Research in Higher Education (2024) 65:871–892
https://doi.org/10.1007/s11162-023-09769-0

STEM Faculty Instructional Beliefs Regarding Assessment, Grading, and Diversity are Linked to Racial Equity Grade Gaps

Elizabeth S. Park[1] · Mike Wilton[2] · Stanley M. Lo[3,4] · Natascha Buswell[5] · Nicole A. Suarez[4,6] · Brian K. Sato[7]

Received: 27 March 2023 / Accepted: 2 December 2023 / Published online: 23 January 2024

Abstract
Studies indicate that racial disparities in STEM achievement or equity grade gaps are associated with faculty fixed mindset beliefs; however, whether specific instructional beliefs are linked to student academic achievement remains unclear. We surveyed 216 STEM faculty to assess their mindset and instructional beliefs and linked these to detailed student transcript data (n = 31,361). Results reveal that faculty with fixed mindset beliefs also endorsed more traditional instructional beliefs regarding assessment, grading, and diversity. Further, the endorsement of these beliefs was associated with larger equity grade gaps. Analysis of faculty characteristics indicate that male faculty, full professors, and instructors in Physical Sciences tended to hold instructional beliefs that are linked to larger equity grade gaps.

Keywords STEM faculty · Grade gap · Mindset · Assessment · Grading · Diversity

✉ Elizabeth S. Park
lizpark@westat.com

✉ Brian K. Sato
bsato@uci.edu

[1] Westat, 1600 Research Blvd., Rockville, MD 20850, USA

[2] Molecular, Cellular, and Developmental Biology, University of California Santa Barbara, Santa Barbara, CA, USA

[3] Section of Cell and Developmental Biology, Division of Biological Sciences, University of California San Diego, La Jolla, CA, USA

[4] Program in Mathematics and Science Education, University of California San Diego, La Jolla, CA, USA

[5] Department of Mechanical and Aerospace Engineering, University of California Irvine, Irvine, CA, USA

[6] Program in Mathematics and Science Education, San Diego State University, San Diego, CA, USA

[7] Department of Molecular Biology and Biochemistry, University of California Irvine, 2238 McGaugh Hall, Irvine, CA 92697, USA

Underrepresented-in-STEM racially minoritized students (including Latina/o/x, African-American, Southeast Asian, Pacific Islander, and Indigenous peoples, hereafter referred to as "minoritized students") experience lower academic success in STEM college courses and report lower sense of belonging in STEM fields relative to White and East/South Asian students (hereafter referred to as "represented" students) (McGee, 2018; Rosenzweig et al., 2021).[1] Empirical research describes the overt and subtle forms of discrimination against minoritized students in college classrooms (Lee & Mccabe, 2021), such as microaggressions, derogatory statements and invalidations, and fulfilling or contradicting a stereotype, all of which contribute to unwelcoming climate (Suárez-Orozco et al., 2015). Unwelcoming educational climate contributes to decreased sense of belonging and negatively impacts students' academic performance (Seymour et al., 2019). Collectively, these experiences contribute to minoritized students leaving STEM programs at a 15–20% higher rate than those from represented backgrounds (Riegle-Crumb et al., 2019).

Previous recommendations on how to minimize these disparate outcomes have utilized a deficit model to shape educational endeavors that seek to remediate students, commonly framed as means to enable the success of underprepared students for the rigors of STEM environments (Harper, 2010; Zeidler, 2016). However, recent efforts have shifted to asset-based frameworks that leverage students' strengths while reconsidering institutional policies and practices that have given rise to and perpetuate structures that may impede the success of minoritized students (McGee, 2020). These approaches promote student equity by shifting the onus of differential student performance away from individual deficiencies and assigning solutions to educational approaches that address systemic racism (Theobald et al., 2020). Major contributors to these institutional structures are the faculty who shape STEM courses, degree programs, and broader university environments (Killpack & Melon, 2016). Within the classroom, faculty are responsible for the course instructional practices and policies that can often serve as barriers to inclusion. Growing evidence demonstrates that STEM faculty's fixed mindset beliefs foster negative classroom experiences and are linked to larger racial grade gaps, hereafter termed equity grade gaps, in their courses (Canning et al., 2019, 2021). The theorized link is that faculty mindset beliefs about innate talent and intelligence in STEM coursework may cue stereotypes and produce negative classroom climate which, in turn, may influence underrepresented-in-STEM racially minoritized students' sense of belonging and academic performance.

This study extends prior literature on faculty mindset to examine faculty instructional beliefs regarding assessment, grading, and diversity in STEM classrooms. While fixed versus growth mindset are beliefs that permeate various actions such as willingness to tackle challenges and consider improvements, instructional beliefs are specific to the classroom setting that have direct consequences on students' educational opportunities and outcomes. Furthermore, faculty beliefs on these instructional practices may influence academic achievement and inclusion in STEM classrooms by way of how the course is structured. For example, viewing course assessments as a means to accountability instead of learning may negatively shape student experiences, particularly for minoritized students (Madaus & Clarke, 2001). Further, the use of a tough-love approach to course grading wherein faculty

[1] While the term minoritized could encompass multiple intersecting dimensions of identities and backgrounds, e.g. gender, race and ethnicity, etc., we use the term "minoritized" to specifically describe the underrepresentation of STEM students in college based on ethnicity and race. We recognize that East and South Asian students are racially minoritized in other ways that are not generally observable in aggregate by representation statistics (Castro & Collins, 2021).

believe that tough grading will ensure retention of only the "most talented" individuals may serve as a gatekeeping function that disproportionately affects minoritized students (Barnes, 1997; Barnes et al., 2001). Faculty conceptions of diversity shape and inform faculty approaches to learning and teaching in the classroom (Gordon et al., 2010). Recent studies has indicated that assignments that focus on topics of diversity within STEM are associated with increased student science identity and interest (Schinske et al., 2016). In contrast, ignoring diversity or "not seeing color" in the classroom may lead to chilly classroom climates (Canning et al., 2019; Suárez-Orozco et al., 2015). For these reasons, we examine the following research questions:

(1) Are there correlations between faculty mindset beliefs and instructional beliefs?
(2) To what extent are there differences in instructional beliefs and mindset by department, faculty rank, faculty gender, and faculty type?
(3) To what extent are certain instructional beliefs associated with larger equity grade gaps?

Below we review literature that suggests potential linkages between observed equity grade gaps and faculty beliefs about intelligence, assessment, grading, and diversity.

Prior Literature on Mindset and Instructional Beliefs

Faculty Mindset Beliefs

Faculty who espouse fixed mindset beliefs endorse the idea that intelligence and ability are innate qualities that can only be minimally changed or developed (Canning et al., 2019). In turn, these beliefs can shape faculty-student interactions in the classroom. For example, a student who scores poorly on a single examination may receive unhelpful advising (based on lower expectations) from a faculty instructor who believes that student ability is fixed (Rattan et al., 2012). Also, when students perceive that their instructors have concluded they are not "cut out" for the rigors of STEM fields (i.e., endorsing fixed mindset), they report a lower sense of belonging in the course, lower interest in the subject material, and a higher degree of imposter syndrome (LaCosse et al., 2021; Leslie et al., 2015; Muenks et al., 2020).

