

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

Elizabeth S. Park
Mike Wilton
Stanley M. Lo
Natascha Buswell
Brian K. Sato

Working Paper #21-7

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

Elizabeth S. Park^{1*}, Mike Wilton², Stanley M. Lo³, Natascha Buswell⁴, Brian K. Sato^{5*}

¹ Center for Research and Evaluation at Extension, University of California San Diego, La Jolla, CA, 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 and ⁴ Program in Mathematics and Science Education, University of California San Diego, La Jolla, CA, USA.

⁴ Department of Mechanical and Aerospace Engineering, University of California Irvine, Irvine, CA, USA.

⁵ Department of Molecular Biology and Biochemistry, University of California Irvine; Irvine, CA, USA.

Email: lizpark@ucsd.edu, bsato@uci.edu

Classification: Social Sciences (Psychological and Cognitive Sciences)

Keywords: social psychology | STEM instruction | undergraduate education | racial equity

Abstract

Growing numbers of studies indicate that STEM faculty's beliefs about fixed ability and intelligence, or mindset beliefs, contribute to racial disparities in STEM course achievement and a lack of inclusion in STEM. In this work, we correlate fixed mindset beliefs to instructional factors, including faculty approaches to assessment, grading, and diversity, and link these approaches with student performance in the context of racial equity grade gaps. We do so using combined survey data on STEM faculty mindset and instructional approaches and student transcript data at a research-intensive, minority-serving institution. We find faculty with fixed mindset beliefs endorse more traditional approaches to assessment, grading, and diversity. In turn, the endorsement of these approaches is associated with lower grades and larger equity grade gaps, indicating that instructional approaches may be driving disparities in STEM achievement. Additionally, we identify a number of faculty characteristics, including gender, position type, and discipline that predict specific approaches. As opposed to focusing on student deficits, this work highlights the need for faculty to carefully consider their instructional approaches as we strive to create inclusive classroom spaces that foster the success of minoritized student populations.

Significance Statement

STEM faculty with fixed mindset beliefs are linked to racial disparities in undergraduate course achievement but mindset beliefs are likely only one of many instructional factors that contribute to racial equity grade gaps. This work reveals that leveraging assessments to keep students accountable, grading through a gatekeeping and tough love lens, and a belief that it is not the instructor's role to incorporate diversity are associated with fixed mindset beliefs and widening equity grade gaps. These results emphasize the importance of faculty decision-making processes on student success. It also places an onus on institutions to provide faculty with opportunities, training, and incentives to reflect and improve on their instructional approaches with the ultimate goal of increasing diversity in STEM.

Introduction

It is estimated that racially minoritized students (including Latinx, African-American, Native American, Southeast Asian, and Pacific Islander, hereafter referred to as “minoritized students”) underperform in STEM college courses and report disenchantment with STEM fields relative to students from represented backgrounds (i.e., White and East Asian, hereafter referred to as “represented” students) (1-2). These experiences contribute to minoritized students switching out of STEM programs at a 15-20% higher rate than those from represented backgrounds (3-4). Previous recommendations on how to minimize these disparate outcomes have focused on efforts to remediate student deficits, commonly framed as means to enable the success of underprepared students for the rigors of STEM environments (5-7). However, integral to our examination of STEM student success is the recognition that institutional policies and practices have given rise to, and perpetuate structures that may prevent the success of minoritized students.

Major contributors to these institutional structures are the faculty who shape STEM courses, degree programs, and broader university environments (8-10). Within the classroom, faculty are responsible for the course instructional practices and policies that can often serve as barriers to inclusion. For example, traditional approaches, such as lecturing, are associated with increased racial equity grade gaps (hereafter termed “equity grade gap”) between minoritized and represented students, relative to active learning pedagogies (11), while implementing assignments that integrate topics of diversity within STEM content improves science identity and interest in science (12). Despite the positive impact of certain evidence-based approaches being clear, not all faculty ascribe to the benefits, or choose to implement them. Faculty reluctance to adopt evidence-based approaches may be associated with the extent to which they hold fixed or growth mindset beliefs (13). Faculty who espouse fixed mindset beliefs endorse the idea that intelligence and ability are innate qualities that can only be minimally changed or developed (14). Mindset beliefs can then impact faculty and student interactions in the classroom. For example, minoritized students may be subject to unhelpful advising and/or lower expectations from their instructor based on a single test performance if said instructor believes that students’ abilities are fixed (15). Previous work has also shown that 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 phenomenon (16-18).

