adoption of active learning pedagogies when controlling for institutional and disciplinary differences. Thus, non-malleable factors, such as race/ethnicity, are not included. In addition, while this study attempts to account for as many of the malleable factors reported in the literature, it is not possible to (1) ask respondents about all possible aspects of the TCSR model in a survey and (2) statistically test all factors that may be present in graphical representations of the TCSR model (cf. Gess-Newsome et al., 2003; Woodbury & Gess-Newsome, 2002); the associated sample size requisite for sufficient statistical power grows with a larger number of tested factors. To balance complexity and parsimony, we include malleable factors found in the literature that have been previously cited as barriers to implementation or as reasons for uptake of active learning strategies. While our study contains factors from all three broad categories (i.e., contextual, personal, and teacher thinking) and all of their subcategories, some factors identified in the model (cf. Gess-Newsome et al., 2003) are either not malleable (e.g., physical location and college president) or difficult to quantify in a statistical model (e.g., instructor's daily/weekly schedule and student personal expectations).
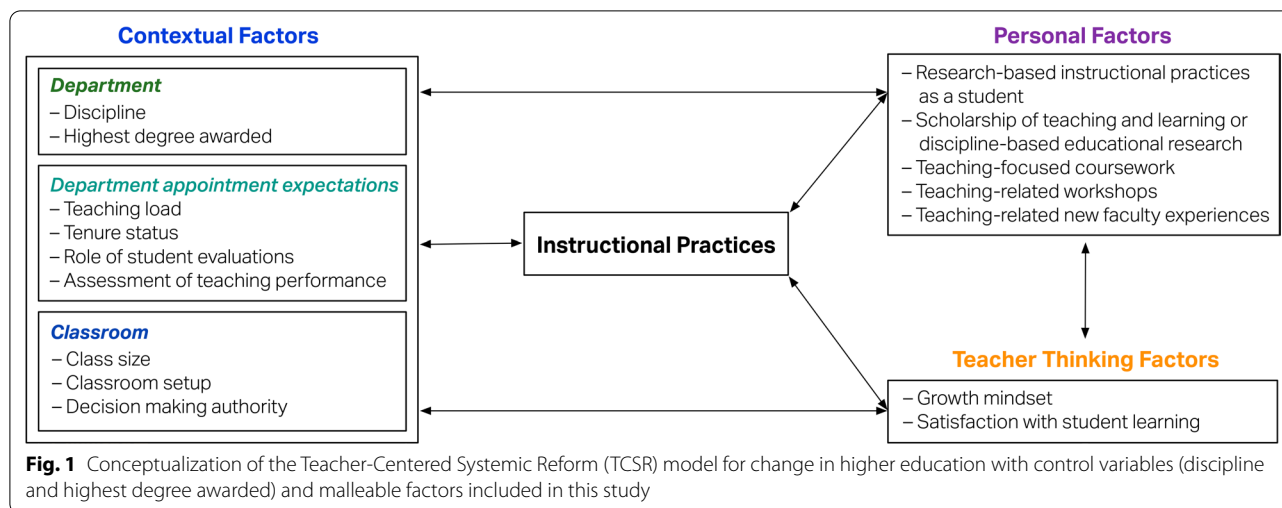
In this study, we include malleable factors that have been found and discussed in many and different STEM fields as specific disciplines may lack literature in that area. While STEM fields may show some disciplinary differences, these factors can be assumed to affect all STEM fields to some extent (Lund & Stains, 2015). In the next sections, we describe the evidence-based factors grounded in the literature that have been found to affect the uptake of active learning under each of the three broad categories in the TCSR model that were tested in this study (see Fig. 1).

Contextual factors
Contextual barriers to instructional change have been widely studied (e.g., Fairweather & Rhoads, 1995; Hora, 2012; Lund et al., 2015; Shadle et al., 2017; Stains et al., 2018). In this study, three types of contextual factors were explored: department characteristics, department appointment expectations, and classroom contexts.

Department characteristics
Department characteristics that have been reported to be associated with the adoption of active learning include: (1) discipline (i.e., chemistry, mathematics, or physics) and (2) the highest degree the department awards (i.e., associates, bachelor's, or graduate). While these factors are not malleable, they are included as control variables for departmental characteristics when testing the association between malleable factors. It has been reported that instructional practices differ between STEM disciplines (Fairweather & Rhoads, 1995; Hora & Anderson,

Fig. 1 Conceptualization of the Teacher-Centered Systemic Reform (TCSR) model for change in higher education with control variables (discipline and highest degree awarded) and malleable factors included in this study

2012; Lund et al., 2015; Stains et al., 2018). Several studies have indicated that the balance between teaching and research at one's institution and department impact how teaching is approached. The highest degree awarded in the department has been shown to be a viable proxy for the extent of focus a department places teaching versus research (cf. Cox et al., 2011; Srinivasan et al., 2018). For example, instructors that teach in departments with graduate degree offerings are presumed to have a greater focus on research than their counterparts in departments with associate degree programs.

Department appointment expectations
Department appointment expectations that have been reported to be associated with the adoption of active learning include: (1) teaching load, (2) tenure status, (3) the role of student evaluations, and (4) the role assessment of teaching in review, promotion, or tenure. While the department may have set teaching loads and standards for the role of student evaluations and assessment of teaching performance in review, promotion, or tenure, there is the possibility, especially in larger departments with an array of teaching personnel, that evaluation may be unique or differentiated by appointment.

Teaching load has been reported to be associated with teaching practices, with higher teaching loads being attributed to an instructor being pressed for time (Hora, 2012). Lack of available time devoted to teaching activities has the potential to result in a lack of innovative pedagogies. Henderson and Dancy (2007) noted that one of the largest barriers reported by physics instructors was a heavy teaching load. In this study, we separate teaching load as a distinct factor from tenure status and institution type. From a broader view, teaching loads might be an indicator of tenure status; for example, an instructor

with no opportunity for tenure (e.g., a lecturer or visiting instructor) may have higher teaching loads. In addition, a teaching load could correlate with institution type; an instructor at an institution with a larger teaching focus (e.g., a primarily undergraduate institution; PUI) may have a higher teaching load. However, upon further inspection, different instructional positions can hold different tenure statuses and be at different institution types. For example, a tenured professor at a PUI may have a large teaching load or a lecturer at a large research-intensive institution may have a high teaching load.

In this study, respondents are grouped into three appointment categories: (1) no opportunity for tenure, (2) tenure-track, and (3) and tenured. Tenure status has been shown to have an association with the amount of adoption of research-based instructional strategies (Landrum et al., 2017; Shadle et al., 2017). Those teaching undergraduate STEM courses, especially those with the privilege of obtaining tenure at research-intensive institutions, have the ability to identify as both an educator/teacher and a researcher. Implementation of active learning strategies may more time consuming (Beatty et al., 2005; Drinkwater et al., 2014), thus, tenure-track and tenured faculty members, in theory, have to weigh time spent on teaching and research, among other responsibilities. Fairweather (2008) reported that untenured faculty members are least likely to be productive in *both* teaching and research compared to being productive in either teaching or research; therefore, tenure status is suggested to be influential in instructional decisions. Landrum et al. (2017) corroborates this idea and found significant differences in evidence-based instructional practice adoption between tenure/tenure-track faculty and their non-tenure-track counterparts, with the former reporting significantly higher use of these practices. Being tenured can

allow for more freedom and flexibility to use innovative teaching methods (Hora, 2012).

Instructor's perceived value of how their department or institution values teaching, both in their assessment of their teaching and from student evaluations, plays an important role in the instructor's role as a teacher; studies have reported that if it was the norm for instructors in a department to integrate research-based methods into their teaching then it was easier for others to do so and that there is no uniform method of evaluating and rewarding one's teaching (e.g., Brownell & Tanner, 2012; Gess-Newsome et al., 2003; Henderson & Dancy, 2007; Hora & Anderson, 2012; Prosser & Trigwell, 1997; Seymour et al., 2011; Sturtevant & Wheeler, 2019; Walczyk et al., 2007). Lund and Stains (2015) reported that a substantial number of STEM instructors describe departmental and tenure pressures as influential to their teaching practices, while few faculty say student evaluations of teaching influence their teaching approaches (Erdmann et al., 2020). In a study of mathematics instructors, Johnson et al. (2018) reported that while the most popular reason reported for not attempting instructional change was a lack of time for course redesign, roughly 20% of respondents reported that they believed their departments would not support them and that instructional change would not be valued in their annual review, promotion, or tenure process.

Classroom contextual factors

Classroom contextual factors include characteristics of the classroom learning environment that influence pedagogical decisions. Classroom contextual factors that have been reported to be associated with the adoption of active learning include: (1) class size, (2) classroom layout, and (iii) decision making authority over instructional choices.

Class size and classroom layout have been reported to be influential and strong barriers to the implementation of active learning strategies (e.g., Henderson & Dancy, 2007; Lund & Stains, 2015; Lund et al., 2015; Michael, 2007; Prosser & Trigwell, 1997; Shadle et al., 2017; Sturtevant & Wheeler, 2019). STEM instructors cite factors such as large class sizes (i.e., over 100 students) as a reason why they have not chosen to adopt interactive teaching methods (Hora & Anderson, 2012; Shadle et al., 2017). and is corroborated by a significant correlation between class size and percent time spent lecturing (Smith et al., 2014). However, fixed classroom layouts are not prohibitive to enacting active learning pedagogies; in an observation study by Lund et al. (2015), various levels of student–student interactions (i.e., active learning) were observed in large classes taught in fixed-seat lecture halls. This may indicate that there is a relationship between class size and classroom layout.

In a case study of doctoral degree granting institutions, it was found that course coordination was one of the seven features that contributed to successful calculus programs (Bressoud & Rasmussen, 2015; Rasmussen et al., 2014). Rasmussen et al. (2019) reported that high-quality active learning can be attributed, in part, by the support systems that course coordination affords. Communication channels, both formal and interpersonal, are key to the diffusion of innovations (Rogers, 2003), for example the dissemination of research-based instructional practices. In the context of adopting innovative teaching practices, such communication could be facilitated through a course coordinator with positional and personal instructional influence (Apkarian & Rasmussen, 2017; Bazett & Clough, 2021; Golnabi et al., 2021; Lane et al., 2019).

By leveraging common tools and resources, encouraging collaboration and shared objectives, and promoting professional development, course coordinators can act as change agents (Williams et al., 2021). Coordination can catalyze community-building and collaboration between instructors; information was also shared at department meetings and retreats which led to meaning conversations centered around teaching (Bazett & Clough, 2021; Williams et al., 2021). Visitors, instructors, teaching assistants, adjuncts, and lecturers (VITAL; Levy, 2019) that teach sections of coordinated mathematics courses have access to course coordinators to discuss pedagogical approaches and active learning activities (Golnabi et al., 2021). In addition, coordinators of active learning courses in mathematics have also been shown to utilize local data (e.g., student performance data, grades in subsequent courses, and student-generated data) to inform curriculum and pedagogy (Martinez & Pilgrim, 2021). In physics, co-teaching between a new instructor and an experienced instructor that uses active learning showed immediate uptake of these teaching practices by the new instructor and a positive shift in their beliefs and intentions of using these strategies in the future (Henderson et al., 2009).

Personal factors

Personal factors include the nature and extent of teaching preparation and experience, and teaching-related training. Personal factors that have been reported to be associated with adoption of active learning that were explored in this study include: (1) experience with research-based instructional strategies (RBIS) as a student, (2) completion of teaching-focused coursework, (3) participation in new faculty experiences or workshops,

and (4) participation in scholarship of teaching (SOTL) or discipline-based education research (DBER).

Previous experience with RBIS as a student impacts instructional decisions (Lund & Stains, 2015; Oleson & Hora, 2014). In a study of upper-division mathematics instructors, it was reported that, by far, the two most influential factors on instructional practices were both their experiences as a teacher and as a student (Fukawa-Connelly et al., 2016). In interviews of STEM instructors, many cite that knowledge regarding teaching explained the selection of teaching techniques including active learning strategies (Oleson & Hora, 2014). In a study of biologists, chemists, and physicists, instructors who had experienced RBIS as a student were more likely to implement RBIS in their own teaching (Lund & Stains, 2015).

In addition, instructors' previous coursework, which can be a part of their teacher preparation, influences teaching decisions (Hora, 2012; Lotter et al., 2007; Southerland et al., 2011a, 2011b; Windschitl & Sahl, 2002). It has been reported that doctoral training influences teaching approaches for some individuals (Hora, 2012; Southerland et al., 2011a, 2011b), and more specifically, university coursework also alters teaching pedagogy (Windschitl & Sahl, 2002).

Dissemination of research-based instructional strategies through new faculty experiences and workshops have been reported to increase instructor awareness and inclusion of these instructional strategies in many STEM fields (e.g., Ebert-May et al., 2015; Henderson, 2008; Stains et al., 2015). In particular, the chemistry education community has introduced instructors to RBIS and advocated for the adoption of RBIS through the Multi-Initiative Dissemination Project workshops via four reform projects: ChemConnections, Molecular Science, New Traditions—now known as Process Oriented Guided Inquiry Learning (POGIL), and Peer-Led Team Learning (cf. Burke et al., 2004; Landis et al., 1998; Peace et al., 2002). More recently, the Cottrell Scholars Collaborative New Faculty Workshop was established to prepare chemistry instructors at becoming teacher–scholars by engaging with evidence-based teaching methods (Baker et al., 2014; Stains et al., 2015). Other initiatives include the Core Collaborators Workshops for biochemistry (Murray et al., 2011), POGIL workshops for physical chemistry laboratory (Stegall et al., 2016), and Active Learning in Organic Chemistry workshops (Houseknecht et al., 2020).