The negative associations between faculty's fixed mindset and student outcomes are likely related to how faculty value and construct aspects of their undergraduate courses. The extant empirical literature, although limited, provides some evidence of how faculty mindset is related to assessment of student learning, grading beliefs, and whether or not topics of diversity are integrated into their curriculum (Archer, 2007; Connell et al., 2016; Lund & Stains, 2015; Sawir, 2005). Collectively, we will refer to these concepts as instructional beliefs. Below we outline how these instructional beliefs may relate to faculty mindset and influence undergraduate student experience and outcomes, particularly for students from minoritized populations.

Assessment in Higher Education Classrooms

Assessment in higher education serves multiple purposes (Fletcher et al., 2012). From the student perspective, assessments can be an opportunity to demonstrate their learning of course content or provide feedback on how to improve their study techniques (Carless et al.,

2007). For faculty, assessments provide insights into teaching effectiveness, provide feedback to students about their performance, and can be used as a mechanism to ensure student and institutional standards are maintained (Fletcher et al., 2012). In contrast, some faculty view assessment as an impediment to the teaching process, such as interfering with the time available for teaching and/or forcing faculty to teach in a way that is against their beliefs (Fletcher et al., 2012). Faculty may also believe that assessments are inaccurate and are generally imprecise when it comes to measuring learning (Fletcher et al., 2012; Harris & Brown, 2010). Additionally, some STEM faculty view assessment as a reason why students violate academic integrity; students face pressure to obtain high grades in STEM courses in order to pursue prestigious careers (Wade-Jaimes, 2021; Turner et al., 2022). Hence, faculty may consider traditional assessment practices and grades as barriers to more meaningful learning and engagement from students (Turner et al., 2022). Mindset beliefs, potentially, are also related to assessment practices implemented by STEM faculty. For example, faculty who believe in the malleability of their students' abilities to develop knowledge in the subject may be more likely to value assessment practices that have been demonstrated to improve student academic and non-academic outcomes (i.e., formative assessments) (Connell et al., 2016; Lund & Stains, 2015).

Tough Love Grading Beliefs

Similar to faculty with fixed mindsets, many STEM faculty view strict grading as a means to ensure that only students with the innate ability to succeed in the STEM discipline are able to earn the highest course grades (Gasiewski et al., 2012). STEM faculty have a well-established reputation for serving as gatekeepers who regulate student progression through the major, which manifests in how grades are assigned (Gandhi-Lee et al., 2017). Grades and grading are also viewed by some faculty as a mechanism to motivate their students. Providing grades on examinations or assignments enables student insight into how they are performing relative to their peers (Barnes, 1997). This study defines tough love grading as a combination of gatekeeping and motivational faculty beliefs. Some research has called to attention that students do not perceive tough love grading as a motivational tool; but rather as one that can instead decrease interest in the course material (Pulfrey et al., 2011), while also contributing to the lack of inclusivity in STEM classrooms (Riley & Surmitis, 2017).

Diversity Advocacy

How faculty conceptualize diversity informs the way they approach teaching and learning in the classroom (Gordon et al., 2010), which can influence student outcomes (Ho et al., 2001). Recent work by Suarez et al. (2022) has found that faculty conceptions of diversity are related to a number of factors in addition to faculty mindset, such as attending to student identities, centering student voices, fostering understanding of different perspectives, and discussing equity and inclusion issues within the discipline. For example, prior research has described how faculty adopt a fixed mindset when teaching international students, placing blame on the students and their language skills for poor classroom performance (Archer, 2007; Sawir, 2005). Such interactions give rise to unwelcoming learning environments and lead students to feel isolated and silenced (Ryan & Viete, 2009). In contrast, attending to student identities and centering student voices suggest cultural competencies in an instructor's pedagogy and foster an inclusive learning environment (Rogers-Sirin & Sirin, 2009). Faculty mindsets are communicated to students in both conscious and unconscious ways, which inadvertently sends students micromessages about their belonging

(Morrell & Parker, 2013). Purposefully fostering an understanding of different perspectives and discussing equity and inclusion issues, such as racial justice, within the discipline suggest an intentional implementation and attention to diversity in the classroom and beyond (Henry et al., 2022; Pasque et al., 2013). The incorporation of diversity-focused teaching interventions can clearly communicate to students that diverse perspectives are valued, which is critical for the retention of minoritized students in STEM (Hartman et al., 2019; Morris & Daniel, 2008). These different aspects of how faculty conceptualize diversity, individually or in combination, will likely contribute to how faculty implement teaching practices in the classroom and ultimately student learning outcomes.

Taken together, extant research has illustrated the potential relationships between faculty mindset and instructional beliefs regarding assessment, grading, and diversity, but has yet to directly link them. Additionally, while prior work has shown a relationship between mindset and equity grade gaps (Canning et al., 2019), similar work has yet to examine these instructional beliefs in the context of equity grade gaps. Given that mindset and instructional beliefs influence student experiences and the classroom environment, there is reason to believe that instructors who harbor a fixed mindset will also hold instructional beliefs that contribute to racial disparities in STEM achievement. Therefore, this study determines (1) whether faculty mindset beliefs correlate with instructional beliefs about grading, assessment, and diversity, (2) the association between these beliefs and faculty characteristics, and (3) whether certain instructional beliefs correlate with larger equity grade gaps, disaggregated by racial subgroups.

Methods

Data and Descriptive Statistics

Our research team developed a detailed survey on STEM faculty's instructional beliefs and mindset. During fall 2020, the study survey was sent out to all STEM faculty across Biological Sciences, Engineering and Computer Sciences, and Physical Sciences at a research-intensive, minority-serving institution. Out of the 594 faculty that were administered the survey, 259 faculty completed the full survey (43% response rate). When we merged faculty survey responses with the student transcript data, we removed 43 faculty as they did not teach an undergraduate-enrolling course during our study timeframe. Once we obtained university records of students taught by our survey respondents, we excluded students in independent study or research courses and courses that were graded exclusively using the Pass/No Pass system. Additionally, we removed repeat students from the sample. Our final sample, therefore, includes 216 STEM faculty who instructed 31,361 unique undergraduate students during Fall 2016 to Fall 2019 academic terms. The majority of students in the data have taken more than one course and also have taken courses with multiple instructors. Specifically, about three quarters of students have taken courses with multiple instructors. The 216 STEM faculty taught 387 courses and offered 1439 sections. On average, faculty taught six courses and 13 sections during the time span of our study. We selected this time period for the study to avoid COVID-19 affected academic terms.