Faculty mindset beliefs are only one of many factors that result in barriers to inclusion in STEM within the classroom setting (Fig. 1). Other contributors include faculty approaches to assessment, grading, and diversity. For example, viewing course assessments as a means for accountability instead of learning negatively shapes student experience, particularly for minoritized students (19). This approach to assessment affects student engagement with learning and has significant impacts on all aspects of students’ experiences, including course outcomes and grades (20-22). Further, having a gatekeeping mentality (23) or using a tough-love approach to course grading (24) results in excluding and sorting students, respectively. Those who view themselves as gatekeepers use grades to sort and select students on the basis of achievement, while those who apply a tough love approach believe that tough grading will ensure retention of only the “most talented” individuals. Faculty conceptions of diversity also inform the way they approach learning and teaching in the classroom (25), which can ultimately influence student

outcomes, as ignoring diversity or “not seeing color” in the classroom leads to chilly climates (14, 26) that place minoritized students in disadvantaged positions (27).

Taken together, these studies indicate that instructors who harbor fixed mindset beliefs will employ instructional approaches that negatively affect students. Nevertheless, whether these instructional approaches contribute to racial disparities in course achievement remains unexplored. Therefore, we aim to determine (1) whether faculty mindset beliefs correlate with these approaches to grading, assessment, and diversity and (2) whether these approaches correlate with lower course grades and larger equity grade gaps, similar to what has been previously demonstrated with faculty mindset beliefs (14).

Results

Relationship between faculty mindset beliefs and approaches to assessment, grading, and diversity

In Fall 2020, faculty were recruited from Biological Sciences, Engineering and Computer Sciences, and Physical Sciences departments at a large, research-intensive, minority-serving institution to complete a survey that measures faculty approaches to grading, assessment, and diversity in the classroom (see Materials and Methods). We combined the survey data with university records of all students taught by faculty survey respondents during Fall 2016 to Fall 2019 (to capture recent student outcomes excluding terms affected by the COVID-19 pandemic). The merged data represent a 36% survey response rate and include 216 STEM faculty who taught 31,361 students in 1,446 courses (Table S1). 72% of the faculty survey respondents are men and about half of the respondents are at the Professor rank (versus Associate and Assistant Professors), both of which are comparable to national STEM faculty trends (28). Over 80% are research faculty (which we define as the traditional tenure-track, research-focused faculty member), with 10% tenure-track teaching-focused faculty (which we will refer to as “teaching faculty”) and 8% adjunct lecturers. In regard to the students taught by the survey participants during the Fall 2016 to Fall 2019 terms, 50% of students are women, 48% are racially minoritized students, one third are low-income students, and the average high school GPA is 3.9 (Table S2). In our analyses, we estimate crossed random effects models to account for the fact that students can take courses with multiple instructors and instructors can teach more than one course (29, 30).

As faculty mindset and approaches to assessment, grading, and diversity have all been shown to impact student outcomes, we sought to determine whether there is a relationship amongst these constructs for our survey participants (Table 1). We find that faculty who endorse fixed mindset beliefs were more likely to approach assessment as a means to enforce student accountability ($\beta = 0.24$ SD on a five point agree-disagree scale, $p < 0.001$), to use grading for gatekeeping purposes ($\beta = 0.28$ SD, $p < 0.001$), and to enact tough-love grading practices ($\beta = 0.35$ SD, $p < 0.001$). In addition, faculty who endorse fixed mindset beliefs were also less likely to advocate that STEM curriculum should promote diversity ($\beta = -0.17$ SD, $p < 0.01$) and that the instructor’s role is to promote equity in their discipline ($\beta = -0.17$ SD, $p < 0.01$).