In the mathematics community, Project NExT (New Experiences in Teaching) and MathFest minicourses, along with other workshops through the Mathematical Association of America are meant to disseminate new teaching pedagogies. However, Fukawa-Connelly et al. (2016) report that only very small percentages of workshop participants found them to be very influential in their teaching and little importance were assigned to these aforementioned workshops. In addition, instructors may participate in workshops through the Academy of Inquiry Based Learning to shape their teaching of inquiry-based learning (Fukawa-Connelly et al., 2016).

Within the physics education community, a long-standing workshop spanning more than two decades has been run for new physics and astronomy faculty to increase awareness of RBIS (Henderson, 2008; Henderson et al., 2012). In addition, these types of workshops that expand awareness and utilization of RBIS have been established in biology through the Faculty Institutes for Reforming Science Teaching programs and National Academies Summer Institutes on Undergraduate Education (Derting et al., 2016; Ebert-May et al., 2015; Handelsman et al., 2004; Wood & Gentile, 2003).

It is the assumption that conducting or participating in SOTL or DBER allow instructors use the knowledge gained from that scholarly work to better their teaching practices. While no study, to our knowledge, exists that specifically explores the enacted instructional practices of DBER instructors, SOTL and DBER are closely related pursuits and participation in either will yield similar changes (Henderson et al., 2012). In a meta-analysis of undergraduate STEM instructional practices (Henderson et al., 2011), instructor participation in SOTL has been found to be associated with improvements in course and program-level curricula.

Teacher thinking factors

Teacher thinking factors include an instructor's beliefs about teaching and level of dissatisfaction with their current practices and student learning. Teacher thinking factors that have been reported to be associated with the adoption of active learning that were explored in this study include: (1) satisfaction with student learning and (2) holding a growth mindset.

Erdmann et al. (2020) reported a small relationship between the level of instructor dissatisfaction and pedagogical revisions in a range of STEM disciplines in a sample of primarily biology, chemistry, mathematics, and physics instructors. Teacher thinking has been shown to bolster and hinder the use of new ideas and technologies in constructing classroom and course learning environment (Moore, 2002; Rogers, 2003). Adoption of new teaching strategies begins with an instructors' dissatisfaction with the current instruction or a belief that students learn better with strategies not being currently utilized (e.g., Andrews & Lemons, 2015; Bauer et al., 2013; Gess-Newsome et al., 2003; Gibbons et al., 2018; Lotter et al., 2007; Windschitl & Sahl, 2002). Pedagogical dissatisfaction is when an instructor

realizes a misalignment of their instructional goals with their instructional practice (Southerland et al., 2011a, 2011b). This disconnect between goals and practice can result in a revision of teaching practice and the adoption of new pedagogical strategies (Feldman, 2000).

Instructors' mindset beliefs have been reported to likely influence how their courses are structured (Rattan et al., 2012). A fixed or growth mindset is a belief in the inflexibility or malleability, respectively, of a human characteristic (e.g., intelligence; Dweck, 1999). Holding a particular mindset, for example beliefs that traits (e.g., student intelligence) are rigid and cannot be changed (fixed mindset) or can be developed with time and experience (growth mindset), is related to teaching practice choices (Canning et al., 2019). In a longitudinal study of STEM faculty, including chemistry, mathematics, and physics, who endorsed growth mindsets used more motivating pedagogical practices (Canning et al., 2019), which can include active learning strategies (Armbruster et al., 2009; Prince, 2004; Springer et al., 1999).

Holding a growth mindset has been reported to be associated with the uptake of evidenced-based practices, such as active learning, in STEM faculty (Bathgate et al., 2019). Instructors' mindset beliefs were found to be related to the adoption of active-learning practices in biology; fixed mindset instructors taught using a teacher-centered focus (e.g., lecturing) and growth mindset instructors taught using a student-centered approach (e.g., active learning; Aragón et al., 2018). In mathematics, fixed mindsets were associated with using teaching strategies that would reduce student engagement and achievement (Rattan et al., 2012); in addition, growth mindset beliefs were held by mathematics instructors that were more willing to consider non-lecture pedagogies (Johnson et al., 2018). In observations of instructors that teach introductory STEM courses, Ferrare (2019) reported that instructors' beliefs about student learning were linked to certain instructional styles; for example, instructors espousing fixed mindset beliefs were more likely to teach using "chalk talks."

Interventions based on growth mindset have been reported to be effective (e.g., Dweck, 1999; Dweck & Leggett, 1988; Yeager, Romero, et al., 2016; Yeager, Walton, et al., 2016). In addition, mindset interventions are generalizable and replicable (Bettinger et al., 2018; Yeager et al., 2019). However, there have been some contest as to their replicability (e.g., Bahník & Vranka, 2017; Li & Bates, 2017). An additional challenge comes from practitioners' misinterpretations of growth mindset and ways to promote it (Dweck, 2019). While current research efforts are underway to evaluate the applications of growth mindset and interventions (McMahon et al.,

2019), the meta-analyses have shown effectiveness in students (Sisk et al., 2018).

Framework conceptualization
The TCSR model is functional for understanding the adoption of active learning strategies due to the interconnectedness of contextual factors, personal factors, and teacher thinking factors along with the interactions between them, when considering why particular teaching practices are enacted in the classroom situated within the larger institutional context. Work in STEM education (e.g., Henderson & Dancy, 2007; Henderson et al., 2011; Johnson et al., 2018, 2019; Lund & Stains, 2015; Oleson & Hora, 2014; Shadle et al., 2017) corroborates these three factors associated with adoption (or barriers to adoption) of more active learning instruction: (1) contextual factors, (2) personal factors, and (3) teacher thinking factors. Figure 1 (see above) summarizes the malleable factors included in this study situated within our conceptualization of the TCSR model.

Research question
The conceptual framework described informed the development of the following research question we seek to answer in this study:

> *To what extent are contextual factors, personal factors, and teacher thinking factors associated with percent time lecturing in gateway chemistry, mathematics, and physics courses when controlling for all other factors and accounting for the nested nature of the data (i.e., instructors within departments within institutions)?*

Methods
We employed survey methodology and used quantitative approaches to answer the research question. Quantitative analysis of the data included multilevel modeling to account for the nested structure of the data (i.e., instructors within departments within institutions). Below, survey development and the nature of the participants included in this study is first described. Then, the multilevel modeling methods are described along with the specific malleable factors (i.e., variables) evaluated in this study.

Survey development
The survey instrument from which specific items are used in this study was developed and informed by previous large-scale studies in postsecondary chemistry (Gibbons et al., 2018; Lund & Stains, 2015; Stains et al., 2018), mathematics (Apkarian et al., 2019; Johnson

Yik *et al. International Journal of STEM Education*   (2022) 9:15

Page 8 of 23

et al., 2018), and physics (Henderson & Dancy, 2009; Walter, Henderson, et al., 2016; Walter et al., 2021). The full survey asked instructors about five main topics: (1) course context, (2) instructional practices, (3) awareness and usage of active learning instructional techniques, (4) perceptions, beliefs, and attitudes related to students, learning, and departmental context, and (5) personal demographics and experience. Where applicable, previous instruments and scales with reliability and validity evidence were used (e.g., mindset: Dweck et al., 1995). Single-item constructs in the survey show content and face validity with expert review. Our interpretation of the survey results in this study is the inherent consequential validity that is presented in the Discussion. Survey items used in this study are detailed in Additional file 1: Table S1.

Participants

A database was constructed of instructors teaching postsecondary introductory chemistry, mathematics, and physics in the United States ($n_{total} = 18,337$). Instructors in this database were identified through stratified random sampling based on institution type; the goal was to create a representative sample of institution types: 2-year institutions (i.e., associate degree-granting, 4-year institutions (i.e., bachelor's degree-granting), and universities (i.e., graduate degree-granting). All instructors at each of the selected institutions teaching the targeted introductory level courses were added to the database and invited to participate in the study.

The database includes 9404 instructors at 4-year institutions (including bachelor's and graduate degree-granting institutions) in the United States that had conferred at least one bachelor's degree in all three disciplines (chemistry, mathematics, and physics) between 2011 and 2016 as recorded by the National Center for Education Statistics' Integrated Postsecondary Education Data System. In addition, the database also includes 8,933 instructors at 2-year institutions in the United States that offer all three of the courses. Contact information for these instructors ($n_{total} = 18,337$) was compiled by the American Institute of Physics Statistical Research Center using publicly accessible online information and through communication with department chairs at the target institutions. Potential survey respondents needed to have taught general chemistry, single-variable calculus, or quantitative-based introductory physics as the primary instructor in 2 years prior to data collection (i.e., in the 2017–18 and/or 2018–19 academic year); in addition, the survey

respondents had to have not taught the course exclusively online.

Data collection

Data were collected via the custom-built survey, as described above, was overseen by the American Institute of Physics Statistical Research Center between March 2019 and May 2019; the survey was open for 58 days. Data were collected in accordance with the Western Michigan University Institutional Review Board application no. 17-06-10 approved June 20, 2017, with informed consent obtained digitally. Instructors were sent up to four email notifications: (1) an announcement about the survey, (2) an invitation to participate, and (3 & 4) up to two follow-up invitation emails. A small number of instructors were contacted via follow-up phone calls ($N = 603$) from the American Institute of Physics Statistical Research Center asking them to complete the survey; instructors were prioritized from institutions, where there were two of the three disciplines already represented to have a data set with representation from all three disciplines.

Respondents included 3769 instructors (20.5% unit response rate) comprised of 2670 instructors at 4-year institutions and 1099 instructors at 2-year institutions; respondents included 1244 chemistry, 1349 mathematics, and 1176 physics instructors. In total, 1387 respondents were listwise deleted from the study described herein, because they did not answer all the survey items used in the construction of the multilevel models. The study sample thus included 2382 instructors from 1405 departments at 749 institutions for which complete data were collected for the survey items used in this study. The study sample included 795 chemistry, 778 mathematics, and 809 physics instructors with 1764 instructors at 4-year and 618 instructors at 2-year institutions.

Multilevel models

A three-level model was used to evaluate the impact of malleable factors on amount lecturing in introductory courses in chemistry, mathematics, and physics to account for the nested structure of the data. In this model, instructors (level 1) can be thought of being nested within departments (level 2) which are nested within institutions (level 3). Instructors may, therefore, be affected by grouping effects at the department and institution levels; this violates the independence of observations assumption required by traditional ordinary least squares regression techniques but can be accounted for in multilevel regression models (Raudenbush & Bryk, 2002; Snijders & Bosker, 2012; Theobald, 2018). If this

nested nature of the data is not accounted for, then data may be analyzed at one level with conclusions drawn at another level; this phenomenon is known as an ecological fallacy (Robinson, 1950). Several studies advocate for this specific three-level model (Porter & Umbach, 2001; Smart & Umbach, 2007) and other studies implement variations of this nesting model to align with their research questions (e.g., Marsh et al., 2002; Smeby & Try, 2005; Sonnert et al., 2007). Descriptive statistics for instructor- and department-level factors are given in Additional file 1: Tables S2 and S3, respectively. Correlations among instructor- and department-level factors are reported in Additional file 1: Tables S4 and S5, respectively. Variance inflation factors (VIF), reported in Additional file 1: Table S6, are well under the suggested cutoff value of 10 and thus do not indicate multicollinearity between predictor variables (Myers, 1990).

Models were constructed using the lme4 package (Bates et al., 2015) in RStudio version 1.2.5033 (R Core Team, 2019) using the full maximum likelihood estimation method due to the large sample size ($N = 2382$ respondents). The lmerTest package was used to obtain p values for fixed effects (Kuznetsova et al., 2017). T tests used the Satterthwaite approximations for degrees of freedom. An iterative model building process was used by adding and subtracting predictor factors using statistical tests for model comparison to obtain the reported model. Models were compared using theory, fit statistics (i.e., deviance), statistical tests (i.e., χ^2 tests), and explained variances. Raudenbuch and Bryk notation (2002) is used to describe the multilevel models. The complete final model is reported in Eq. 1.

$$
\begin{aligned}
\text{LECTURE}_{ijk} = {} & \gamma_{000} + \beta_{001}\text{CHEM}_j + \beta_{002}\text{PHYS}_j + \beta_{003}\text{BACH}_j \\
& + \beta_{004}\text{GRAD}_j + \pi_{100}\text{SIZE}_i + \pi_{200}\text{ROOM}_i \\
& + \pi_{300}\text{SIZE}_i \times \text{ROOM}_i + \pi_{400}\text{DECISION}_i \\
& + \pi_{500}\text{LOAD}_i + \pi_{600}\text{TENURED}_i \\
& + \pi_{700}\text{TENURETRACK}_i + \pi_{800}\text{SET}_i \\
& + \pi_{900}\text{ATP}_i + \pi_{1000}\text{RBIS}_i + \pi_{1100}\text{SOTL}_i \\
& + \pi_{1200}\text{TFC}_i + \pi_{1300}\text{WKSP}_i \\
& + \pi_{1400}\text{NFE}_i + \pi_{1500}\text{GROWTH}_i \\
& + \pi_{1600}\text{SATISFACTION}_i
\end{aligned}
$$

$$(1)$$

Effect sizes indicate the magnitude of the relationship between a predictor variable and the outcome variable. Typically, effect sizes are used as a standardized measure for the comparison of effects within a study or between studies. However, there is no consensus method for calculating effect size in multilevel models (Lorah, 2018; Selya et al., 2012). Cohen's f^2 is advantageous as it is compatible with the nested nature of the data and will be calculated as a measure of local fixed effect sizes and global effect size (Cohen, 1988). Local effect size is the

proportion of explained variance by a given effect relative to the proportion of the unexplained outcome (i.e., percent lecturing) variance, whereas global effect size is the proportion of explained variance by all effects relative to the proportion of the unexplained outcome variance (Lorah, 2018; Selya et al., 2012). Random effect sizes are related to the intraclass correlation coefficient (ICC), which is the proportion of variance in the outcome accounted for by a level in a multilevel model and is, therefore, a measure of strength of association between level membership (department or institution) and the outcome (Lorah, 2018). The ICC represents and effect size index as the magnitude of the association can be interpreted analogously to a correlation coefficient (Snijders & Bosker, 2012).