We show sample statistics of the faculty respondents and of the students taught by the faculty in our sample in Table 1. Approximately three-quarters of the faculty survey respondents are male and about half of the respondents are at the Professor rank (versus Associate and Assistant Professors), both of which are comparable to national STEM

Table 1 Sample Statistics of Survey Respondents

Faculty (n = 216)	
Department	
Biological Sciences	22%
Engineering and Computer Science	41%
Physical Science	37%
Gender	
Male	72%
Female	28%
Rank	
Assistant	18%
Associate	20%
Full	54%
Type	
Teaching faculty	10%
Research faculty	81%
Lecturer	8%
Students (n = 31,361)	
Women	50%
Racially Minoritized	48%
White	12%
Black	3%
Latinx	24%
Chinese	26%
East Indian	5%
Japanese	1%
Korean	4%
Filipino	7%
Vietnamese	12%
Other/Decline to State	6%
First Generation	47%
Low Income	32%
SAT Math	638.089
SAT Verbal	576.683
Weighted High School GPA	3.954

Note. Faculty sample limited to respondents who responded to every question in the survey and taught at least one course that fit our course criteria during fall 2016 to fall 2019. Survey was administered in Fall 2020. The students are those who were taught by the faculty in our sample during Fall 2016 through Fall 2019. API = Asian and Pacific Islander. The "Other/Decline to state" category includes: Thai or other Asian (2%), American Indian (0.2%), Polynesian (0.15%), and Decline to State (3%)

faculty trends (National Center for Education Statistics [NCES], n.d.). Over 80% are tenured or tenure-track research faculty (which we define as the traditional tenure-track, research-focused faculty member), with 10% tenured or tenure-track teaching-focused faculty (which we will refer to as "teaching faculty") and 8% adjunct lecturers. Among the

Table 2 Correlation of Faculty Instructional and Mindset Beliefs

	Assessment Beliefs			Grading	Diversity Advocacy	
	Student Ranking	Student Improvement	Irrelevant	Tough Love	Curriculum Diversity	Instructor Role
Assessment: Student Ranking						
Assessment: Student Improvement	0.601***					
Assessment: Negative Perceptions	−0.262***	−0.554***				
Tough Love Grading	0.365***	0.129	−0.116			
Curriculum Diversity	−0.212**	−0.0826	0.203**	−0.410***		
Instructor Role Diversity	−0.144*	−0.0449	0.180**	−0.394***	0.637***	
Fixed Mindset	0.337***	0.0549	−0.0383	0.442***	−0.228***	−0.219**

*$p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 3 Relationship between Faculty Mindset, Faculty Characteristics and Instructional Beliefs

Variables	Assessment Beliefs			Grading	Diversity Advocacy	
	(1)	(2)	(3)	(4)	(5)	(6)
	Student Ranking	Student Improvement	Irrelevant	Tough Love	Curriculum	Instructor Role
Fixed Mindset	0.241***	0.027	0.006	0.349***	−0.166**	−0.168**
	(0.051)	(0.071)	(0.065)	(0.054)	(0.062)	(0.057)
Gender						
Female	−0.288*	−0.354*	0.223	−0.299*	0.473**	0.015
	(0.127)	(0.178)	(0.162)	(0.143)	(0.156)	(0.142)
Type						
Research faculty	0.447*	−0.027	0.045	0.212	0.044	−0.129
	(0.187)	(0.263)	(0.239)	(0.211)	(0.229)	(0.209)
Lecturer	0.653*	−0.004	−0.307	0.257	−0.143	−0.307
	(0.270)	(0.379)	(0.345)	(0.304)	(0.331)	(0.302)
Rank						
Associate	−0.229	−0.414	−0.410+	0.222	−0.226	−0.314
	(0.180)	(0.252)	(0.230)	(0.202)	(0.220)	(0.201)
Full	−0.263+	−0.244	−0.675***	0.065	−0.227	−0.350*
	(0.153)	(0.215)	(0.195)	(0.172)	(0.188)	(0.171)
Department						
Biological Sciences	0.021	−0.139	0.064	−0.379*	0.387*	0.265
	(0.148)	(0.208)	(0.189)	(0.167)	(0.181)	(0.165)
Physical Sciences	0.350**	0.318+	−0.003	0.002	−0.076	−0.147
	(0.125)	(0.175)	(0.159)	(0.140)	(0.153)	(0.139)
R-squared	0.223	0.071	0.077	0.250	0.144	0.103
Observations	216	216	216	216	216	216

Note. Each column is a separate regression estimate and all of the survey response coefficients are in standard deviation units. Reference categories are: Male, Assistant Professor, Engineering and Computer Science, Teaching Professor. Mindset is identified with two questions "Intelligence is something about a person that cannot be changed very much." and "To be honest, students have a certain amount of intelligence, and they really can't do much to change it."

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 4 Relationship between Instructional Beliefs and Racial Equity Grade Gap

	Mindset	*Assessment Beliefs*			*Grading*	*Diversity Advocacy*	
	Fixed	Student Ranking	Student Learning	Irrelevant	Tough Love	Curriculum Diversity	Instructor Role
Racial Equity Grade Gap	−0.033***	−0.038***	−0.052***	0.019**	−0.055***	0.029***	0.049***
	(0.007)	(0.007)	(0.005)	(0.007)	(0.006)	(0.007)	(0.007)
Student-by-class-term	138,797						
Students	31,361						
Classes	1439						
Instructors	216						

Note Each column represents separate regression models. The coefficient represents the interaction between students' racially minoritized status and instructors' instructional beliefs on standardized course grades. Students are identified as racially minoritized if they are American Indian, Black, Chicano/Latinx, Philipino, Polynesian, Thai, and Vietnamese. All models include student and section fixed effects. Standard errors, clustered at the class level, are in parentheses

*$p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

31,361 students in our analytic sample, half are female and 48% of the students identify as racially minoritized. Notably, a quarter of the STEM students in our sample identify as Latinx. 47% of students are first-generation college students and a third of the students are low-income (Tables 2, 3, 4). The average weighted high school GPA of the students is 3.9.

Survey Constructs

We administered a survey to assess STEM faculty's mindset beliefs and instructional beliefs about assessment, grading, and diversity in STEM. These survey constructs are based on prior published work (see Appendix Table 5 for survey items and citations).

The standardized measure of faculty mindset beliefs is based on two questions "intelligence is something about a person that cannot be changed very much" and "to be honest, students have a certain amount of intelligence and they can't really do much to change it" as in Canning et al. (2019).

For the assessment questions, we concluded that the assessment construct is better suited as three sub-constructs based on confirmatory factor analysis. Therefore, the 15 assessment-related survey items were separated into three scales: *assessment for student ranking*, *assessment for student improvement*, and *assessment as irrelevant for teaching and learning*. The assessment for student ranking fit statistics are: SRMR: 0.038; RMSEA: 0.068; CFI: 0.993. These fit indices are among the most widely reported in the literature with lower RMSEA suggesting good fit, and CFI greater than or equal to 0.95 and SRMR less than 0.08 suggesting acceptable fit (please see Kline, 2016 for explanation). The student improvement scale's fit statistics are: SRMR: 0.048; RMSEA: 0.128; CFI: 0.971. Based on these results, the two scales have good fit. The three-item negative perceptions regarding assessment measure did not have enough variation in responses and therefore were treated as an index. The assessment for ranking scale focuses on whether students meet qualification standards and the scale reliability coefficient is 0.724. The assessment for student improvement scale examines whether assessment provides feedback to students and whether assessments help students improve their learning. This six-item scale has a reliability coefficient of 0.826. The three-item negative perceptions of assessment identify instructors' beliefs about fairness of assessments and whether instructors believe assessments interfere with their teaching. This index has a reliability coefficient of 0.806.

For faculty grading beliefs, we adapted the items from Barnes (1997) and created a six-item faculty *tough love grading beliefs* scale. The correlated two-factor model, gatekeeping and tough love scale has excellent fit: SRMR: 0.037, RMSEA: 0.083, CFI: 0.975. We see that gatekeeping is highly correlated with tough love. Gatekeeping is mostly driven by the question "assigning grades to students based on their performance relative to others gives them a dose of reality" and the tough love scale is driven by the question "tough grading is needed to identify students who are not prepared for college-level work." The grading scale has an internal consistency of 0.680.