Relationship between mindset beliefs, approaches and student outcomes

Before investigating whether a relationship between these approaches and student outcomes exists, we examined whether the mindset and student grade relationship identified by Canning et al. (14) was replicated in our sample. Similar to this work, students in courses taught by faculty who hold fixed mindset beliefs earned lower grades and experienced larger equity grade gaps relative to students in courses taught by faculty who hold growth mindset beliefs (Fig. 2, Table S3). Specifically, students, on average, underperform by 0.05 GPA points when taught by faculty who endorse fixed mindset beliefs ($p < 0.05$). In addition, represented students perform 0.04 GPA points higher ($p < 0.05$) but the equity grade gap increases to 0.07 GPA points in courses taught by faculty with fixed mindset beliefs (+1 SD: represented GPA: 2.86; minoritized GPA: 2.80). In courses taught by instructors who endorse growth mindset beliefs (1 SD below the average), the equity grade gap is just 0.01 GPA points.

We next investigated the relationship between faculty approaches regarding assessment, grading, and diversity and course grades, while controlling for faculty mindset beliefs (Fig. 2 and Table S4). We find that students, on average, perform 0.07 GPA points ($p < 0.05$) lower in courses taught by instructors who endorse assessment for accountability purposes relative to courses taught by instructors who do not. Similarly, students perform, on average, 0.11 GPA points ($p < 0.001$) lower in courses where instructors espouse gatekeeping approaches and 0.08 GPA points ($p < 0.01$) lower in courses where instructors espouse tough love. Aligning with our hypothesis regarding equity grade gaps, instructors' approaches disproportionately impact minoritized students. Represented students, on average, perform 0.03 GPA points higher ($p < 0.05$) but the gap increases to 0.08 GPA points in courses taught by instructors who advocate for assessment for accountability purposes (+1 SD: represented GPA: 2.86; minoritized GPA: 2.78). In regard to grading approaches, represented students perform 0.04 GPA points higher ($p < 0.01$) than minoritized students, but the gap increases to 0.09 in courses where faculty endorse gatekeeping approaches (+1 SD: represented GPA: 2.80; minoritized GPA: 2.72) and 0.07 GPA points ($p < 0.01$) in courses with tough love grading approaches (+1 SD: represented GPA: 2.85; minoritized GPA: 2.77).

In contrast, students earn *higher* grades if their instructors endorse the approach that faculty play a role in fostering acceptance of different beliefs and ideas as well as advocate for the importance of diversity in the curriculum. Students on average achieve 0.06 GPA ($p < 0.05$) points higher when their instructor fosters acceptance of diverse beliefs and ideas and agrees on the importance of diversity and equity in the curriculum. As for the equity grade gap, represented students on average achieve 0.04 GPA units higher ($p < 0.01$) but this gap shrinks to zero in courses where instructors advocate for diversity (+1 SD). We also observe this pattern in courses where instructors advocate for race and ethnic diversity to be more strongly reflected in the curriculum. Table S5 and S6 estimates the models presented above using fixed effects with similar results (31, 32).

Identifying faculty characteristics that predict mindset beliefs and approaches

Similar to previously published work (14), faculty type, rank, and discipline had no bearing on whether an individual holds fixed versus growth mindset beliefs. In our sample though, female faculty were significantly less likely to have a fixed mindset than male faculty ($\beta = -0.51$, $p < 0.01$) (Table 2). In regard to the instructional approaches, female faculty were also more likely to endorse the importance of diversity conversation in STEM curricula than male faculty ($\beta = 0.56$

SD, $p < 0.001$), less likely to believe that assessment should be used for accountability purposes ($\beta = -0.41$ SD, $p < 0.01$) or student improvement ($\beta = -0.37$ SD, $p < 0.05$), and less likely to take a tough love approach to grading ($\beta = -0.52$ SD, $p < 0.001$) (Table 2). Both research faculty and adjunct lecturers were more likely to hold assessment for accountability purposes views (research: $\beta = 0.52$ SD, $p < 0.01$; lecturer: $\beta = 0.67$ SD, $p < 0.05$) relative to teaching faculty. Comparing faculty rank, full professors were more likely to hold favorable perceptions of assessment ($\beta = -0.68$ SD, $p < 0.001$) and less likely to believe it is the instructor's role to foster diversity in the classroom ($\beta = -0.37$ SD, $p < 0.05$) relative to assistant professors. Looking at disciplinary differences, we note that Biological Sciences faculty compared to Engineering and Computer Science faculty were less likely to take a tough love and gatekeeping approach to grading, although these results are marginally significant ($p < 0.10$). Biological Sciences faculty were also more likely to support the idea that diversity should be promoted in the curriculum ($\beta = 0.38$ SD, $p < 0.05$) relative to their Engineering and Computer Sciences peers. In addition, Physical Sciences faculty were more likely to value assessment for accountability purposes ($\beta = 0.32$ SD, $p < 0.05$) relative to Engineering and Computer Sciences faculty.