Explanation of variables

Table 1 includes all factors used in the multilevel model and their coding. In the final multilevel model (Eq. 1), LECTURE$_{ijk}$ is the overall percent lecturing for an instructor i in department j within institution k. Subscripts i and j correspond to variables that apply to individual instructors or their departments, respectively. In the final multilevel model, γ_{000} represents the overall mean intercept, β_{001} through β_{004} represent the department-level (level 2) predictor coefficients, and π_{100} through π_{1600} represent the instructor-level (level 1) predictor coefficients.

In the custom-built survey, the item corresponding to class size ("What was the approximate enrollment in a typical lecture section?") featured text entry for the response. As a result, this resulted in a range of entry formats from specific values to ranges. In addition, we do not expect instructors to make meaningful instructional decisions based on the exact number of students enrolled in a lecture section (e.g., 32 vs. 33 students), we opted to use bins for the class size variable. The Common Data Set Initiative (n.d.), which is used by *U.S. News & World Report* rankings, uses well-known bins for class sizes: 2–19, 20–29, 30–39, 40–59, 60–99, and 100 or more students. The bins allow for a more interpretable multilevel coefficient versus the value for a single individual student in a class (i.e., class size as a continuous variable). For these reasons, binning of the class size variable was chosen to better inform pedagogical decisions.

Results

We report, herein, the results of a national survey on instructional practices (i.e., percent time lecturing) in introductory, gateway chemistry, mathematics, and physics courses (i.e., general chemistry, single-variable calculus, and introductory quantitative physics). These gateway courses have long been identified as a cause for

Table 1 Factors used in the final multilevel model and their coding

Factor	Coding
Dependent variable	
LECTURE: Overall percent lecturing	The overall percent of time during regular class meetings that students spend listening to the instructor lecture or solve problems
Department characteristics	
CHEM: Chemistry (reference group: math)	0 = no, 1 = yes
PHYS: Physics (reference group: math)	0 = no, 1 = yes
BACH: Bachelor's degree is highest offered by department (reference group: associate degree)	0 = no, 1 = yes
GRAD: Graduate degree is highest offered by department (reference group: associate degree)	0 = no, 1 = yes
Department appointment expectations	
LOAD: Teaching load	
TENURE: Tenured instructor (reference group: no opportunity to earn tenure)	0 = no, 1 = yes
TENURETRACK: Tenure-track instructor (reference group: no opportunity to earn tenure)	0 = no, 1 = yes
SET: Role of student evaluation of teaching in review, promotion, or tenure compared to other measures	0 = not used, 1 = less weight, 2 = equal weight, 3 = more weight, 4 = only used
APT: Role of assessment of teaching performance in review, promotion, or tenure compared to other measures	0 = not used, 1 = not influential, 2 = somewhat influential, 3 = influential, 4 = very influential
Classroom contextual	
SIZE: Class size	−2 = 2–19, −1 = 20–29, 0 = 30–39, +1 = 40–59, +2 = 60–99, +3 = 100 + students[a]
ROOM: Classroom setup	0 = fixed seats, 1 = allows for group work
SIZE × ROOM: Class size × classroom setup interaction effect	Interaction effect of class size and classroom setup with coding used above
DECISION: Decision making authority	0 = respondent has sole decision-making authority, 1 = is in collaboration with others to make decisions
Personal factors	
RBIS: Had been in a course as a student that had used RBIS	0 = no, 1 = yes
SOTL: Conducts scholarship of teaching and learning or discipline-based education research	0 = no, 1 = yes
TFC: Has taken teaching-focused coursework at the undergraduate, graduate, or postdoctoral level	0 = no, 1 = yes
WKSP: Has participated in teaching-related workshops	0 = no, 1 = yes
NFE: Has participated in teaching-related new faculty experiences	0 = no, 1 = yes
Teacher thinking	
GROWTH: Degree of holding a growth mindset by the instructor	−2.5 = fixed mindset to +2.5 = growth mindset[b]
SATISFACTION: Instructor's level of satisfaction with student learning	−2 = very dissatisfied to +2 = very satisfied[c]

[a] Grand-median centered at 30–39 students. [b] Average of three items on a six-point Likert scale from 1 to 6 that describe fixed mindset (Dweck et al., 1995); items were reverse coded to represent increasing growth mindset and centered at the middle of the scale. [c] The single item was answered on a five-point Likert scale from 1 (very dissatisfied) to 5 (very satisfied); values were centered at middle of the scale

students not completing STEM degrees (Koch, 2017; Seymour & Hewitt, 1997; Seymour & Hunter, 2019). Data were collected and modeled based on the nested nature of departments and institutions from which the 2382 respondents were sampled.

A multilevel regression model was constructed to evaluate the association of malleable factors with percent lecturing in introductory chemistry, mathematics, and physics courses. Data were modeled by department and institution to account for the lack of independence of observations at these levels: instructors (i, level 1) were nested within departments (j, level 2) nested within institutions (k, level 3). The unconditional model has an ICC of 0.13 for level 2 (department) and an ICC of 0.11 for level 3 (institution); thus ~13% of variability in the dependent variable (i.e., percent time lecturing) is accounted by nesting observations at the department level and ~11% of variability in the outcome variable by nesting at the institution level. When more than 10%

Yik *et al. International Journal of STEM Education*   (2022) 9:15

Page 11 of 23

Table 2 Factors in explaining percent lecturing in introductory chemistry, mathematics, and physics courses

	Estimate	SE	p	Cohen's f^2 (size)[a]
Intercept	82.66	2.78	< 0.001	
Chemistry	0.34	1.25	0.789	< 0.01 (n)
Physics	−4.30	1.20	< 0.001	< 0.01 (n)
Bachelor's program	0.95	1.43	0.505	< 0.01 (n)
Graduate program	1.56	1.63	0.340	< 0.01 (n)
Teaching load	−0.08	0.45	0.855	< 0.01 (n)
Tenured faculty	0.47	1.12	0.671	< 0.01 (n)
Tenure-track faculty	−1.18	1.44	0.414	< 0.01 (n)
Student evaluation of teaching	−0.41	0.47	0.391	< 0.01 (n)
Assessment of teaching performance	−0.04	0.55	0.939	< 0.01 (n)
Class size	1.14	0.50	0.021	< 0.01 (n)
Classroom setup	−10.71	1.08	< 0.001	0.05 (s)
Class size × classroom setup	−1.17	0.67	0.080	< 0.01 (n)
Decision making	−4.60	1.34	< 0.001	< 0.01 (n)
RBIS use as a student	−3.83	1.31	0.003	< 0.01 (n)
Scholarship of teaching and learning	−8.56	0.96	< 0.001	0.03 (s)
Teaching-focused coursework	−3.22	0.97	< 0.001	< 0.01 (n)
Teaching-related workshops	−8.24	1.59	< 0.001	< 0.01 (n)
Teaching-related new faculty experiences	−4.69	1.14	< 0.001	< 0.01 (n)
Growth mindset	−3.08	0.41	< 0.001	0.02 (s)
Satisfaction with student learning	−1.17	0.67	0.080	< 0.01 (n)

[a] Cohen's (1992) f^2 effect size: < 0.02 = negligible (n); ≥ 0.02 = small (s); ≥ 0.15 = medium (m); ≥ 0.35 = large (l)

of variance occurs between levels, a multilevel model is appropriate (Tai et al., 2005).

Seventeen factors—ten contextual, five personal, and two teacher thinking—were examined (see full model in Table 2); 10% of respondent-level, 68% of department-level, and 52% of institutional-level variances (analogous to R^2 in multiple regression) were accounted for in the full model. In addition, all the factors considered in the full model collectively explained 22% of the variance in the data; when all factors are accounted for in the full model, there is a medium to large global effect size with Cohen's $f^2 = 0.28$ (Cohen, 1992). The full model intercept (i.e., 82.66) represents the percent time lecturing reported by an instructor in a mathematics department that awards an associate degree as the highest degree (i.e., reference instructor) at zero (or the median value) for all other evaluated factors (see "Methods"). All multilevel regression coefficients for an individual factor are reported with all other factors held constant, with the exception of the single reported interaction effect between class size and classroom setup.

This model is visualized in Fig. 2; the regression is plotted with the intercept at 82.66 with multilevel regression coefficients below the intercept showing a decrease in percent lecturing and estimates above the intercept indicating an in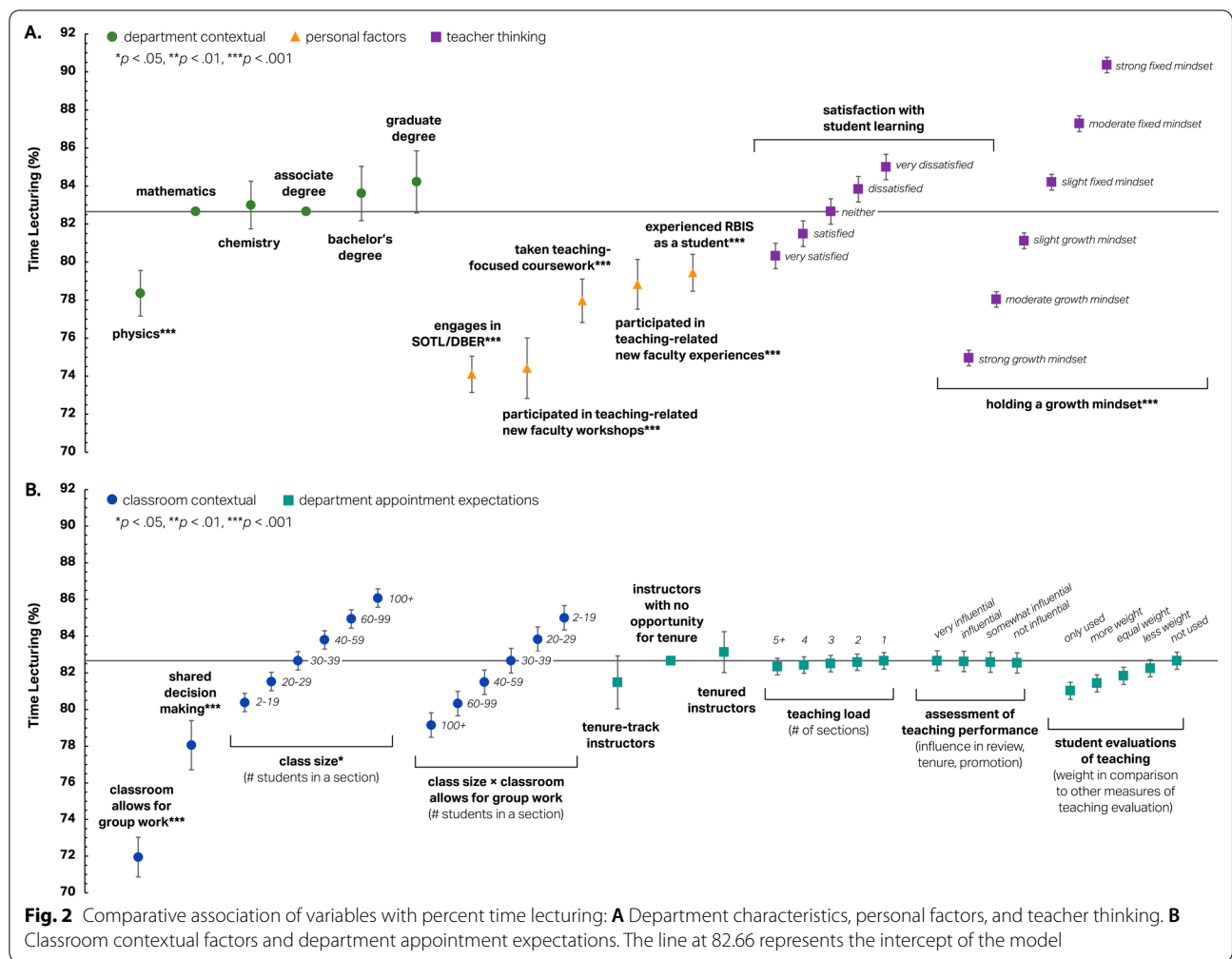crease in percent lecturing. For scaled factors, estimates for each scale point is shown. Multilevel regression coefficients indicate the strength of the relationship between a predictor variable and the outcome variable (i.e., percent lecturing) when all other predictor variables are accounted for and held constant.