For diversity advocacy related survey questions, we asked four items related to beliefs about the instructor's role in attending to student identities and centering student voices in the classroom, which we call *instructor's role in diversity advocacy* and another four items related to curriculum diversity such as fostering understanding of different perspectives and discussing equity and inclusion issues within the discipline, which we call *curriculum diversity*. Survey items were adapted based on findings from Pasque et al. (2013) and Suarez et al. (2022). Both instructor role diversity and curriculum diversity scales have good fit.The instructor role diversity has the following fit statistics: SRMR: 0.028; RMSEA: 0.060; CFI: 0.998. The

curriculum diversity scale has the following fit statistics: SRMR: 0.036; RMSEA: 0.071; CFI: 0.998. The diversity advocacy scales have an internal consistency of 0.780 for instructor role diversity and 0.790 for curriculum diversity. All of the items are on a six point response ranging from strongly disagree to strongly agree, which were standardized such that results can be comparable across model estimations. The confirmatory factor analyses were conducted using R and the lavaan package.

Analytical Strategy

We began our inquiry by examining the correlation between faculty mindset beliefs and instructional beliefs about grading, assessment, and diversity. Then, we examined whether instructional beliefs differ depending on department, faculty rank, gender, and type controlling for faculty mindset as specified in Eq. (1):

$$\begin{aligned} Instructionalbeliefs_i = \beta_0 + \beta_1\left(facultymindset_i\right) + \beta_2\left(gender_i\right) \\ + \beta_3\left(rank_i\right) + \beta_4\left(type_i\right) + \beta_5\left(department_i\right) + \varepsilon_i \end{aligned} \tag{1}$$

where $Instructionalbeliefs_i$ represents standardized measures of assessment, grading, and diversity advocacy beliefs for faculty *i.* $facultymindset_i$ indicates a standardized measure of faculty mindset beliefs. To examine the association between faculty characteristics and instructional beliefs, we included a list of faculty covariates, with β_1 through β_5 representing the coefficients of faculty characteristics. We included the following faculty characteristics in the model: faculty gender (male vs. female), type (research-focused vs. teaching-focused), rank (Lecturer, Assistant, Associate, Full), and department (Computer Science and Engineering, Biological Sciences, Physical Sciences). β_0 represents the intercept and β_1 represents the relationship between faculty mindset and instructional beliefs in standard deviation units, and ε_i represents random error.

To answer our third research question, we estimated the relationship between the standardized measure of instructional beliefs regarding assessment, grading, and diversity and a standardized measure of equity grade gaps using a series of regression models that included student and section (i.e., classroom) fixed effects (Fairlie et al., 2014; Xu & Solanki, 2020). We estimated the changes in GPA of the represented students in a particular section relative to the changes in GPA of minoritized students in that same section. We included section fixed effects to account for the fact that instructors may impose different instructional standards depending on the class. This also removed any observed and unobserved aspects of an instructor that is constant overtime. We further included student fixed effects because students are not randomly assigned to faculty with specific instructional beliefs about assessment, grading, or diversity. Students may have selected into certain classes based on unobserved factors: for example, a student may desire to prove to oneself that they can master the course material with an instructor who exercises tough love and thus have opted to take a course with a certain type of instructor. Because 95% of the students in the data took more than one course within each term and across terms, we included student fixed effects to account for both observable and unobservable time-invariant characteristics of the student that could be correlated with taking courses with a certain type of instructor. Equation (2) formalizes the model we estimate:

$$y_{ic} = \delta\left(minoritized_i * facultybeliefs_c\right) + \gamma_i + \theta_c + \varepsilon_{ic} \tag{2}$$

y_{ic} indicates standardized measure of equity grade gap in a particular class with i indexing students and c indexing section offered during a specific term. $facultybeliefs_c$ represents

standardized measures of assessment, grading, and diversity advocacy as well as faculty mindset beliefs. δ represents the equity grade gap, which is the gap in course grades between racially minoritized and represented students, measured on a 0 to 4 scale with 0 indicating an "F" and a 4 indicating an "A+." A significant positive coefficient δ indicates that racially minoritized students perform better in courses taught by instructors who espouse certain instructional beliefs whereas a significant negative coefficient δ indicates that racially minoritized students perform worse.

Rather than including instructor and student level controls, we included section (θ_c) and student fixed effects (γ_i). The inclusion of section fixed effects addresses selection that may arise due to factors such as different grading policies. By including section fixed effects, we are comparing student groups within the same classroom and who are subject to the same faculty standards, expectations, and assignments. Furthermore, the inclusion of student fixed effects addresses student sorting into different instructors—although this does not address differential sorting into instructors among racially minoritized and represented students which we discuss in the limitations section. By modeling both fixed effects, we essentially compare the performance of racially minoritized and represented students in the same section taught by an instructor with a particular type of instructional belief. By including student and section fixed effects we capitalize on within-student variation as well as between students within class variation. Because the section fixed effects subsumes faculty level characteristics and is a combination of faculty, course, section, quarter and year, all of the faculty-level and course-level variables dropped out of the model. Similarly, all student-level variables dropped out of the model with the inclusion of the student fixed effects. In addition to student and section fixed effects, we included an indicator for every term-year combination. This allowed us to compare course grades within each term in a given year. The standard errors of all the regression estimates were clustered at the section-level. All of the fixed effects estimations were conducted in Stata, a statistical software developed by StataCorp.

Results

Are There Correlations Between Faculty Mindset Beliefs and Instructional Beliefs?

As prior studies indicated that faculty mindset and beliefs of assessment, grading, and diversity may be linked to student outcomes, we sought to determine whether there is a relationship among these constructs for the STEM faculty survey participants. Examining the correlation coefficients of the survey constructs, assessment for student ranking purposes and assessment for improvement were positively correlated ($r=0.601$). Assessment for student ranking was also positively correlated with tough love grading approaches ($r=0.365$) and fixed mindset beliefs ($r=0.337$). Faculty perceptions that assessments help determine students' qualifications (i.e., assessment for student ranking) and faculty perceptions that assessments interfere with teaching (i.e., assessments are irrelevant) were negatively correlated with one another ($r=-0.554$).

Similar to assessment for student ranking beliefs, tough love grading beliefs were positively correlated with fixed mindset beliefs ($r=0.442$). Tough love grading beliefs were negatively correlated with diversity advocacy. In contrast, instructors' perceived role in discussing diversity and promoting diversity in STEM curriculum were positively correlated ($r=0.637$). Despite these patterns, we noted that the faculty average responses in our study sample on fixed mindset and tough love grading approaches was somewhat disagree while

the average responses to diversity advocacy beliefs were somewhat agree on a six-point disagree to agree scale (See Appendix Table 6).

To What Extent Are There Differences in Instructional Beliefs and Mindset by Department, Faculty Rank, Faculty Gender, and Faculty Type?

We next examined the relationship between these constructs holding constant faculty characteristics using Eq. (1) (Table 3). Faculty who endorsed fixed mindset beliefs were more likely to believe that assessment is a means to rank students (β= 0.24 standard deviation unit on a five point agree-disagree scale, $p<0.001$) and to hold tough love grading beliefs (β= 0.35 SD, $p<0.001$). In addition, faculty who endorsed fixed mindset beliefs were also less likely to advocate that STEM curriculum should promote diversity (β= −0.17 SD, $p<0.01$) and that the instructor's role is to promote equity in their discipline (β= −0.17 SD, $p<0.01$).