Discussion

The present research complements the existing literature (14) by examining how fixed mindset beliefs connect with traditional approaches to assessment, grading, and diversity that ultimately result in racial inequalities in STEM courses. Specifically, leveraging assessments to keep students accountable as opposed to a means to provide formative feedback on their 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. These equity grade gaps may stem from how these traditional approaches likely contribute to and shape undergraduate students' classroom experiences. Empirical research describes the overt and subtle forms of discrimination against minoritized students in college classrooms, a phenomenon referred to as chilly classroom climate (33). Originally described in the context of women in STEM, such experiences have also been tied to students' personal identities including race (34,35) and sexual orientation (36), among others. For example, microaggressions, derogatory statements and invalidations, and fulfilling or contradicting a stereotype can contribute to the unwelcoming climate experienced (37,38). As students' sense of belonging in class is diminished, their motivation and academic performance is negatively impacted (39).

Analyses of faculty demographics revealed significant differences in mindset as well as faculty approaches. Female faculty possessed more of a growth mindset, expressed that assessment should not be used for student accountability, that grades should not be assigned based on tough love, and that the curriculum should incorporate diversity (Table 2). The finding that female faculty are likely to employ approaches that benefit minoritized students aligns with prior work that they are more likely to implement pedagogies that disproportionately benefit minoritized students (11, 40, 41). While this is a promising result, it also aligns with prior research showing that female, and other minoritized faculty, see increased burdens for diversity-related activities - including advising minoritized students and increased service loads, which lead to less time spent on research tasks (42). 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.

While female faculty felt that assessment should not be for student accountability purposes, they also were less likely to state that assessment is intended for student improvement. It may be that female faculty's conception of assessment is more tied to improving their teaching as opposed to how assessment can drive student learning (35). More research is needed to better understand how faculty's racial and gender identities interact with their perceptions and use of classroom assessment as well as how assessment is related to racial equity grade gaps in STEM classrooms.

We also found that instructor-type influenced faculty perspective on the purpose of assessment. Research faculty (82%) and adjunct lecturers (8%) were significantly more likely to endorse the approach of assessment as a mechanism of student accountability than tenure-track teaching-focused faculty (10%; Table 2). As tenure-track teaching faculty are hired on the basis of pedagogical expertise and knowledge (43, 44), it might be expected that they ascribe to approaches that are more focused on student learning and creating a more positive classroom environment. Conversely, individuals in this particular faculty position tend to be formally trained similarly to research faculty in their discipline which may explain why there were few differences overall amongst the measured approaches (43). We are currently working to better capture the background characteristics and conceptions of teaching and diversity of tenure-track teaching faculty relative to their research track and adjunct lecturer colleagues to better understand the origin of these approaches, classroom practices, and student outcomes (45).

In regard to discipline, Biological Sciences faculty were less likely to take tough love and gatekeeping approaches 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 (46) found that women are the majority 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 found that all STEM fields are similarly unsuccessful at graduating Black and Hispanic 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 racially minoritized students relative to represented STEM students. While research has identified the benefits of particular classroom practices (i.e. active learning), we also know that faculty choose to implement such practices based on their own personal opinions and experiences (47). This, in combination with our findings reinforces the importance of faculty reflection regarding their approaches and beliefs 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 (48) and the writing of reflective teaching statements (49), are means to transform the classroom space. We, and others (14, 50), propose that providing faculty with the training to constructively reflect on their beliefs about and approaches to teaching, could lead to changes in classroom policies and practices and ultimately greater inclusion in the classroom. 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 (51, 52). If institutions expect their faculty to reconsider their approaches, the appropriate reward structures must exist to enable them to prioritize such activities.