Department characteristics factors

Department characteristics factors evaluated in this model include discipline and highest degree awarded by department.

Academic discipline (i.e., chemistry, mathematics, physics) is evaluated with mathematics as reference: Holding all other evaluated variables constant, percent time lecturing for instructors from chemistry departments ($\beta_j = 0.34$, $p > 0.05$, $f^2 < 0.01$) is not statistically different from instructors in mathematics departments; percent time lecturing for instructors from physics departments is 4.30% less ($p < 0.001$, $f^2 < 0.01$) than instructors from mathematics departments.

Highest degree awarded by the department is evaluated with an associate degree as the highest degree awarded as reference: Holding all other evaluated variables constant, average percent time lecturing for instructors in departments awarding as the highest degree bachelor's degrees ($\beta_j = 0.95$, $p > 0.05$, $f^2 < 0.01$) or graduate degrees ($\beta_j = 1.56$, $p > 0.05$, $f^2 < 0.01$) are not statistically different

Fig. 2 Comparative association of variables with percent time lecturing: **A** Department characteristics, personal factors, and teacher thinking. **B** Classroom contextual factors and department appointment expectations. The line at 82.66 represents the intercept of the model

from instructors in departments awarding associate degrees and its highest degree.

Department appointment expectations

Department appointment expectations evaluated in this model include teaching load, tenure status, and perceived professional review and reward structure factors.

Teaching load and tenure status are evaluated: Teaching load is non-significant, negatively associated with percent time lecturing ($\pi_j = -0.08$, $p > 0.05$, $f^2 < 0.01$). In addition, tenure-status is evaluated ("not in a tenure-track position" is the reference): Changes in percent time lecturing for instructors in a tenure-track position ($\pi_j = 0.47$, $p > 0.05$, $f^2 < 0.01$) or in a tenured position ($\pi_j = -1.18$, $p > 0.05$, $f^2 < 0.01$) are not statistically from instructors without opportunity for tenure.

Perceived role of assessment of teaching performance in review, promotion, and tenure is evaluated: A non-significant, negligible decrease in percent time lecturing ($\pi_j = -0.41$, $p > 0.05$, $f^2 < 0.01$) is associated with increases in the role of student evaluations of teaching in evaluating teaching performance. In addition, a non-significant, negligible decrease in percent time lecturing ($\pi_j = -0.04$, $p > 0.05$, $f^2 < 0.01$) is associated with an increase in the perceived role overall assessment of teaching performance matters in decisions of review, promotion, and tenure.

Classroom contextual factors

Classroom contextual factors evaluated in this model include class size, physical layout, and course administration (i.e., involvement in decision making).

Physical space, including maximum number of students for a given classroom, and configuration and type of furniture, are evaluated: A positive, small increase in percent time lecturing is associated with larger course sizes ($\pi_j = 1.14$, $p = 0.021$, $f^2 < 0.01$). A large decrease in percent time lecturing is associated with classroom spaces conducive to group work ($\pi_j = -10.71$, $p < 0.001$, $f^2 = 0.05$, small effect). There is a non-significant, small interaction effect between course size and classroom setup ($\pi_j = -1.17$, $p > 0.05$, $f^2 < 0.01$); this interaction effect essentially cancels out fluctuations in course size for classrooms spaces conducive to group work.

The primary decision maker for course instructional methods is evaluated: A significant decrease in percent time lecturing ($\pi_j = -4.60$, $p < 0.001$, $f^2 < 0.01$) is associated with courses in which decisions are made primarily by the instructor in conjunction with at least one additional instructor of the department.

Personal factors
Instructor's personal factors evaluated in this model include prior experience in courses taught with research-based instructional strategies, participation in SOTL or DBER, and also involvement in pedagogical professional development.

Experience as a student in courses taught with research-based instructional strategies is evaluated: Such experience is associated with a significant decrease in percent time lecturing ($\pi_j = -3.38$, $p < 0.01$, $f^2 < 0.01$).

Participation in SOTL or conducting DBER is evaluated: A significant decrease in percent time lecturing ($\pi_j = -8.56$, $p < 0.001$, $f^2 = 0.03$, small effect) is associated with such engagement.

Participation in teaching-related professional development experiences is evaluated: A significant decrease in percent time lecturing ($\pi_j = -3.22$, $p < 0.001$, $f^2 < 0.01$) is associated with participation in teaching-focused coursework at the undergraduate, graduate, or postdoctoral levels. A significant decrease in percent time lecturing ($\pi_j = -8.24$, $p < 0.001$, $f^2 < 0.01$) is associated with participation in teaching-focused workshops including half-day to multiple day workshops and attending teaching-focused conferences. A significant decrease in percent time lecturing ($\pi_j = -4.69$, $p < 0.001$, $f^2 < 0.01$) is associated with participation in teaching-related new faculty experiences internal or external to the respondent's institution.

Teacher thinking factors
Teacher thinking factors evaluated in this model include holding a growth mindset and satisfaction with student learning.

Growth mindset (i.e., the belief that ability can be developed) is evaluated: A significant decrease in percent time lecturing ($\pi_j = -3.08$, $p < 0.001$, $f^2 = 0.02$, small effect) is associated with an increase in growth mindset. Satisfaction with student learning is evaluated: A non-significant decrease in percent time lecturing ($\pi_j = -1.17$, $p > 0.05$, $f^2 < 0.01$) is associated with increased satisfaction with student learning.

Discussion
Six themes emerged from our multilevel regression model results that are associated with decreased percent time lecturing, at the instructor-level, when all other factors are held constant: (1) classroom spaces conducive to group work, (2) shared decision making on instructional methods, (3) participation as a student in courses utilizing research-based instructional strategies, (4) participation in teaching-related professional development experiences, (5) participation in scholarship of teaching and learning and discipline-based education research, and (6) espousing a growth mindset.

Only malleable factors are considered and interpreted in this Discussion as only these factors can lead to tangible implications for institutions and potential actions to support the adoption of more engaging instructional practices. In addition, these malleable factors have tangible implications for professional organizations and communities of practice to support transformation and reform efforts. Therefore, we do not consider department-level factors (e.g., discipline and highest degree awarded) as these cannot be changed; it is not practical for instructors to abandon their disciplinary training and switch to a new STEM field or take up a new position at a different institution for the sake of enacting more active learning practices.

Classroom spaces
Instructors consistently indicate that large class sizes, coupled with fixed-seating classroom layouts (i.e., lecture halls with bolted seats to the floor), are not conducive to interactions between students and make it difficult to implement research-based instructional strategies (Gess-Newsome et al., 2003; Henderson & Dancy, 2007; Hora, 2012; Lund & Stains, 2015). Studies routinely show that smaller-sized classes held in spaces that allow for active learning (e.g., moveable tables or desks) are associated with implementation of more student-centered teaching methods (Cotner et al., 2013; Lund & Stains, 2015; Shadle et al., 2017). Our findings corroborate these studies indicating that class size and classroom spaces matter, with the latter having a large, significant association with decreased percent time lecturing.

These findings beg the question: "If we build it, will they come?" Or more pointedly, "If classrooms spaces are built or renovated to be more conducive to group work, will instructors implement more active-learning pedagogies in courses taught in such spaces?" The underlying ambiguity is: *what is the cause and what is the effect?* Our results can be interpreted that the space catalyzes implementation of non-lecture-based pedagogies. Some evidence indicates that building a classroom for active learning increases the like likelihood for such intended purposes, because it motivates and encourages instructors to try active learning pedagogies (Foote et al., 2016). Conversely, our results could be interpreted in such a way that instructors wanting to use pedagogies that are more easily implemented in classroom spaces conducive to group work seek out and request to teach in such spaces. Having a decided active learning classroom increases the sustainability of active learning implementations as it takes effort to undue significant structural change because of the buy in and support of many individuals, including instructors and administrators (Knaub et al., 2016). Both causal explanations are only possible if such classroom spaces exist and are available. Regardless, we argue that such classrooms spaces should be advocated for and built regardless of the cause−effect relationship.

There are resources readily available for designing classroom spaces that promote implementation of active learning pedagogies: One such example is SCALE-UP (Student-Centered Active Learning Environment with Upside-down Pedagogies) which aims to reform teaching practices by manufacturing physical changes to the classroom layout that in turn minimize lecture (Knaub et al., 2016; SCALE-UP, 2011). SCALE-UP has demonstrated improvements in students' problem-solving abilities, conceptual understanding, attitudes toward science, retention in introductory courses, and later performance in subsequent courses in chemistry, mathematics, and physics; studies have also shown a reduction in failure rates, especially for women and minoritized students in SCALE-UP classrooms (Beichner, 2008; Beichner et al., 2007). In addition, FLEXspace® (Flexible Learning Environments Exchange space) is a tool that supports communities of experts, practitioners, and institutional decision makers to improve active learning space planning, design, and implementation (FLEXspace, 2018).

Shared decision on instructional methods

Introductory courses in chemistry, mathematics, and physics are typically large enrollment courses, even at smaller sized institutions (Koch, 2017; Seymour & Hewitt, 1997; Seymour & Hunter, 2019). We define "large" to be considered relative to other upper level courses within a given institution. Irrespective of institution size, though, typically large enrollment courses are divided into smaller classes (i.e., sections). It is common for more than one instructor to be teaching the set of sections for a given course. The degree of coordination of these sections can vary from complete independence (including different textbooks and syllabi) to complete coordination, wherein all aspects of the course are common across sections including examinations and instructional practices (Apkarian & Kirin, 2017). Decision making authority for aspects of the course may lie with the individual instructor or with a committee. Our findings suggest that when an instructor shares decision making authority with one or more instructors on the instructional methods used, that such coordination is associated with less time lecturing.

Decisions on teaching methods are best when done as a team. A recent study revealed that instructors at three research-intensive universities who use innovative teaching practices preferentially interact with other users due to their similar teaching values (Lane et al., 2020). As Lane et al. (2020) note, co-teaching and teaching teams can encourage the interaction between instructors of different teaching practices (Henderson et al., 2009). Instructors that have experience in implementing active learning strategies should collaborate or co-teach with non-users to assist in the uptake of these pedagogies (Gess-Newsome et al., 2003; Henderson & Dancy, 2007; Henderson et al., 2009). Instructors state that it is easier to integrate research-based methods if other instructors are implementing new methods at the same time (Foote et al., 2016; Henderson & Dancy, 2007); implementation requires supportive departmental and institutional mentors (Henderson & Dancy, 2007; Shadle et al., 2017). Coordinating a course across multiple class sections is a possible route to execute this change and offers an opportunity for instructional support outside of formal avenues of professional development (Apkarian et al., 2019; Golnabi et al., 2021; Henderson et al., 2009; Martinez & Pilgrim, 2021; Williams et al., 2021).

We advocate for course coordination to encourage the exchange of ideas and experience among instructors as means to reduce the time constraints and uncertainties of implementing new instructional practices (Shadle et al., 2017). As lack of time has been previously noted as a barrier for the implementation of new teaching pedagogies (e.g., Brownell & Tanner, 2012; Henderson & Dancy, 2007; Lund & Stains, 2015), the appointment of course coordinators with long-term roles would provide necessary and ample time to support implementation of active learning pedagogies (Rasmussen & Ellis, 2015). In addition, co-teaching can be used in cases of a two-section course, where one instructor is experienced in

Yik *et al. International Journal of STEM Education*   (2022) 9:15

Page 15 of 23

the implementation of active learning strategies (Henderson et al., 2009). Course coordinators can discuss RBIS implementation with instructors in the teaching team (Golnabi et al., 2021) and others at department meetings (Bazett & Clough, 2021; Williams et al., 2021) to facilitate conversation around active learning uptake.

Experience as a student in a course using research-based instructional strategies

It has been reported that teaching methods experienced by STEM instructors when they were students influence their current teaching practices (Fukawa-Connelly et al., 2016; Lund & Stains, 2015; Oleson & Hora, 2014). Prior classroom experiences as an instructor, professional development programs, and interactions with other instructors are also influential in teaching pedagogies (Oleson & Hora, 2014). Our results corroborate these findings with experiences of RBIS as a student aligning with lower reported percent time lecturing.

A continual cycle of future instructors (i.e., current undergraduate and graduate students) using active learning strategies can pave the way for an effective, longitudinal, and sustainable method for implementing pedagogical reform supported the results of this study and the notion that previous experiences a student influence present teaching pedagogies practices (Fukawa-Connelly et al., 2016; Lund & Stains, 2015; Oleson & Hora, 2014). For this to succeed though, current instructors need to be made aware of and be willing to implement RBIS. Active learning pedagogies used in today's classrooms will influence future instructors.