Controlling for mindset beliefs, female faculty were more likely to endorse the inclusion of diversity topics in STEM curricula than male faculty (β= 0.47 SD, $p<0.001$), less likely to believe that assessment should be used for ranking purposes (β= −0.29 SD, $p<0.05$) or to improve student learning (β= −0.35 SD, $p<0.05$), and less likely to hold tough love beliefs regarding grading (β= −0.29 SD, $p<0.05$) (Table 3). Both tenure-track research faculty and lecturers were more likely to hold the belief that assessment should be used for ranking purposes (research: β= 0.45 SD, $p<0.01$; lecturer: β= 0.65 SD, $p<0.05$) relative to tenure-track teaching faculty. Across faculty rank, full professors were less likely to believe that assessment is irrelevant (β= −0.68 SD, $p<0.001$) and less likely to believe it is the instructor's role to foster diversity in the classroom (β= −0.35 SD, $p<0.05$) relative to assistant professors. Looking at disciplinary differences, we noted that Biological Sciences faculty compared to Engineering and Computer Science faculty were less likely to hold tough love beliefs regarding grading (β= −0.38 SD, $p<0.05$) and more likely to support the idea that diversity should be promoted in the curriculum (β= 0.39 SD, $p<0.05$). In addition, Physical Sciences faculty were more likely to value assessment for ranking purposes (β= 0.35 SD, $p<0.01$) relative to Engineering and Computer Sciences faculty.

To What Extent Are Certain Instructional Beliefs Associated with Larger Equity Grade Gaps?

In order to investigate the relationship between instructional beliefs and disparities in grades between minoritized and represented populations, we sought to replicate the prior finding that faculty mindset is correlated with student grades (Canning et al., 2019). We found a larger equity grade gap between minoritized and non-minoritized students when taught by STEM faculty with fixed, as opposed to growth, mindset beliefs, holding constant student-level and classroom-level differences (Table 4). This is consistent with findings from prior studies (e.g., Canning et al., 2019). Specifically, the equity grade gap increased by 0.03 standard deviations ($p<0.001$) when taught by instructors who endorsed fixed mindset beliefs.

We next investigated the relationship between faculty beliefs regarding mindset, assessment, grading, and diversity and student equity grade gaps using Eq. 2. Similar to prior work, students taught by STEM faculty who held fixed mindset beliefs earned lower grades and experienced larger equity grade gaps relative to students in courses taught by STEM faculty who held growth mindset beliefs. Examining the interaction between racially minoritized

indicator and faculty instructional beliefs on assessment, we found that, on average, minoritized students were awarded 0.04—0.05 standard deviations lower course grades ($p < 0.001$) than their represented peers when instructors believed that assessment should be used in the learning process and for student ranking purposes. Similarly, in courses where instructors endorsed a tough love belief regarding grading, minoritized students received 0.05 standard deviations lower course grades ($p < 0.001$). The equity grade gaps are notable given that 0.04 to 0.05 standard deviations of grades represents 18–23 percent of the total gap between minoritized and represented students' grades at this university.

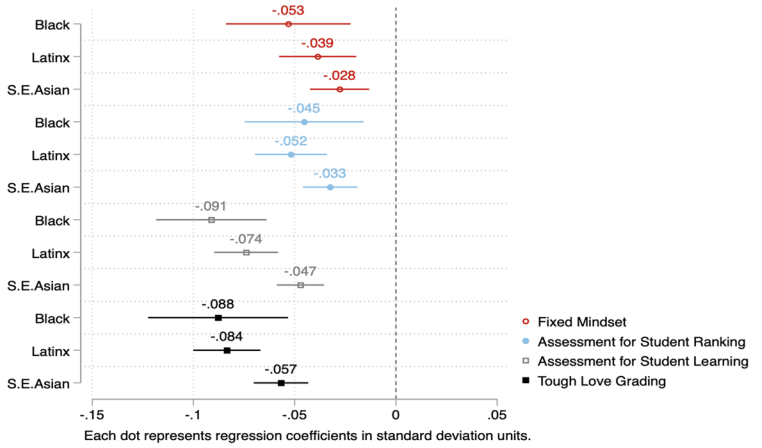
In contrast, the equity grade gap was reversed when students took a course with an instructor who endorses the belief that it is the instructor's role to center student identities and voices as well as to foster discussions of diversity, equity, and inclusion in the STEM curriculum. Minoritized students, on average, were awarded 0.03 standard deviation higher course grades ($p < 0.001$) in courses where the instructor agreed on the importance of diversity and equity in the curriculum relative to students taught by instructors who did not agree with these instructional beliefs, after accounting for student sorting, instructor and section-level differences, and subject-specific differences. We also observed this pattern in courses where instructors advocate for race and ethnic diversity to be more strongly reflected in the STEM curriculum. In addition, we found that minoritized students received 0.02 standard deviations higher course grades ($p < 0.001$) when taught by instructors who held the belief that assessments are irrelevant relative to peers taught by instructors who did not endorse these beliefs.

Disaggregation by Racial Identification

The term "minoritized" masks the varied experiences of different racial/ethnic groups. It may be that faculty instructional beliefs are associated with greater/lesser equity grade gaps when we examine Latinx students, Black students, or Southeast Asian students. We therefore disaggregated the results shown in Table 3 and re-estimated the model by interacting instructional beliefs with a student indicator for Latinx, Black, and Southeast Asian (we do not include Native American and Pacific Islander students in this analysis due to very small sample sizes). We found that minoritized students regardless of different subgroups, on average, received lower grades in STEM courses when instructors subscribed to fixed mindset beliefs and tough love grading beliefs while students received higher grades in courses where instructors espoused diversity advocacy beliefs in STEM (Fig. 1). However, Black and Latinx students particularly underperformed when instructors held assessment for improvement perceptions and tough love grading beliefs (Fig. 2). Specifically, Black and Latinx students were awarded 0.07–0.09 standard deviations lower course grades ($p < 0.001$) when taught by these instructors whereas Southeast Asians received 0.05–0.06 standard deviations lower course grades ($p < 0.001$). When taught by STEM instructors who reported that it is their role to foster conversations about diversity, Black and Latinx students received 0.05–0.06 standard deviations higher course grades ($p < 0.001$).

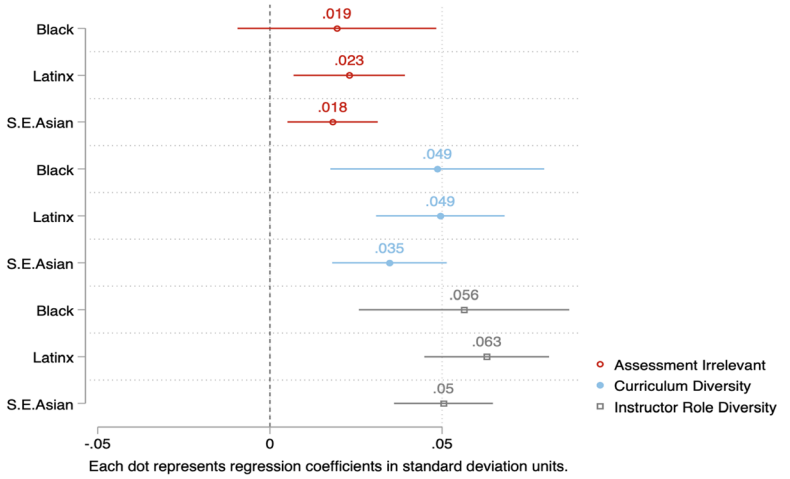
Limitations and Future Research

We note a few limitations and discuss future research suggestions. First, our estimations are not causal. The student and section fixed effects do not account for differential sorting of student groups into different sections. The underlying assumption is that, absent faculty with different beliefs, racially minoritized and represented groups would have followed a common GPA trend. Yet, the current model cannot rule out confounding factors such as



Note. Each dot represents the interaction coefficient between students' racially minoritized status and instructors' instructional beliefs on standardized course grades using a student and section fixed effects model. The lines indicate 95 percent confidence interval of the coefficients.