Materials and Methods

Sample

The study survey was sent out to 594 faculty across three STEM departments (Biological Sciences, Engineering and Computer Sciences, Physical Sciences) at a single, research-intensive, minority-serving institution. 259 of these individuals completed the full survey (43% response rate). When we merged faculty survey responses with the student transcript data, 43 faculty were removed as they did not teach an undergraduate-enrolling course during our study time-frame. Once we obtained university records of students taught by our survey respondents, we excluded students who took courses that were independent study or research courses and courses that were predominantly graded using the Pass/No Pass system. Finally, we removed repeat students from the sample. Our final sample therefore includes 216 STEM faculty who instructed 31,361 students in 1446 courses during the Fall 2016 to Fall 2019 academic terms. This time period was selected for the study to avoid COVID-19 affected academic terms.

Among the 31,361 students in our analytic sample, half are women and 48% of the students are identified as racially minoritized. 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 college students based on ethnicity and race (53). 47% of students are first-generation college students and a third of the students are low-income. The average weighted high school GPA of the students is 3.9.

To assess the racial equity grade gap, we examine the relationship between various measures and student grades in the course. University records provided course grades and they were recorded on a 0-4 scale (A/A+ = 4.0, A- = 3.7, B+ = 3.3, B = 3.0, B- = 2.7, C+ = 2.3, C = 2.0, C- = 1.7, D+ = 1.3, D = 1.0, D- = 0.7, F = 0.0).

Survey Constructs

We administered a survey to assess STEM faculty’s instructional approaches including assessment, grading, diversity in STEM curriculum, and mindset beliefs. The assessment items were separated into three subscales and the grading items were separated into two subscales. Upon conducting confirmatory factor analysis, we concluded that some of our constructs are better suited as sub-constructs. As such, the 15 assessment-related survey items were separated into three scales: assessment for accountability purposes, assessment for student improvement, and negative perceptions about assessment. The assessment for accountability scale focuses on whether students meet qualification standards; the assessment for student improvement scale examines whether assessment provides feedback to students; and negative perceptions of learning identifies instructors’ opinions about fairness of assessments. Similarly, the six grading-related survey items were separated into two sub-constructs: tough love and gatekeeping. The tough love grading scale asks instructors whether tough grading motivates students to perform

their best, and the gatekeeping scale asks instructors whether distributing A's to half of the class is unreasonable. We asked four questions related to instructor role diversity and another four questions related to curriculum diversity. All survey constructs, their scale fit, and psychometric properties are included in Table S7 and S8 (see Supplementary Text for more information). Furthermore, we plot a histogram of each survey construct to examine the distribution of faculty responses. While responses to the assessment and grading constructs are evenly distributed, the diversity constructs are slightly left-skewed, suggesting that respondents in our study sample, on average, are more likely to agree to statements regarding the importance of diversity in STEM (see Fig. S1-S5). The psychometric analyses were conducted using R version 4.0.3 (2020-10-10) and the lavaan package (54).

Empirical Strategy

We modeled crossed random effects to account for the fact that students can take courses with multiple instructors and instructors can teach more than one course (e.g., 14). For example, students in the same course also take courses with other instructors and therefore students are not neatly nested within courses and courses are not neatly nested within instructors (29, 30). We estimate the relationship between the standardized measure of faculty mindset and student grades, and include an interaction term on the right hand side of the equation to identify differential effects for racially minoritized students. We include a vector of student-level, course-level, faculty-level covariates, and a dummy variable for each year. The student-level covariates include: high school GPA, first-generation status, low-income status, SAT math, and SAT verbal, race, and gender. The course-level covariates include class size and a binary variable indicating whether the course is an upper division course or a lower division course. Finally, we include instructor level covariates such as faculty gender, rank (e.g., assistant), and type (e.g., research-track versus teaching-track faculty). We estimate a random intercept for each instructor and a separate random intercept for differential effect on grades by race within each course.

To examine whether the racial equity grade gaps are driven by faculty approaches associated with fixed mindset, we estimate the same cross random effects but replace fixed mindset beliefs with approaches regarding assessment, grading, and diversity. In estimating the relationship between approaches and grades, we also control for faculty's mindset beliefs to see whether the approaches independently explain racial equity grade gaps above and beyond what is explained by faculty mindset. The variance in grades across instructors within students is 0.11, and the variance in grades across courses within instructor and student is 0.06. Our intraclass correlation for level 2 (instructors) is 0.56 indicating substantial variation in student grades across instructors. All of the measures on instructional approaches were standardized. The cross random effects estimation were conducted using R version 4.0.3 (2020-10-10) and the lme4 package (55).