Participation in teaching-related professional development

Our findings show that instructors who have taken teaching-focused coursework, participated in teaching-related workshops, or teaching-related new faculty experiences report a lower percentage time lecturing than instructors who have not engaged in these opportunities. Centers for teaching and learning and professional organizations (e.g., the American Chemical Society, the Mathematical Association of America, and the American Physical Society) provide opportunities for instructors, in addition to graduate students and postdoctoral scholars, to participate in professional development workshops. Teaching-focused workshops are shown to be effective in informing teaching decisions (Henderson et al., 2011; Oleson & Hora, 2014). In addition, some centers for teaching and learning now provide structured for-credit teaching-focused coursework independently or in collaboration with colleges of education. Topics of these professional development opportunities range from discipline-specific pedagogical training to teaching and learning theory,

effective pedagogical practices, and design of instructional materials (cf. Coppola, 2016; Gardner & Jones, 2011; Wheeler et al., 2017; Wyse et al., 2014). Although few instructors receive pedagogical training as a part of their graduate programs, STEM instructors who have received training were found to more likely to have referenced sources of instructional innovation (Walczyk et al., 2007). While most professional development programs focus on new or early-career faculty and instructors, programs should also be developed for and tailored to those at mid-career or late-stages (Austin & Sorcinelli, 2013), and institutionalized and sustained professional development is necessary for lasting pedagogical change (Borda et al., 2020; Henderson et al., 2011).

In addition to pedagogical training, centers for teaching and learning can provide practical assistance to all instructors in their active learning endeavors. These centers can also sponsor communities of practice for instructors to discuss their teaching practices and instructional projects, which have also been shown to increase use of RBIS and transfer knowledge between disciplines to enhance instruction (Benabentos et al., 2020; Dancy et al., 2019; Henderson et al., 2017; Pelletreau et al., 2018; Tomkin et al., 2019); for adopters of active learning pedagogies, participation in a community of practice has been shown to have greater use of student-centered practices (Benabentos et al., 2020). If these centers for teaching and learning do exist, we propose that centers are given sufficient resources to achieve their goals. A common reason why these centers are under-utilized is because center staff, while experts in education, may not have broad disciplinary expertise; to increase credibility, centers for teaching and learning should incorporate more discipline-specific skills training and hire more persons with a broad spectrum of DBER experience (Pelletreau et al., 2018; Seymour et al., 2011). At institutions, where centers are not financially feasible, peer-coaching may be a useful form of professional development (Desimone & Pak, 2017); instructors can observe one another's teaching and provide feedback and discuss teaching methods (Gormally et al., 2014; Smith et al., 2014).

Dissemination of RBIS through new faculty experiences has been shown to increase instructor awareness and inclusion of these instructional strategies in biology (Derting et al., 2016; Ebert-May et al., 2015; Handelsman et al., 2004; Wood & Gentile, 2003), chemistry (Baker et al., 2014; Houseknecht et al., 2020; Murray et al., 2011; Stains et al., 2015; Stegall et al., 2016), and physics (Henderson, 2008; Henderson et al., 2012). Institutions may require, or highly incentivize, participation in these professional development experiences for all new instructors, or through tenure and promotion requirements (Bathgate et al., 2019a, 2019b; Seymour et al., 2011).

Institutions can also help sustain implementation of RBIS through, at the very minimum highly incentive, programs designed for mid- and late-stage instructors (Austin & Sorcinelli, 2013; Borda et al., 2020).

Scholarship of teaching and learning or discipline-based education research

Conducting or participating in SOTL or DBER had the largest impact on reducing the percent time teaching of all non-contextual factors in our study. In an analytical review of literature on undergraduate STEM practices, Henderson et al. (2011) noted that engagement in SOTL is a means for developing reflective educators. Instructors that engaged in STEM education research were found to use more student-centered instructional practices (Benabentos et al., 2020; Dancy et al., 2019; Henderson et al., 2017; Pelletreau et al., 2018; Tomkin et al., 2019). This is corroborated by our findings that participation in SOTL is associated with adoption of more active learning pedagogies. Thus, instructors should be encouraged to engage and be recognized for SOTL, and by extension, DBER work (Henderson et al., 2012). Institutions should encourage and reward instructors for their efforts in SOTL and treat it as a valuable scholarly outlet (Walczyk et al., 2007). Collaborative projects with education scholars, both discipline-based and within colleges of education, can serve to catalyze purposeful investigation of teaching and learning in STEM courses (Dancy et al., 2019; Oleson & Hora, 2014; Shadle et al., 2017). In addition, collaborative SOTL projects between instructors and graduate students enrolled in future faculty programs provide an additional pathway for involving current and future educators in teaching-oriented scholarship (e.g., Coppola, 2016), which would give both short- and long-term benefits to individuals and the field.

Holding a growth mindset

Our results suggest that growth mindset beliefs are associated with a reduction in the amount of time spent lecturing. Instructors of different identities (i.e., gender and race/ethnicity) and experiences (i.e., teaching experience and tenure status) across multiple STEM disciplines have been found to espouse fixed mindsets (Canning et al., 2019). This is problematic such that fixed mindsets may make stereotype threats more evident and concerning; stereotyped stigmatized students have been shown to experience more anxiety, lower sense of belonging, and become less interested (Bian et al., 2018; Emerson & Murphy, 2015). In addition, instructors holding fixed mindsets may inhibit the pursual of graduate-level education by women and minoritized students (Leslie et al., 2015). Fixed mindsets are malleable and large-scale studies have shown that more of a growth mindset can be

developed (e.g., Broda et al., 2018; Yeager & Dweck, 2012; Yeager, Romero, et al., 2016; Yeager, Walton, et al., 2016), although do not yet know the effect that growth mindset interventions have on the implementation of active learning (Aragón et al., 2018).

Studies have shown that instructors' growth mindsets have the potential to improve student learning as well as address equity in the classroom (e.g., Canning et al., 2019; Gasiewski et al., 2012; Leslie et al., 2015; T. Smith et al., 2018). Safe environments promote engagement; students felt more comfortable asking questions when instructors held a growth mindset, which is exemplified by their feeling that a too rudimentary question was nonexistent (Gasiewski et al., 2012). Instructor feedback can communicate their mindset beliefs; students that received growth mindset comments themselves moved toward more growth mindset beliefs and scored higher on a summative assessment than their counterparts who received fixed mindset comments (Smith et al., 2018). This is advantageous as growth mindsets have been associated with higher achievement in students (e.g., Blackwell et al., 2007; Burnette et al., 2013; Yeager & Dweck, 2012; Yeager et al., 2019). Thus, professional development activities including dissemination of active learning pedagogies, teaching assistant training, and new faculty experiences should include a component on beliefs about teaching and learning. These training and experiences should aim to make instructors aware of how their mindset influences the culture in their classes and student motivation and achievement (Canning et al., 2019).

Non-significant factors

Several factors are not statistically significant in explaining variability in percent time lecturing. These factors include teaching load, tenure status, student evaluations of teaching, assessment of teaching performance, and satisfaction with student learning.

The insignificance of teaching load and tenure status potentially indicates that instructors have a high work expectation that is irrespective of the division of time between teaching and research; in other words, instructors are busy. We know that current incentives are lacking and there is a limited focus on teaching in annual and merit reviews for faculty members to enact research-based instructional strategies (Lund & Stains, 2015; Shadle et al., 2017). When the association between percent time lecturing and the importance of student evaluations of teaching or the role of teaching performance in reviews are considered without controlling for other predictors, a significant, but small association has been found (Apkarian et al., 2021); this suggests that there may be a more nuanced relationship between instructional practices and evaluations of teaching performance than

what is reported in the study reported herein. Nonetheless, departmental and institutional pressures or incentives can become more influential if a larger emphasis is placed on the assessment of teaching performance in annual, tenure, or promotion evaluations (Lund & Stains, 2015).

Finally, decreased satisfaction with student learning has been shown in other studies to be associated with adoption of active learning strategies, with dissatisfaction being a central tenant to the TCSR model that frames our study and analyses (Gess-Newsome et al., 2003). However, we acknowledge that the relationship between this factor and percent time lecturing may be more complex than what we have modeled in our study given the array of modeled factors. One way to interpret this is that instructors who lecture less are happier with their student outcomes; this would be consistent with dissatisfaction leading to change.

Limitations
This study has a limited scope; we only sampled instructors who teach introductory courses from chemistry, mathematics, and physics. In addition, previous studies show differences exist between lower- and upper-division STEM courses (Benabentos et al., 2020; Lund et al., 2015). To better characterize STEM courses as a whole, a broader study should be conducted that includes a wider array of STEM disciplines at both introductory and advanced levels from a representative sample of different institution types.

The TCSR model was originally developed to understand reform in K-12 education (Woodbury & Gess-Newsome, 2002) and was then adapted for a college classroom (Gess-Newsome et al., 2003). Consequently, due to our review of the literature, our conceptualization of the TCSR framework required further modification of the framework to delineate department appointment expectations apart from departmental contextual factors to account for the complex nature of different instructional positions in higher education. In addition, it should be noted that the TCSR framework has not been previously used in multilevel analyses; therefore, some of the considerations of testing a statistical model such as the nested nature of the data (i.e., instructors within departments within institutions) have not been previously addressed. While a modification of the model (i.e., an articulated differentiation of appointment expectations) was necessary, this does not invalidate the TCSR model, but furthers theoretical and empirical possibilities for using the model to evaluate teaching practices.

The self-reported nature of our outcome measure (i.e., percent time lecturing) and a respondent's contextual, personal, and beliefs about teaching and learning factors result in some loss of empirical strength due to potential reliability threats and may potentially mischaracterize the complex nature of the classroom. Discrepancies between self-reported data and researcher-evaluated observations of observed classroom practices have previously been found (e.g., Bodzin & Beerer, 2003; Herrington et al., 2016; Koziol & Burns, 1986). Although, observational studies are prone to observer subjectivity, particularly in regard to rater agreement (Hill et al., 2012; Waxman & Padrón, 2004). Low numbers of observations, insufficient training of raters, and non-representative snapshots of instructional practices can raise uncertainty of observation data (Cohen & Goldhaber, 2016; Hill et al., 2012). However, evidence suggests that self-reported data about teaching practices align well with observational studies (Durham et al., 2018; Gibbons et al., 2018; Hayward et al., 2018). While large-scale studies are necessary to identify teaching practices (Williams et al., 2015), we note that observational studies, though, would parallel self-report data but is impractical with the large sample of our study (i.e., $N = 2382$). Balancing error associated with self-report data and the opportunity to conduct large-scale study, such as ours, supported by the work of others (Durham et al., 2018; Gibbons et al., 2018; Hayward et al., 2018), the results of the study herein are trustworthy.

Conclusions
Based on these results from a national survey of gateway chemistry, mathematics, and physics instructors that considers a large number of factors associated with uptake of active learning and accounts for the nested nature of institutional contexts, we provide four broad recommendations for sustaining active learning strategies in introductory STEM courses:

1. Construct classroom spaces that support and promote active learning (i.e., moveable table/desks for shared group work and activities, whiteboards to support collaboration, etc.). Provide and incentivize professional development to assist instructors in maximizing the use of active learning spaces.
2. Coordinate large enrollment courses with multiple course sections. Collaborate with other instructors on instructional methods, allowing for discussion and reflection on instructional practices.
3. Offer and encourage participation in professional development programs and communities of practice for widespread awareness and implementation of research-based instructional strategies. Promote a growth mindset and develop constructive beliefs about teaching and learning in professional development opportunities.

4. Engage in the scholarship of teaching and learning. Recognize and value SOTL work accomplished by instructors.

Our results demonstrate to instructors, departmental, and institutional leaders on how contextual, personal, and teacher thinking factors are associated with decisions about instructional practices at their institutions. The research literature suggests that instructor-centered reforms and instructor-, departmental-, and institutional-led reform initiatives are essential to meaningful and sustained cultural change about how STEM courses are taught (Beichner et al., 2007; Henderson et al., 2011; Reinholz & Apkarian, 2018; Shadle et al., 2017).

Abbreviations
ATP: Assessment of teaching performance; DBER: Discipline-based educational research; FLEXspace: Flexible learning environments exchange space; ICC: Intraclass correlation coefficient; NeXT: New experiences in teaching; NFE: New faculty experience; POGIL: Process oriented guided inquiry learning; RBIS: Research-based instructional strategies; SCALE-UP: Student-centered active learning environment with upside-down pedagogies; SET: Student evaluations of teaching; SOTL: Scholarship of teaching and learning; STEM: Science, technology, engineering, and mathematics; TFC: Teaching-focused coursework; TCSR: Teacher-centered systemic reform; VITAL: Visitors, instructors, teaching assistants, adjuncts, and lecturers; WKSP: Workshop.

Supplementary Information
The online version contains supplementary material available at https://doi.org/10.1186/s40594-022-00333-3.

Additional file 1. Supplementary Table 1. Survey items. **Supplementary Table 2.** Descriptive statistics for level 1 variables (instructors; n = 2,382). Supplementary Table 3. Descriptive statistics for level 2 variables (department; n = 1,405). **Supplementary Table 4.** Correlations among the variables at level 1 (instructors; n = 2,382). **Supplementary Table 5.** Correlations among the variables at level 2 (department; n = 1,405). **Supplementary Table 6.** Variance inflation factors (VIF) on the standardized model.