Fig. 1 Instructional Beliefs Associated with Negative Racial Equity Grade Gaps



Note. Each dot represents the interaction coefficient between students' racially minoritized status and instructors' instructional beliefs on standardized course grades using a student and section fixed effects model. The lines indicate 95 percent confidence interval of the coefficients.

Fig. 2 Instructional Beliefs Associated with Positive Racial Equity Grade Gaps

systematic sorting into sections taught by faculty with certain instructional beliefs. For instance, highly motivated racially-minoritized students may have obtained information on faculty's instructional beliefs on assessment, grading, and/or diversity and may have opted to take certain sections. We cannot control for all possible confounding factors in our model. Second, we do not know to what extent instructors enacted these instructional beliefs in the classroom. Instructors may report agreement in the survey but not necessarily implement their beliefs in their classroom. While studies show that surveys are valid proxies for actual behavior (Li et al., 2020), we

acknowledge that our results hinge on reported responses rather than observed behavior. Future research that describes instructional practices as they relate to assessment, grading, and diversity advocacy in STEM classrooms will elucidate the extent to which beliefs align with practices in STEM classrooms.

Discussion

This study extends existing literature by examining how mindset beliefs connect with instructor beliefs regarding assessment, grading, and diversity that ultimately contribute to racial disparities in STEM course achievement. Specifically, leveraging assessments for student ranking as opposed to a means to improve student learning, grading through a gatekeeping and tough love lens, and a belief that it is not the instructor's role to incorporate diversity into the classroom and curriculum were all associated with fixed mindset beliefs and widening equity grade gaps. This study refers to these beliefs as "traditional" beliefs as they contribute to the chilly climate traditionally found in STEM programs (Seymour et al., 2019).

Notably, racially minoritized students performed better relative to represented students when they took courses with STEM instructors who agreed on the importance of diversity in STEM classrooms. The equity grade gap reversed when students took courses with instructors who believed it is their role to foster conversations about diversity, and the positive trend was more pronounced for Black and Latinx students. These results point to the value of incorporating topics related to diversity into the curriculum, and intentionally discussing structural racism in STEM higher education as a way to help racially minoritized students feel seen and heard in underrepresented spaces. In contrast, the equity grade gap widened when students took courses with STEM instructors who believed that tough grading motivates students to perform their best. Our research finds that ranking students and administering harsh grades were associated with lower performance among racially minoritized students in STEM courses. Research has indicated that racially minoritized students are more likely to face microaggressions in STEM classrooms and feel extreme pressure to outperform and prove themselves (McGee, 2020). And, these students are prone to feeling judged when they perform suboptimally. Despite STEM faculty's intentions to increase student achievement through tough grading, our study indicates that this type of grading belief has the opposite effect.

Analyses of faculty demographics reveal significant differences in mindset as well as faculty beliefs. Our results indicate that female faculty tend to espouse growth mindset beliefs, express that assessment should not be used for student ranking, and that the curriculum should incorporate diversity. The finding that female faculty are likely to hold beliefs that benefit minoritized students aligns with prior work that they are more likely to implement equity-oriented pedagogies (Hurtado et al., 2012). While this is a promising finding, it also aligns with prior research showing that female faculty (and other minoritized faculty) face increased burden for diversity-related activities, including advising minoritized students and increased service loads, which leads to less time spent on research tasks (O'Meara et al., 2017). As such, institutions need to be aware of this increased burden on female faculty as they consider means to improve outcomes for minoritized STEM students.

We also find that instructor-type correlates with faculty perspective on the purpose of assessment. Research faculty and adjunct lecturers were significantly more likely to endorse the belief of assessment as a mechanism of student ranking than tenure-track teaching-focused faculty. As

tenure-track teaching faculty are hired on the basis of pedagogical expertise and knowledge (Harlow et al., 2022), institutions might expect that teaching faculty hold beliefs that are more focused on student learning and creating a more positive classroom environment rather than ranking students. At the same time, individuals in this tenure-track teaching faculty position tend to be trained similarly to research faculty in their discipline in doctoral programs, which may explain why there were few differences overall among the measured beliefs (Harlow et al., 2020; Harlow et al., 2022).

In regard to discipline, Biological Sciences faculty were less likely to possess tough love and gatekeeping beliefs for grading and more likely to support diversity in the curriculum. This may be related to a more inclusive classroom climate in Life Sciences as the field reaches gender enrollment parity. A recent Pew Research Center report (Fry et al., 2021) has found that females are the majority of degree earners in Life Sciences fields while making up only 22% and 40% of graduates in Engineering and Physical Sciences fields, respectively. Yet, the same report has found that all STEM fields are similarly unsuccessful at graduating Black and Latinx students, which reflects the ongoing need to address this issue across STEM disciplines.

As opposed to focusing on perceived student deficits, this work highlights the need to address the instructor's role in the commonly observed disparities in success of underrepresented-in-STEM racially minoritized students relative to represented students. While research has identified the benefits of particular classroom practices, we also know that faculty choose to implement such practices based on their own personal opinions and experiences (Andrews & Lemons, 2015). This, in combination with our findings, reinforces the importance of faculty reflection regarding their beliefs and practices to better understand what they do in the classroom and why. Reflection, in the form of faculty development on inclusive and evidence-based teaching practices (Beach et al., 2016) and the writing of reflective teaching statements (Hubball et al., 2005), are means to transform the classroom space. Such reflection though cannot be expected to occur in a vacuum. Particularly at research-intensive institutions, there is a well-established misalignment between faculty promotion processes that prioritize research success, and the time needed to reflect on and improve one's teaching (Brownell & Tanner, 2012). If institutions expect their faculty to reconsider their beliefs, the appropriate reward structures must exist to enable them to prioritize such activities. This can include encouraging faculty to submit a reflective teaching statement to complement their student evaluations of teaching as a means to demonstrate teaching excellence, while also providing them with professional development opportunities to train faculty to create and leverage such statements to improve their teaching.

Conclusion

Our study sheds light on the critical role that faculty mindset and instructional beliefs play in perpetuating racial equity grade gaps. Extending prior research, our findings reveal an alignment between fixed mindset and traditional beliefs surrounding assessment, grading, and diversity. The correlation between these instructional beliefs and racial equity grade gaps emphasizes the need for faculty to engage in reflective practices, critically examining the impact their beliefs have on their teaching practices. Moreover, our exploration of faculty characteristics, such as gender, instructor type, and academic discipline, provides valuable insights for institutional leadership. By understanding these factors, institutions can implement strategic measures to promote faculty development and work towards closing racial equity grade gaps.

Appendix

See Tables 5 and 6.