Acknowledgments

The authors would like to thank the survey participants for their time in completing the study survey. We would also like to acknowledge Nicole Suarez for construction of the diversity

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

References

1. E. McGee, "Black genius, Asian fail": The detriment of stereotype lift and stereotype threat in high-achieving Asian and Black STEM students. *AERA Open*. **4**, 1-16 (2018).
2. E. Q. Rosenzweig, C. A. Hecht, S. J. Priniski, E. A. Canning, M. W. Asher, Y. Tibbetts, J. S. Hyde, J. M. Harackiewicz, Inside the STEM pipeline: Changes in students' biomedical career plans across the college years. *Sci. Adv.* **7**, eabe0985 (2021).
3. C. Rieggle-Crumb, B. King, Y. Irizarry, Does STEM stand out? Examining racial/ethnic gaps in persistence across postsecondary fields. *Educ. Res.* **48**, 133–144 (2019).
4. S. J. Gates, C. Mirkin, Engage to excel. *Science*. **335**, 1545–1545 (2012).
5. S. R. Harper, An anti-deficit achievement framework for research on students of color in STEM. *New Dir. Ins. Res.* **2010**, 63–74 (2010).
6. E. L. Castro, "Underprepared" and "At-Risk": Disrupting deficit discourses in undergraduate STEM recruitment and retention programming. *J. Stud. Aff. Res. Pract.* **51**, 407–419 (2014).
7. D. L. Zeidler, STEM education: A deficit framework for the twenty first century? A sociocultural socioscientific response. *Cult. Stud. of Sci. Educ.* **11**, 11–26 (2016).
8. C. Haynes, L. D. Patton, From racial resistance to racial consciousness: Engaging white STEM faculty in pedagogical transformation. *J. Cases Ed. Lead.* **22**, 85–98 (2019).
9. T. L. Killpack, L. C. Melón, Toward inclusive STEM classrooms: What personal role do faculty play? *CBE Life Sci. Educ.* **15**, es3 (2016).
10. J. J. Park, Y. K. Kim, C. Salazar, S. Hayes, Student–faculty interaction and discrimination from faculty in STEM: The link with retention. *Res. High. Educ.* **61**, 330–356 (2020).
11. E. J. Theobald, M. J. Hill, E. Tran, S. Agrawal, E. N. Arroyo, *et al.*, Active learning narrows achievement gaps for underrepresented students in undergraduate science, technology, engineering, and math. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 6476–6483 (2020).
12. J. N. Schinske, H. Perkins, A. Snyder, M. Wyer, Scientist spotlight homework assignments shift students' stereotypes of scientists and enhance science identity in a diverse introductory science class. *CBE Life Sci. Educ.* **15**, ar47 (2016).
13. O. R. Aragón, S. L. Eddy, M. J. Graham, Faculty beliefs about intelligence are related to the adoption of active-learning practices. *CBE Life Sci. Educ.* **17**, ar47 (2018).
14. E. A. Canning, K. Muenks, D. J. Green, M. C. Murphy, STEM faculty who believe ability is fixed have larger racial achievement gaps and inspire less student motivation in their classes. *Sci. Adv.* **5**, eaau4734 (2019).
15. A. Rattan, C. Good, C. S. Dweck, "It's ok — Not everyone can be good at math": Instructors with an entity theory comfort (and demotivate) students. *J. Exp. Soc. Psychol.* **48**, 731–737 (2012).
16. K. Muenks, E. A. Canning, J. LaCosse, D. J. Green, S. Zirkel, J. A. Garcia, M. C. Murphy, 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. *J. Exp. Psychol. Gen.* **149**, 2119–2144 (2020).
17. J. LaCosse, M. Murphy, J. Garcia, S. Zirkel, The role of STEM professors' mindset beliefs on students' anticipated psychological experiences and course interest. *J. Educ. Psychol.* (2020).
18. S.-J. Leslie, A. Cimpian, M. Meyer, E. Freeland, Expectations of brilliance underlie gender distributions across academic disciplines. *Science*. **347**, 262–265 (2015).