Availability of data and materials
The data sets used and/or analyzed during the current study are 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 Chemistry, University of South Florida, Tampa, FL 33620-5250, USA. [2]School of Mathematical and Statistical Sciences, Arizona State University, Tempe, AZ 85287-1804, USA. [3]Department of Chemistry, University of Virginia, Charlottesville, VA 22904-4319, USA. [4]Department of Physics and Mallinson Institute for Science Education, Western Michigan University, Kalamazoo, MI 49008-5252, USA. [5]Department of Physics and Center for STEM Learning, University of Colorado, Boulder, CO 80309-0390, USA. [6]Department of Mathematics, Virginia Tech, Blacksburg, VA 24061-0123, USA.

References
Andrews, T. C., & Lemons, P. P. (2015). It's personal: biology instructors prioritize personal evidence over empirical evidence in teaching decisions. *CBE Life Sciences Education, 14*(1), ar7. 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), e0247544. https://doi.org/10.1371/journal.pone.0247544

Apkarian, N., & Kirin, D. (2017). *Progress through calculus: Census survey technical report*. Mathematical Association of America. https://www.maa.org/sites/default/files/PtC%20Technical%20Report_Final.pdf

Apkarian, N., & Rasmussen, C. (2017). Mathematics instruction leadership in undergraduate departments. In A. Weinberg, C. Rasmussen, J. Rabin, M. Wawro, & S. Brown, Proceedings of the 20th Annual Conference on Research in Undergraduate Mathematics Education. http://sigmaa.maa.org/rume/RUME20.pdf

Apkarian, N., Smith, W. M., Vroom, K., Voigt, M., Gehrtz, J., Team, P. P., & Team, S. P. (2019). *X-PIPS-M survey suite*. 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), ar47. https://doi.org/10.1187/cbe.17-05-0084

Armbruster, P., Patel, M., Johnson, E., & Weiss, M. (2009). Active learning and student-centered pedagogy improve student attitudes and performance in introductory biology. *CBE Life Sciences Education, 8*(3), 203–213. https://doi.org/10.1187/cbe.09-03-0025

Austin, A. E., & Sorcinelli, M. D. (2013). The future of faculty development: where are we going? *New Directions for Teaching and Learning, 2013*(133), 85–97. https://doi.org/10.1002/tl.20048

Bahník, Š, & Vranka, M. A. (2017). Growth mindset is not associated with scholastic aptitude in a large sample of university applicants. *Personality and Individual Differences, 117*, 139–143. https://doi.org/10.1016/j.paid.2017.05.046

Baker, L. A., Chakraverty, D., Columbus, L., Feig, A. L., Jenks, W. S., Pilarz, M., Stains, M., Waterman, R., & Wesemann, J. L. (2014). Cottrell Scholars Collaborative New Faculty Workshop: professional development for new chemistry faculty and initial assessment of its efficacy. *Journal of Chemical Education, 91*(11), 1874–1881. https://doi.org/10.1021/ed500547n

Ballen, C. J., Wieman, C., Salehi, S., Searle, J. B., & Zamudio, K. R. (2017). Enhancing diversity in undergraduate science: self-efficacy drives performance gains with active learning. *CBE Life Sciences Education, 16*(4), ar56. https://doi.org/10.1187/cbe.16-12-0344

Bates, D., Mächler, M., Bolker, B. M., & Walker, S. C. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Bathgate, M. E., Aragón, O. R., Cavanagh, A. J., Frederick, J., & Graham, M. J. (2019a). Supports: a key factor in faculty implementation of

evidence-based teaching. *CBE Life Sciences Education, 18*(2), ar22. https://doi.org/10.1187/cbe.17-12-0272

Bathgate, M. E., Aragón, O. R., Cavanagh, A. J., Waterhouse, J. K., Frederick, J., & Graham, M. J. (2019b). Perceived supports and evidence-based teaching in college STEM. *International Journal of STEM Education, 6*(1), 11. https://doi.org/10.1186/s40594-019-0166-3

Bauer, C., Libby, R. D., Scharberg, M., & Reider, D. (2013). Transformative research-based pedagogy workshops for chemistry graduate students and postdocs. *Journal of College Science Teaching, 43*(2), 36–43.

Bazett, T., & Clough, C. L. (2021). Course coordination as an avenue to departmental culture change. *Primus, 31*(3–5), 467–482. https://doi.org/10.1080/10511970.2020.1793853

Beatty, I. D., Gerace, W. J., Leonard, W. J., & Dufresne, R. J. (2005). Designing effective questions for classroom response system teaching. *American Journal of Physics, 74*(1), 31–39. https://doi.org/10.1119/1.2121753

Beichner, R. J. (2008). *The SCALE-UP project: A student-centered active learning environment for undergraduate programs*. National Academy of Sciences. https://sites.nationalacademies.org/cs/groups/dbassesite/documents/webpage/dbasse_072628.pdf

Beichner, R. J., Saul, J. M., Abbott, D. S., Morse, J. J., Deardorff, D. L., Allain, R. J., Bonham, S. W., Dancy, M. H., & Risley, J. S. (2007). *The Student-Centered Activities for Large Enrollment Undergraduate Programs (SCALE-UP) project* (Vol. 1). American Association of Physics Teachers. http://www.per-central.org/document/ServeFile.cfm?ID=4517

Benabentos, R., Hazari, Z., Stanford, J. S., Potvin, G., Marsteller, P., Thompson, K. V., Cassone, V. M., Murasko, D., & Kramer, L. (2020). Measuring the implementation of student-centered teaching strategies in lower- and upper-division STEM courses. *Journal of Geoscience Education, 69*(4), 342–356. https://doi.org/10.1080/10899995.2020.1768005

Bettinger, E., Ludvigsen, S., Rege, M., Solli, I. F., & Yeager, D. (2018). Increasing perseverance in math: evidence from a field experiment in Norway. *Journal of Economic Behavior & Organization, 146*, 1–15. https://doi.org/10.1016/j.jebo.2017.11.032

Bian, L., Leslie, S.-J., Murphy, M. C., & Cimpian, A. (2018). Messages about brilliance undermine women's interest in educational and professional opportunities. *Journal of Experimental Social Psychology, 76*, 404–420. https://doi.org/10.1016/j.jesp.2017.11.006

Blackwell, L. S., Trzesniewski, K. H., & Dweck, C. S. (2007). Implicit theories of intelligence predict achievement across an adolescent transition: a longitudinal study and an intervention. *Child Development, 78*(1), 246–263. https://doi.org/10.1111/j.1467-8624.2007.00995.x

Bodzin, A. M., & Beerer, K. M. (2003). Promoting inquiry-based science instruction: the validation of the science teacher inquiry rubric (stir). *Journal of Elementary Science Education, 15*(2), 39. https://doi.org/10.1007/BF03173842

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), 4. https://doi.org/10.1186/s40594-020-0203-2

Bressoud, D., & Rasmussen, C. (2015). Seven characteristics of successful calculus programs. *Notices of the American Mathematical Society, 62*(2), 144–146. https://doi.org/10.1090/noti1209

Broda, M., Yun, J., Schneider, B., Yeager, D. S., Walton, G. M., & Diemer, M. (2018). Reducing inequality in academic success for incoming college students: a randomized trial of growth mindset and belonging interventions. *Journal of Research on Educational Effectiveness, 11*(3), 317–338. https://doi.org/10.1080/19345747.2018.1429037

Brownell, S. E., & Tanner, K. D. (2012). Barriers to faculty pedagogical change: lack of training, time, incentives, and tensions with professional identity? *CBE Life Sciences Education, 11*(4), 339–346. https://doi.org/10.1187/cbe.12-09-0163

Burke, K. A., Greenbowe, T. J., & Gelder, J. I. (2004). The multi-initiative dissemination project workshops: who attends them and how effective are they? *Journal of Chemical Education, 81*(6), 897. https://doi.org/10.1021/ed081p897

Burnette, J. L., O'Boyle, E. H., VanEpps, E. M., Pollack, J. M., & Finkel, E. J. (2013). Mind-sets matter: a meta-analytic review of implicit theories and self-regulation. *Psychological Bulletin, 139*(3), 655–701. https://doi.org/10.1037/a0029531

Canning, E. A., Muenks, K., Green, D. J., & Murphy, M. C. (2019). STEM faculty who believe ability is fixed have larger racial achievement gaps and inspire less student motivation in their classes. *Science Advances, 5*(2), eaau4734. https://doi.org/10.1126/sciadv.aau4734

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates Inc. https://doi.org/10.4324/9780203771587

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112*(1), 155–159. https://doi.org/10.1037/0033-2909.112.1.155

Cohen, J., & Goldhaber, D. (2016). Building a more complete understanding of teacher evaluation using classroom observations. *Educational Researcher, 45*(6), 378–387. https://doi.org/10.3102/0013189X16659442

Common Data Set Initiative. (n.d.). Retrieved 6 July 2021 from https://commondataset.org

Coppola, B. P. (2016). Broad & capacious: a new norm for instructional development in a research setting. *Change the Magazine of Higher Learning, 48*(2), 34–43. https://doi.org/10.1080/00091383.2016.1163206

Cotner, S., Loper, J., Walker, J. D., & Brooks, D. C. (2013). "It's not you, it's the room"—are the high-tech, active learning classrooms worth it? *Journal of College Science Teaching, 42*(6), 82–88.

Cox, B. E., McIntosh, K. L., Reason, R. D., & Terenzini, P. T. (2011). A culture of teaching: policy, perception, and practice in higher education. *Research in Higher Education, 52*(8), 808–829. https://doi.org/10.1007/s11162-011-9223-6

Dancy, M., Lau, A. C., Rundquist, A., & Henderson, C. (2019). Faculty online learning communities: a model for sustained teaching transformation. *Physical Review Physics Education Research, 15*(2), 020147. https://doi.org/10.1103/PhysRevPhysEducRes.15.020147

Derting, T. L., Ebert-May, D., Henkel, T. P., Maher, J. M., Arnold, B., & Passmore, H. A. (2016). Assessing faculty professional development in STEM higher education: sustainability of outcomes. *Science Advances, 2*(3), e1501422. https://doi.org/10.1126/sciadv.1501422

Desimone, L. M., & Pak, K. (2017). Instructional coaching as high-quality professional development. *Theory into Practice, 56*(1), 3–12. https://doi.org/10.1080/00405841.2016.1241947

Drinkwater, M. J., Gannaway, D., Sheppard, K., Davis, M. J., Wegener, M. J., Bowen, W. P., & Corney, J. F. (2014). Managing active learning processes in large first year physics classes: the advantages of an integrated approach. *Teaching & Learning Inquiry THe ISSOTL Journal, 2*(2), 75–90. https://doi.org/10.2979/teachlearninqu.2.2.75

Durham, M. F., Knight, J. K., Bremers, E. K., DeFreece, J. D., Paine, A. R., & Couch, B. A. (2018). Student, instructor, and observer agreement regarding frequencies of scientific teaching practices using the Measurement Instrument for Scientific Teaching-Observable (MISTO). *International Journal of STEM Education, 5*, 31. https://doi.org/10.1186/s40594-018-0128-1

Dweck, C. S. (1999). *Self-theories: their role in motivation, personality, and development*. Psychology Press.

Dweck, C. S. (2019). The choice to make a difference. *Perspectives on Psychological Science, 14*(1), 21–25. https://doi.org/10.1177/1745691618804180

Dweck, C. S., Chiu, C.-Y., & Hong, Y.-Y. (1995). Implicit theories and their role in judgments and reactions: a word from two perspectives. *Psychological Inquiry, 6*(4), 267–285. https://doi.org/10.1207/s15327965pli0604_1

Dweck, C. S., & Leggett, E. L. (1988). A social-cognitive approach to motivation and personality. *Psychological Review, 95*(2), 256–273. https://doi.org/10.1037/0033-295X.95.2.256

Ebert-May, D., Derting, T. L., Henkel, T. P., Middlemis Maher, J., Momsen, J. L., Arnold, B., & Passmore, H. A. (2015). Breaking the cycle: future faculty begin teaching with learner-centered strategies after professional development. *CBE Life Sciences Education, 14*(2), ar22. https://doi.org/10.1187/cbe.14-12-0222

Emerson, K. T. U., & Murphy, M. C. (2014). A company I can trust? Organizational lay theories moderate stereotype threat for women. *Personality and Social Psychology Bulletin, 41*(2), 295–307. https://doi.org/10.1177/0146167214564969

Erdmann, R., Miller, K., & Stains, M. (2020). Exploring STEM postsecondary instructors' accounts of instructional planning and revisions. *International Journal of STEM Education, 7*(1), 7. https://doi.org/10.1186/s40594-020-00206-7

Fairweather, J. (2008). *Linking evidence and promising practices in science, technology, engineering, and mathematics (STEM) undergraduate education*.

https://sites.nationalacademies.org/cs/groups/dbassesite/documents/webpage_dbasse_072637.pdf

Fairweather, J. S., & Rhoads, R. A. (1995). Teaching and the faculty role: enhancing the commitment to instruction in American colleges and universities. *Educational Evaluation and Policy Analysis, 17*(2), 179–194. https://doi.org/10.3102/01623737017002179

Feldman, A. (2000). Decision making in the practical domain: a model of practical conceptual change. *Science Education, 84*(5), 606–623.