Table 5 Constructs and survey items

Construct	Survey Items
Faculty Mindset (Canning, 2019)	Two-items on fixed mindset Intelligence is something about a person that cannot be changed very much To be honest, students have a certain amount of intelligence, and they really can't do much to change it
Assessment Beliefs (Fletcher et al., 2012)	Fifteen items were separated into three scales: assessment for accountability purposes (6 items), assessment for student improvement (6 items), and negative perceptions about assessment (3 items) *Assessment for Student Ranking* Assessment is assigning a grade or level to student work Assessment determines if students meet qualification standards Assessment results can be depended on Assessment places students into ranks Assessment results are trustworthy Assessment results are consistent *Assessment for Student Improvement* Assessment provides feedback to students about their performance Assessment establishes what students have learned Assessment helps students improve their learning Assessment measures students' higher order thinking skills Assessment feeds back to students their learning needs Assessment is a way to determine how much students have learned from teaching *Assessment is Irrelevant* Assessment forces instructors to teach in a way that is against their beliefs Assessment is unfair to students Assessment interferes with teaching
Tough Love Grading Beliefs (Barnes, 1997; Barnes et al., 2001)	Even with effective instruction, the idea that half the class deserves A's is unreasonable Generally, a high percentage of "A's" in a class indicates low standards or a lack of rigor in assessing achievement The distribution of scores on a well-written exam following effective instruction should be "piled up" at the upper end of the scale range Tough grading is needed to identify students who are not prepared for college-level work Tough grading motivates students to perform their best Assigning grades to students on the basis of their performance relative to others gives them a dose of reality

Table 5 (continued)

Construct	Survey Items
Diversity Advocacy (Park & Denson 2009; Suarez et al., 2022)	Eight items were asked regarding faculty's beliefs about embedding diversity discussions in the STEM curriculum as well as beliefs about instructor role in learning about students' life experiences and fostering acceptance of different beliefs and perspectives *Curriculum Diversity* Racial and ethnic diversity should be more strongly reflected in the curriculum A racially/ethnically diverse student body enhances the educational experience of all students The goal of undergraduate education should be to enhance students' knowledge of and appreciation for other racial/ethnic groups A personal objective of mine is to help promote racial understanding *Instructor Role Diversity* Please indicate the extent to which you agree it is your role to: Learn about individual students' life experiences Encourage students to voice disagreement with ideas being presented in class Foster acceptance of different beliefs and perspectives Foster classroom discussions on equity issues in your discipline

Table 6 Means and standard deviations of survey construct and indices

	Mean	Median	Standard deviation
Assessment: Student Ranking	4.118	4.167	0.653
Assessment: Student Improvement	4.607	4.667	0.689
Assessment: Negative Perceptions	2.363	2.000	0.956
Tough Love Grading	2.901	2.833	0.868
Curriculum Diversity	4.438	4.500	0.974
Instructor Role Diversity	4.507	4.500	0.958
Fixed Mindset	2.081	2.000	1.019
N	216		

Note These indices and scales are on a strongly disagree (1) to strongly agree (6) scale. Some items within scales were reverse coded

Acknowledgements The authors would like to thank the survey participants for their time in completing the study survey.

Funding This work was supported by the National Science Foundation (NSF DUE 1821724).

Data Availability Data are not publicly available. All inquiries about the data and the paper should be directed to the corresponding authors.

Declarations

Conflict of interest The authors have no competing interests to declare that are relevant to the content of this article.

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), 7.

Archer, L. (2007). Diversity, equality and higher education: A critical reflection on the ab/uses of equity discourse within widening participation. *Teaching in Higher Education, 12*(5–6), 635–653.

Barnes, L. L. (1997). Development of the faculty beliefs about grades inventory. *Educational and Psychological Measurement, 57*(3), 459–468.

Barnes, L. L., Bull, K. S., Campbell, N. J., & Perry, K. M. (2001). Effects of academic discipline and teaching goals in predicting grading beliefs among undergraduate teaching faculty. *Research in Higher Education, 42*(4), 455–467.

Beach, A. L., Sorcinelli, M. D., Austin, A. E., & Rivard, J. K. (2016). *Faculty development in the age of evidence: Current practices, future imperatives*. Stylus Publishing, LLC.

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.

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.

Canning, E. A., Ozier, E., Williams, H. E., AlRasheed, R., & Murphy, M. C. (2021). Professors who signal a fixed mindset about ability undermine women's performance in STEM. *Social Psychological and Personality Science, 13*(5), 927–937.

Carless, D. (2007). Learning-oriented assessment: Conceptual bases and practical implications. *Innovations in Education and Teaching International, 44*(1), 57–66.

Castro, A. R., & Collins, C. S. (2021). Asian American women in STEM in the lab with "White Men Named John." *Science Education, 105*(1), 33–61.

Connell, G. L., Donovan, D. A., & Chambers, T. G. (2016). Increasing the use of student-centered pedagogies from moderate to high improves student learning and attitudes about biology. *CBE—Life Sciences Education*, *15*(1), ar3.

Fairlie, R. W., Hoffmann, F., & Oreopoulos, P. (2014). A community college instructor like me: Race and ethnicity interactions in the classroom. *American Economic Review, 104*(8), 2567–2591.

Fletcher, R. B., Meyer, L. H., Anderson, H., Johnston, P., & Rees, M. (2012). Faculty and students conceptions of assessment in higher education. *Higher Education, 64*(1), 119–133.

Fry, R., Kennedy, B., & Funk, C. (2021). STEM jobs see uneven progress in increasing gender, racial and ethnic diversity. *Pew Research Center Science & Society*.

Gandhi-Lee, E., Skaza, H., Marti, E., Schrader, P. G., & Orgill, M. (2017). Faculty perceptions of student recruitment and retention in STEM fields. *European Journal of STEM Education, 2*(1), 2.

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.

Gordon, S., Reid, A., & Petocz, P. (2010). Educators' conceptions of student diversity in their classes. *Studies in Higher Education, 35*(8), 961–974.

Harlow, A., Lo, S. M., Saichaie, K., & Sato, B. K. (2020). Characterizing the University of California's tenure-track teaching position from the faculty and administrator perspectives. *PLoS ONE, 15*(1), e0227633.

Harlow, A. N., Buswell, N. T., Lo, S. M., & Sato, B. K. (2022). Stakeholder perspectives on hiring teaching-focused faculty at research-intensive universities. *International Journal of STEM Education, 9*(1), 54.

Harris, L. R., & Brown, G. T. (2010). Mixing interview and questionnaire methods: Practical problems in aligning data. *Practical Assessment, Research, and Evaluation, 15*(1), 1.

Harper, S. R. (2010). An anti-deficit achievement framework for research on students of color in STEM. *New Directions for Institutional Research, 2010*(148), 63–74.

Hartman, H., Forin, T., Sukumaran, B., Farrell, S., Bhavsar, P., Jahan, K., & Macey, D. (2019). Strategies for improving diversity and inclusion in an engineering department. *Journal of Professional Issues in Engineering Education and Practice, 145*(2), 04018016.

Ho, A., Watkins, D., & Kelly, M. (2001). The conceptual change approach to improving teaching and learning: An evaluation of a Hong Kong staff development programme. *Higher Education*, *42*(2), 143–169.