Ferrare, J. J. (2019). A multi-institutional analysis of instructional beliefs and practices in gateway courses to the sciences. *CBE Life Sciences Education, 18*(2), ar26. https://doi.org/10.1187/cbe.17-12-0257

FLEXspace. (2018). Retrieved 1 June 2020 from https://flexspace.org

Foote, K., Knaub, A., Henderson, C., Dancy, M., & Beichner, R. J. (2016). Enabling and challenging factors in institutional reform: the case of SCALE-UP. *Physical Review Physics Education Research, 12*(1), 010103. https://doi.org/10.1103/PhysRevPhysEducRes.12.010103

Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., & Wenderoth, M. P. (2014). Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the National Academy of Sciences, 111*(23), 8410–8415. https://doi.org/10.1073/pnas.1319030111

Fukawa-Connelly, T., Johnson, E., & Keller, R. (2016). Can math education research improve the teaching of abstract algebra? *Notices of the American Mathematical Society, 63*(3), 276–281. https://doi.org/10.1090/noti1339

Gardner, G. E., & Jones, M. G. (2011). Pedagogical preparation of the science graduate teaching assistant: challenges and implications. *Science Educator, 20*(2), 31–41.

Gasiewski, J. A., Eagan, M. K., Garcia, G. A., Hurtado, S., & Chang, M. J. (2012). From gatekeeping to engagement: a multicontextual, mixed method study of student academic engagement in introductory STEM courses. *Research in Higher Education, 53*(2), 229–261. https://doi.org/10.1007/s11162-011-9247-y

Gess-Newsome, J., Southerland, S. A., Johnston, A., & Woodbury, S. (2003). Educational reform, personal practical theories, and dissatisfaction: the anatomy of change in college science teaching. *American Educational Research Journal, 40*(3), 731–767. https://doi.org/10.3102/00028312040003731

Gibbons, R. E., Villafañe, S. M., Stains, M., Murphy, K. L., & Raker, J. R. (2018). Beliefs about learning and enacted instructional practices: an investigation in postsecondary chemistry education. *Journal of Research in Science Teaching, 55*(8), 1111–1133. https://doi.org/10.1002/tea.21444

Golnabi, A. H., Murray, E., & Su, H. (2021). How precalculus course coordination can impact students' academic performance. *Journal of Higher Education Theory and Practice.* https://doi.org/10.33423/jhetp.v21i5.4278

Gormally, C., Evans, M., & Brickman, P. (2014). Feedback about teaching in higher ed: neglected opportunities to promote change. *CBE Life Sciences Education, 13*(2), 187–199. https://doi.org/10.1187/cbe.13-12-0235

Haak, D. C., HilleRisLambers, J., Pitre, E., & Freeman, S. (2011). Increased structure and active learning reduce the achievement gap in introductory biology. *Science, 332*(6034), 1213. https://doi.org/10.1126/science.1204820

Handelsman, J., Ebert-May, D., Beichner, R., Bruns, P., Chang, A., DeHaan, R., Gentile, J., Lauffer, S., Stewart, J., Tilghman, S. M., & Wood, W. B. (2004). Scientific teaching. *Science, 304*(5670), 521. https://doi.org/10.1126/science.1096022

Harris, R. B., Mack, M. R., Bryant, J., Theobald, E. J., & Freeman, S. (2020). Reducing achievement gaps in undergraduate general chemistry could lift underrepresented students into a "hyperpersistent zone." *Science Advances, 6*(24), eaaz5687. https://doi.org/10.1126/sciadv.aaz5687

Hayward, C. N., Weston, T., & Laursen, S. L. (2018). First results from a validation study of TAMI: Toolkit for Assessing Mathematics Instruction. In A. Weinberg, C. Rasmussen, J. Rabin, M. Wawro, & S. Brown, Proceedings of the 21st Annual Conference on Research in Undergraduate Mathematics Education. http://sigmaa.maa.org/rume/RUME21.pdf

Henderson, C. (2008). Promoting instructional change in new faculty: an evaluation of the physics and astronomy new faculty workshop. *American Journal of Physics, 76*(2), 179–187. https://doi.org/10.1119/1.2820393

Henderson, C., Beach, A., & Famiano, M. (2009). Promoting instructional change via co-teaching. *American Journal of Physics, 77*(3), 274–283. https://doi.org/10.1119/1.3033744

Henderson, C., Beach, A., & Finkelstein, N. (2011). Facilitating change in undergraduate STEM instructional practices: an analytic review of the literature. *Journal of Research in Science Teaching, 48*(8), 952–984. https://doi.org/10.1002/tea.20439

Henderson, C., Connolly, M., Dolan, E. L., Finkelstein, N., Franklin, S., Malcom, S., Rasmussen, C., Redd, K., & St. John, K. (2017). Towards the STEM DBER alliance: why we need a discipline-based STEM education research community. *International Journal of STEM Education, 4*(1), 14. https://doi.org/10.1186/s40594-017-0076-1

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 Physics Education Research, 3*, 020102. https://doi.org/10.1103/PhysRevSTPER.3.020102

Henderson, C., & Dancy, M. H. (2009). Impact of physics education research on the teaching of introductory quantitative physics in the united states. *Physical Review Special Topics Physics Education Research, 5*(2), 020107. https://doi.org/10.1103/PhysRevSTPER.5.020107

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 Physics Education Research, 8*(2), 020104. https://doi.org/10.1103/PhysRevSTPER.8.020104

Herrington, D. G., Bancroft, S. F., Edwards, M. M., & Schairer, C. J. (2016). I want to be the inquiry guy! How research experiences for teachers change beliefs, attitudes, and values about teaching science as inquiry. *Journal of Science Teacher Education, 27*(2), 183–204. https://doi.org/10.1007/s10972-016-9450-y

Hill, H. C., Charalambous, C. Y., & Kraft, M. A. (2012). When rater reliability is not enough: teacher observation systems and a case for the generalizability study. *Educational Researcher, 41*(2), 56–64. https://doi.org/10.3102/0013189X12437203

Hora, M. T. (2012). Organizational factors and instructional decision-making: a cognitive perspective. *The Review of Higher Education, 35*(2), 207–235. https://doi.org/10.1353/rhe.2012.0001

Hora, M. T., & Anderson, C. (2012). Perceived norms for interactive teaching and their relationship to instructional decision-making: a mixed methods study. *Higher Education, 64*(4), 573–592. https://doi.org/10.1007/s10734-012-9513-8

Houseknecht, J. B., Bachinski, G. J., Miller, M. H., White, S. A., & Andrews, D. M. (2020). Effectiveness of the active learning in organic chemistry faculty development workshops. *Chemistry Education Research and Practice, 21*(1), 387–398. https://doi.org/10.1039/C9RP00137A

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

Johnson, E., Keller, R., Peterson, V., & Fukawa-Connelly, T. (2019). Individual and situational factors related to undergraduate mathematics instruction. *International Journal of STEM Education, 6*(1), 23. https://doi.org/10.1186/s40594-019-0175-2

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), 8. https://doi.org/10.1186/s40594-016-0042-3

Koch, A. K. (2017). It's about the gateway courses: defining and contextualizing the issue. *New Directions for Higher Education, 2017*(180), 11–17. https://doi.org/10.1002/he.20257

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

Koziol, S. M., & Burns, P. (1986). Teachers' accuracy in self-reporting about instructional practices using a focused self-report inventory. *The Journal of Educational Research, 79*(4), 205–209. https://doi.org/10.1080/00220671.1986.10885678

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: tests in linear mixed effects models. *Journal of Statistical Software, 82*(13), 1–26. https://doi.org/10.18637/jss.v082.i13

Landis, C. R., Peace, G. E., Scharberg, M. A., Branz, S., Spencer, J. N., Ricci, R. W., Zumdahl, S. A., & Shaw, D. (1998). The new traditions consortium: shifting from a faculty-centered paradigm to a student-centered paradigm.

Journal of Chemical Education, 75(6), 741. https://doi.org/10.1021/ed075p741

Landrum, R. E., Viskupic, K., Shadle, S. E., & Bullock, D. (2017). Assessing the STEM landscape: the current instructional climate survey and the evidence-based instructional practices adoption scale. *International Journal of STEM Education, 4*(1), 25. https://doi.org/10.1186/s40594-017-0092-1

Lane, A. K., McAlpin, J. D., Earl, B., Feola, S., Lewis, J. E., Mertens, K., Shadle, S. E., Skvoretz, J., Ziker, J. P., Couch, B. A., Prevost, L. B., & Stains, M. (2020). Innovative teaching knowledge stays with users. *Proceedings of the National Academy of Sciences, 117*(37), 22665. https://doi.org/10.1073/pnas.2012372117

Lane, A. K., Skvoretz, J., Ziker, J. P., Couch, B. A., Earl, B., Lewis, J. E., McAlpin, J. D., Prevost, L. B., Shadle, S. E., & Stains, M. (2019). Investigating how faculty social networks and peer influence relate to knowledge and use of evidence-based teaching practices. *International Journal of STEM Education, 6*(1), 28. https://doi.org/10.1186/s40594-019-0182-3

Leslie, S.-J., Cimpian, A., Meyer, M., & Freeland, E. (2015). Expectations of brilliance underlie gender distributions across academic disciplines. *Science, 347*(6219), 262–265. https://doi.org/10.1126/science.1261375

Levy, R. (2019). *VITAL faculty: a growing workforce in colleges and universities*. Mathematical Association of America. Retrieved 10 July 2021 from https://www.mathvalues.org/masterblog/vital-faculty

Li, Y., & Bates, T. C. (2017). Does growth mindset improve children's IQ, educational attainment or response to setbacks? Active-control interventions and data on children's own mindsets. *SocArXiv.* https://doi.org/10.31235/osf.io/tsdwy

Lorah, J. (2018). Effect size measures for multilevel models: definition, interpretation, and times example. *Large-Scale Assessments in Education, 6*(1), 8. https://doi.org/10.1186/s40536-018-0061-2

Lorenzo, M., Crouch, C. H., & Mazur, E. (2006). Reducing the gender gap in the physics classroom. *American Journal of Physics, 74*(2), 118–122. https://doi.org/10.1119/1.2162549

Lotter, C., Harwood, W. S., & Bonner, J. J. (2007). The influence of core teaching conceptions on teachers' use of inquiry teaching practices. *Journal of Research in Science Teaching, 44*(9), 1318–1347. https://doi.org/10.1002/tea.20191

Lund, T. J., Pilarz, M., Velasco, J. B., Chakraverty, D., Rosploch, K., Undersander, M., & Stains, M. (2015). The best of both worlds: building on the copus and rtop observation protocols to easily and reliably measure various levels of reformed instructional practice. *CBE Life Sciences Education, 14*(2), ar18. https://doi.org/10.1187/cbe.14-10-0168

Lund, T. J., & Stains, M. (2015). The importance of context: an exploration of factors influencing the adoption of student-centered teaching among chemistry, biology, and physics faculty. *International Journal of STEM Education, 2*(1), 13. https://doi.org/10.1186/s40594-015-0026-8

Marsh, H. W., Rowe, K. J., & Martin, A. (2002). Phd students' evaluations of research supervision: Issues, complexities, and challenges in a nationwide Australian experiment in benchmarking universities. *The Journal of Higher Education, 73*(3), 313–348. https://doi.org/10.1353/jhe.2002.0028

Martinez, A. E., & Pilgrim, M. E. (2021). Online grading platforms: establishing a connection between course coordination and local data. *International Journal of Mathematical Education in Science and Technology.* https://doi.org/10.1080/0020739X.2021.1937731

McMahon, K., Yeh, C.S.-H., & Etchells, P. J. (2019). The impact of a modified initial teacher education on challenging trainees' understanding of neuromyths. *Mind, Brain, and Education, 13*(4), 288–297. https://doi.org/10.1111/mbe.12219

Michael, J. (2007). Faculty perceptions about barriers to active learning. *College Teaching, 55*(2), 42–47.

Moore, G. A. (2002). *Crossing the chasm*. HarperCollins.

Murray, T. A., Higgins, P., Minderhout, V., & Loertscher, J. (2011). Sustaining the development and implementation of student-centered teaching nationally: the importance of a community of practice. *Biochemistry and Molecular Biology Education, 39*(6), 405–411. https://doi.org/10.1002/bmb.20537

Myers, R. (1990). *Classical and modern regression with applications* (2nd ed.). Duxbury.

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. https://doi.org/10.1007/s10734-013-9678-9

Peace, G. E., Lewis, E. L., Burke, K. A., & Greenbowe, T. J. (2002). The multi-initiative dissemination project: active-learning strategies for college chemistry. *Journal of Chemical Education, 79*(6), 699. https://doi.org/10.1021/ed079p699

Pelletreau, K. N., Knight, J. K., Lemons, P. P., McCourt, J. S., Merrill, J. E., Nehm, R. H., Prevost, L. B., Urban-Lurain, M., & Smith, M. K. (2018). A faculty professional development model that improves student learning, encourages active-learning instructional practices, and works for faculty at multiple institutions. *CBE Life Sciences Education, 17*(2), es5. https://doi.org/10.1187/cbe.17-12-0260

Popova, M., Kraft, A., Harshman, J., & Stains, M. (2021). Changes in teaching beliefs of early-career chemistry faculty: a longitudinal investigation. *Chemistry Education Research and Practice.* https://doi.org/10.1039/D0RP00313A

Popova, M., Shi, L., Harshman, J., Kraft, A., & Stains, M. (2020). Untangling a complex relationship: teaching beliefs and instructional practices of assistant chemistry faculty at research-intensive institutions. *Chemistry Education Research and Practice, 21*(2), 513–527. https://doi.org/10.1039/C9RP00217K

Porter, S. R., & Umbach, P. D. (2001). Analyzing faculty workload data using multilevel modeling. *Research in Higher Education, 42*(2), 171–196. https://doi.org/10.1023/A:1026573503271

Prince, M. (2004). Does active learning work? A review of the research. *Journal of Engineering Education, 93*(3), 223–231. https://doi.org/10.1002/j.2168-9830.2004.tb00809.x

Prosser, M., & Trigwell, K. (1997). Relations between perceptions of the teaching environment and approaches to teaching. *British Journal of Educational Psychology, 67*(1), 25–35. https://doi.org/10.1111/j.2044-8279.1997.tb01224.x

R Core Team. (2019). *R: A language and environment for statistical computing.* R Foundation for Statistical Computing. https://www.R-project.org/

Rasmussen, C., Ellis, J., Zazkis, D., & Bressoud, D. (2014). *Features of successful calculus programs at five doctoral degree granting institutions.* Proceedings of the 38th Conference of the International Group for the Psychology of Mathematics Education and the 36th Conference of the North American Chapter of the Psychology of Mathematics Education, Vancouver, Canada. https://www.maa.org/sites/default/files/pdf/cspcc/rasmussen_ellis_zazkis_bressoud.pdf

Rasmussen, C., Apkarian, N., Hagman, J. E., Johnson, E., Larsen, S., & Bressoud, D. (2019). Brief report: characteristics of precalculus through calculus 2 programs: insights from a national census survey. *Journal for Research in Mathematics Education, 50*(1), 98–111. https://doi.org/10.5951/jresematheduc.50.1.0098

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. 111–120). MMA Press.

Rattan, A., Good, C., & Dweck, C. S. (2012). "It's ok—not everyone can be good at math": instructors with an entity theory comfort (and demotivate) students. *Journal of Experimental Social Psychology, 48*(3), 731–737. https://doi.org/10.1016/j.jesp.2011.12.012

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: applications and data analysis methods* (2nd ed.). Sage.

Reinholz, D. L., & Apkarian, N. (2018). Four frames for systemic change in STEM departments. *International Journal of STEM Education, 5*(1), 3. https://doi.org/10.1186/s40594-018-0103-x

Robinson, W. S. (1950). Ecological correlations and the behavior of individuals. *American Sociological Review, 15*(3), 351–357. https://doi.org/10.2307/2087176

Rogers, E. M. (2003). *Diffusion of innovations*. The Free Press.

Ruiz-Primo, M. A., Briggs, D., Iverson, H., Talbot, R., & Shepard, L. A. (2011). Impact of undergraduate science course innovations on learning. *Science, 331*(6022), 1269. https://doi.org/10.1126/science.1198976

SCALE-UP. (2011). Retrieved 15 July 2020 from https://sites.google.com/view/scale-up/home

Selya, A., Rose, J., Dierker, L., Hedeker, D., & Mermelstein, R. (2012). A practical guide to calculating Cohen's f^2, a measure of local effect size, from

Yik *et al. International Journal of STEM Education* (2022) 9:15

Page 22 of 23

PROC MIXED. *Frontiers in Psychology, 3*, 111. https://doi.org/10.3389/fpsyg.2012.00111

Seymour, E., DeWelde, K., & Fry, C. (2011). Determining progress in improving undergraduate STEM education: The reformers' tale. A White Paper commissioned for the Forum, "Characterizing the Impact and Diffusion of Engineering Education Innovations". https://www.nae.edu/File.aspx?id=36664

Seymour, E., & Hewitt, N. M. (Eds.). (1997). *Talking about leaving: Why undergraduates leave the sciences*. Westview Press. https://doi.org/10.1086/495437

Seymour, E., & Hunter, A.-B. (Eds.). (2019). *Talking about leaving revisited: persistence, relocation, and loss in undergraduate STEM education*. Springer International Publishing. 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), 8. https://doi.org/10.1186/s40594-017-0062-7

Sisk, V. F., Burgoyne, A. P., Sun, J., Butler, J. L., & Macnamara, B. N. (2018). To what extent and under which circumstances are growth mind-sets important to academic achievement? Two Meta-Analyses. *Psychological Science, 29*(4), 549–571. https://doi.org/10.1177/0956797617739704

Smart, J. C., & Umbach, P. D. (2007). Faculty and academic environments: Using Jolland's theory to explore differences in how faculty structure undergraduate courses. *Journal of College Student Development, 48*(2), 183–195. https://doi.org/10.1353/csd.2007.0021

Smeby, J.-C., & Try, S. (2005). Departmental contexts and faculty research activity in Norway. *Research in Higher Education, 46*(6), 593–619. https://doi.org/10.1007/s11162-004-4136-2

Smith, T., Brumskill, R., Johnson, A., & Zimmer, T. (2018). The impact of teacher language on students' mindsets and statistics performance. *Social Psychology of Education, 21*(4), 775–786. https://doi.org/10.1007/s11218-018-9444-z

Smith, M. K., Vinson, E. L., Smith, J. A., Lewin, J. D., & Stetzer, M. R. (2014). A campus-wide study of STEM courses: new perspectives on teaching practices and perceptions. *CBE Life Sciences Education, 13*(4), 624–635. https://doi.org/10.1187/cbe.14-06-0108

Snijders, T. A. B., & Bosker, R. J. (2012). *Multilevel analysis: an introduction to basic and advanced multilevel modeling* (2nd ed.). Sage Publishing.

Sonnert, G., Fox, M. F., & Adkins, K. (2007). Undergraduate women in science and engineering: effects of faculty, fields, and institutions over time. *Social Science Quarterly, 88*(5), 1333–1356. https://doi.org/10.1111/j.1540-6237.2007.00505.x

Southerland, S. A., Sowell, S., Blanchard, M., & Granger, E. M. (2011a). Exploring the construct of pedagogical discontentment: A tool to understand science teachers' openness to reform. *Research in Science Education, 41*(3), 299–317. https://doi.org/10.1007/s11165-010-9196-5

Southerland, S. A., Sowell, S., & Enderle, P. (2011b). Science teachers' pedagogical discontentment: Its sources and potential for change. *Journal of Science Teacher Education, 22*(5), 437–457. https://doi.org/10.1007/s10972-011-9242-3

Springer, L., Stanne, M. E., & Donovan, S. S. (1999). Effects of small-group learning on undergraduates in science, mathematics, engineering, and technology: a meta-analysis. *Review of Educational Research, 69*(1), 21–51. https://doi.org/10.3102/00346543069001021

Srinivasan, S., Gibbons, R. E., Murphy, K. L., & Raker, J. (2018). Flipped classroom use in chemistry education: results from a survey of postsecondary faculty members. *Chemistry Education Research and Practice, 19*(4), 1307–1318. https://doi.org/10.1039/C8RP00094H

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., … (2018). Anatomy of STEM teaching in North American universities. *Science, 359*(6383), 1468–1470. https://doi.org/10.1126/science.aap8892

Stains, M., Pilarz, M., & Chakraverty, D. (2015). Short and long-term impacts of the Cottrell scholars collaborative new faculty workshop. *Journal of Chemical Education, 92*(9), 1466–1476. https://doi.org/10.1021/acs.jchemed.5b00324

Stegall, S. L., Grushow, A., Whitnell, R., & Hunnicutt, S. S. (2016). Evaluating the effectiveness of POGIL-PCL workshops. *Chemistry Education Research and Practice, 17*(2), 407–416. https://doi.org/10.1039/C5RP00225G

Sturtevant, H., & Wheeler, L. (2019). The STEM Faculty Instructional Barriers and Identity Survey (FIBIS): development and exploratory results. *International Journal of STEM Education, 6*, 35. https://doi.org/10.1186/s40594-019-0185-0

Styers, M. L., Van Zandt, P. A., & Hayden, K. L. (2018). Active learning in flipped life science courses promotes development of critical thinking skills. *CBE Life Sciences Education, 17*(3), ar39. https://doi.org/10.1187/cbe.16-11-0332

Tai, R. H., Sadler, P. M., & Loehr, J. F. (2005). Factors influencing success in introductory college chemistry. *Journal of Research in Science Teaching, 42*(9), 987–1012. https://doi.org/10.1002/tea.20082

Theobald, E. (2018). Students are rarely independent: when, why, and how to use random effects in discipline-based education research. *CBE Life Sciences Education, 17*(3), rm2. https://doi.org/10.1187/cbe.17-12-0280

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., … (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

Tomkin, J. H., Beilstein, S. O., Morphew, J. W., & Herman, G. L. (2019). Evidence that communities of practice are associated with active learning in large STEM lectures. *International Journal of STEM Education, 6*, 1. https://doi.org/10.1186/s40594-018-0154-z

Walczyk, J. J., Ramsey, L. L., & Zha, P. (2007). Obstacles to instructional innovation according to college science and mathematics faculty. *Journal of Research in Science Teaching, 44*(1), 85–106. https://doi.org/10.1002/tea.20119

Walter, E. M., Beach, A. L., Henderson, C., Williams, C. T., & Ceballos-Madrigal, I. (2021). Understanding conditions for teaching innovation in postsecondary education: development and validation of the Survey of Climate for Instructional Improvement (SCII). *International Journal of Technology in Education, 4*(2), 166–199.

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), ar53. https://doi.org/10.1187/cbe.15-09-0193

Waxman, H. C., & Padrón, Y. N. (2004). The uses of the classroom observation schedule to improve classroom instruction. In H. C. Waxman, R. S. Hilberg, & R. G. Tharp (Eds.), *Observational research in U.S. Classrooms: new approaches for understanding cultural and linguistic diversity* (pp. 72–96). Cambridge University Press. https://doi.org/10.1017/CBO9780511616419.004

Wheeler, L. B., Maeng, J. L., & Whitworth, B. A. (2017). Characterizing teaching assistants' knowledge and beliefs following professional development activities within an inquiry-based general chemistry context. *Journal of Chemical Education, 94*(1), 19–28. https://doi.org/10.1021/acs.jchemed.6b00373

Williams, M., Apkarian, N., Uhing, K., Martinez, A. E., Rasmussen, C., & Smith, W. M. (2021). In the driver's seat: course coordinators as change agents for active learning in university precalculus to calculus 2. *International Journal of Research in Undergraduate Mathematics Education*. https://doi.org/10.1007/s40753-021-00153-w

Williams, C. T., Walter, E. M., Henderson, C., & Beach, A. L. (2015). Describing undergraduate STEM teaching practices: a comparison of instructor self-report instruments. *International Journal of STEM Education, 2*(1), 18. https://doi.org/10.1186/s40594-015-0031-y

Windschitl, M., & Sahl, K. (2002). Tracing teachers' use of technology in a laptop computer school: the interplay of teacher beliefs, social dynamics, and institutional culture. *American Educational Research Journal, 39*(1), 165–205. https://doi.org/10.3102/00028312039001165

Wood, W., & Gentile, J. (2003). Meeting report: the first national academies summer institute for undergraduate education in biology. *Cell Biology Education, 2*(4), 207–209. https://doi.org/10.1187/cbe.03-10-0042

Woodbury, S., & Gess-Newsome, J. (2002). Overcoming the paradox of change without difference: a model of change in the arena of fundamental school reform. *Educational Policy, 16*(5), 763–782. https://doi.org/10.1177/089590402237312

Wyse, S. A., Long, T. M., & Ebert-May, D. (2014). Teaching assistant professional development in biology: designed for and driven by multidimensional

data. *CBE Life Sciences Education, 13*(2), 212–223. https://doi.org/10.1187/cbe.13-06-0106

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. https://doi.org/10.1080/00461520.2012.722805

Yeager, D. S., Hanselman, P., Walton, G. M., Murray, J. S., Crosnoe, R., Muller, C., Tipton, E., Schneider, B., Hulleman, C. S., Hinojosa, C. P., Paunesku, D., Romero, C., Flint, K., Roberts, A., Trott, J., Iachan, R., Buontempo, J., Yang, S. M., Carvalho, … (2019). A national experiment reveals where a growth mindset improves achievement. *Nature, 573*(7774), 364–369. https://doi.org/10.1038/s41586-019-1466-y

Yeager, D. S., Romero, C., Paunesku, D., Hulleman, C. S., Schneider, B., Hinojosa, C., Lee, H. Y., O'Brien, J., Flint, K., Roberts, A., Trott, J., Greene, D., Walton, G. M., & Dweck, C. S. (2016). Using design thinking to improve psychological interventions: the case of the growth mindset during the transition to high school. *Journal of Educational Psychology, 108*(3), 374–391. https://doi.org/10.1037/edu0000098

Yeager, D. S., Walton, G. M., Brady, S. T., Akcinar, E. N., Paunesku, D., Keane, L., Kamentz, D., Ritter, G., Duckworth, A. L., Urstein, R., Gomez, E. M., Markus, H. R., Cohen, G. L., & Dweck, C. S. (2016). Teaching a lay theory before college narrows achievement gaps at scale. *Proceedings of the National Academy of Sciences, 113*(24), E3341. https://doi.org/10.1073/pnas.1524360113

Publisher's Note