Henry, J., Flaieh, K., & Buswell, N. (2022, August). *The importance of female crash test dummies: Bringing equity discussions into engineering classrooms through questioning inequitable product design.* In 2022 ASEE *Annual Conference & Exposition.*

Hubball, H., Collins, J., & Pratt, D. (2005). Enhancing reflective teaching practices: Implications for faculty development programs. *Canadian Journal of Higher Education, 35*(3), 57–81.

Hurtado, S., Eagan, K., Pryor, J. H., Whang, H., & Tran, S. (2012). *Undergraduate teaching faculty: The 2010–2011 HERI faculty survey*. University of California.

Kline, R. B. (2016). Principles and Practice of Structural Equation Modeling (4th ed.). New York, NY: The Guilford Press.

Killpack, T. L., & Melón, L. C. (2016). Toward inclusive STEM classrooms: what personal role do faculty play? *CBE—Life Sciences Education, 15*(3), es3.

LaCosse, J., Murphy, M. C., Garcia, J. A., & Zirkel, S. (2021). The role of STEM professors' mindset beliefs on students' anticipated psychological experiences and course interest. *Journal of Educational Psychology, 113*(5), 949.

Lee, J. J., & Mccabe, J. M. (2021). Who speaks and who listens: Revisiting the chilly climate in college classrooms. *Gender & Society, 35*(1), 32–60.

Leslie, S. J., Cimpian, A., Meyer, M., & Freeland, E. (2015). Expectations of brilliance underlie gender distributions across academic disciplines. *Science, 347*(6219), 262–265.

Li, Q., Baker, R., & Warschauer, M. (2020). Using clickstream data to measure, understand, and support self-regulated learning in online courses. *The Internet and Higher Education, 45*, 100727.

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), 1–21.

Madaus, G. F., & Clarke, M. (2001). The adverse impact of high stakes testing on minority students: Evidence from 100 years of test data.

McGee, E. (2018). "Black Genius, Asian Fail": The detriment of stereotype lift and stereotype threat in high-achieving Asian and black STEM students. *AERA Open, 4*(4), 2332858418816658.

McGee, E. O. (2020). Interrogating structural racism in STEM higher education. *Educational Researcher, 49*(9), 633–644.

Morrell, C., & Parker, C. (2013). Adjusting micromessages to improve equity in STEM. *Diversity & Democracy, 16*, 2.

Morris, L. K., & Daniel, L. G. (2008). Perceptions of a chilly climate: Differences in traditional and nontraditional majors for women. *Research in Higher Education, 49*, 256–273.

Muenks, K., Canning, E. A., LaCosse, J., Green, D. J., Zirkel, S., Garcia, J. A., & Murphy, M. C. (2020). Does my professor think my ability can change? Students' perceptions of their STEM professors' mindset beliefs predict their psychological vulnerability, engagement, and performance in class. *Journal of Experimental Psychology: General, 149*(11), 2119.

National Center for Education Statistics. (2022). Characteristics of Postsecondary Faculty. *Condition of Education*. U.S. Department of Education, Institute of Education Sciences. Retrieved Aug 5, 2022 from https://nces.ed.gov/programs/coe/indicator/csc

O'Meara, K., Kuvaeva, A., Nyunt, G., Waugaman, C., & Jackson, R. (2017). Asked more often: Gender differences in faculty workload in research universities and the work interactions that shape them. *American Educational Research Journal, 54*(6), 1154–1186.

Park, J. J., & Denson, N. (2009). Attitudes and advocacy: Understanding faculty views on racial/ethnic diversity. *The Journal of Higher Education*, 415–438.

Pasque, P. A., Chesler, M. A., Charbeneau, J., & Carlson, C. (2013). Pedagogical approaches to student racial conflict in the classroom. *Journal of Diversity in Higher Education, 6*(1), 1.

Pulfrey, C., Buchs, C., & Butera, F. (2011). Why grades engender performance-avoidance goals: The mediating role of autonomous motivation. *Journal of Educational Psychology, 103*(3), 683.

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.

Riley, C. L., & Surmitis, K. A. (2017). Inclusivity on campus: Strategies for counselors creating community for all students with an emphasis on Muslim college students. *Counseling & Wellness: A Professional Counseling Journal, 6*, 1.

Riegle-Crumb, C., King, B., & Irizarry, Y. (2019). Does STEM stand out? Examining racial/ethnic gaps in persistence across postsecondary fields. *Educational Researcher, 48*(3), 133–144.

Rogers-Sirin, L., & Sirin, S. R. (2009). Cultural competence as an ethical requirement: Introducing a new educational model. *Journal of Diversity in Higher Education, 2*(1), 19.

Rosenzweig, E. Q., Hecht, C. A., Priniski, S. J., Canning, E. A., Asher, M. W., Tibbetts, Y., Harackiewicz, J. M., et al. (2021). Inside the STEM pipeline: Changes in students' biomedical career plans across the college years. *Science Advances*, *7*(18), eabe0985.

Ryan, J., & Viete, R. (2009). Respectful interactions: Learning with international students in the English-speaking academy. *Teaching in Higher Education, 14*(3), 303–314.

Sawir, E. (2005). Language difficulties of international students in Australia: The effects of prior learning experience. *International Education Journal, 6*(5), 567–580.

Schinske, J. N., Perkins, H., Snyder, A., & Wyer, M. (2016). Scientist spotlight homework assignments shift students' stereotypes of scientists and enhance science identity in a diverse introductory science class. *CBE—Life Sciences Education*, *15*(3), ar47.

Seymour, E., Hunter, A. B., & Weston, T. J. (2019). Why we are still talking about leaving. In *Talking about leaving revisited* (pp. 1–53). Springer, Cham.

Suarez, N. A., Wang, S., Brydges, S., & Lo, S. M. (2022) *How higher education instructors conceptualize diversity: A phenomenographic study.* In C. Chinn, E. Tan, C. Chan, & Y. Kali (Eds.), *Proceedings of the 16th international conference of the learning sciences* - ICLS 2022 (pp. 287–296). International Society of the Learning Sciences.

Suárez-Orozco, C., Casanova, S., Martin, M., Katsiaficas, D., Cuellar, V., Smith, N. A., & Dias, S. I. (2015). Toxic rain in class: Classroom interpersonal microaggressions. *Educational Researcher, 44*(3), 151–160.

Theobald, E. J., Hill, M. J., Tran, E., Agrawal, S., Arroyo, E. N., Behling, S., & Freeman, S. (2020). Active learning narrows achievement gaps for underrepresented students in undergraduate science, technology, engineering, and math. *Proceedings of the National Academy of Sciences, 117*(12), 6476–6483.

Turner, K. L., Adams, J. D., & Eaton, S. E. (2022). Academic integrity, STEM education, and COVID-19: A call to action. *Cultural Studies of Science Education, 17*(2), 331–339.

Wade-Jaimes, K. (2021). 'STEM girls should be': A discourse analysis of school structures and their impact on African American, middle school girls' positioning in science. *Discourse: Studies in the Cultural Politics of Education, 42*(6), 828–839.

Xu, D., & Solanki, S. (2020). Tenure-track appointment for teaching-oriented faculty? The impact of teaching and research faculty on student outcomes. *Educational Evaluation and Policy Analysis, 42*(1), 66–86.

Zeidler, D. L. (2016). STEM education: A deficit framework for the twenty first century? A sociocultural socioscientific response. *Cultural Studies of Science Education, 11*(1), 11–26.