19. G. F. Madaus, Clarke, M, The adverse impact of high stakes testing on minority students: evidence from 100 years of test data (2001).
20. H. Coates, The value of student engagement for higher education quality assurance. *Quality in Higher Education*. 11, 25–36 (2005).
21. M. Prosser, P. Ramsden, K. Trigwell, E. Martin, Dissonance in experience of teaching and its relation to the quality of student learning. *Studies in Higher Education*. 28, 37–48 (2003).
22. J. T. E. Richardson, Instruments for obtaining student feedback: a review of the literature. *Assess Eval. High. Edu.* 30, 387–415 (2005).
23. L. L. B. Barnes, K. S. Bull, N. J. Campbell, K. M. Perry, Effects of academic discipline and teaching goals in predicting grading beliefs among undergraduate teaching faculty. *Res. Higher Ed.* 42, 455–467 (2001).
24. L. L. B. Barnes, Development of the faculty beliefs about grades inventory. *Educ. Psychol. Meas.* 57, 459–468 (1997).
25. S. Gordon, A. Reid, P. Petocz, Educators' conceptions of student diversity in their classes. *Studies in Higher Education*. 35, 961–974 (2010).
26. C. Suárez-Orozco, S. Casanova, M. Martin, D. Katsiaficas, V. Cuellar, N. A. Smith, S. I. Dias. Toxic rain in class: Classroom interpersonal microaggressions. *Educ. Res.* 44 (3), 151-160 (2015).
27. P.L Carter, R. Skiba, M. I. Arredondo, M. Pollock, You can't fix what you don't look at: Acknowledging race in addressing racial discipline disparities. *Urban Educ.* 52 (2), 207-235 (2017).
28. National Center for Education Statistics Condition of Education, "Characteristics of postsecondary faculty", (available at <https://nces.ed.gov/programs/coe/indicator/csc>).
29. S. W. Raudenbush, A. S. Bryk, *Hierarchical Linear Models: Applications and Data Analysis Methods* (SAGE, 2002).
30. S. Rabe-Hesketh, A. Skrondal, *Multilevel and Longitudinal Modeling using STATA*. (Stata Press, Texas, ed. 3, 2012).
31. C. T. Clotfelter, H. F. Ladd, J. L. Vigdor, Teacher credentials and student achievement in high school a cross-subject analysis with student fixed effects. *J. Hum. Resour.* 45(3), 655-681 (2010).
32. R. W. Fairlie, F. Hoffmann, P. Oreopoulos, A community college instructor like me: Race and ethnicity interactions in the classroom. *Am. Econ. Rev.* 104(8), 2567-91 (2014).
33. J. J. Lee, J. M. McCabe, Who speaks and who listens: Revisiting the chilly climate in college classrooms. *Gender & Society*. 35, 32–60 (2021).
34. D. Solorzano, M. Ceja, T. Yosso, Critical race theory, racial microaggressions, and campus racial climate: The experiences of African American college students. *The Journal of Negro Education*, 69, 70-73.
35. D. Ogunyemi, C. Clare, Y. M. Astudillo, M. Marseille, E. Manu, S. Kim, Microaggressions in the learning environment: A systematic review. *J. Divers. High. Educ.* 13, 97–119 (2020).
36. A. Vaccaro, R. M. Koob, A critical and intersectional model of LGBTQ microaggressions: Toward a more comprehensive understanding. *Journal of Homosexuality*. 66, 1317–1344 (2019).
37. C. Harrison, K. D. Tanner, Language Matters: Considering Microaggressions in Science. *CBE Life Sci. Educ.* 17, fe4 (2018).
38. C. Suárez-Orozco, S. Casanova, M. Martin, D. Katsiaficas, V. Cuellar, N. A. Smith, S. I. Dias, Toxic rain in class: Classroom interpersonal microaggressions. *Educ. Res.* 44, 151–160 (2015).

39. T. M. Freeman, L. H. Anderman, J. M. Jensen, Sense of belonging in college freshmen at the classroom and campus levels. *J. Exp. Educ.* **75**, 203–220 (2007).
40. S. Hurtado, K. Eagan, J. H. Pryor, H. Whang and S. Tran. “Undergraduate Teaching Faculty: The 2013–2014 HERI Faculty Survey”, The Higher Education Research Institute (2011).
41. S. Freeman, S. L. Eddy, M. McDonough, M. K. Smith, N. Okoroafor, H. Jordt, M. P. Wenderoth, Active learning increases student performance in science, engineering, and mathematics. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 8410–8415 (2014).
42. K. O’Meara, A. Kuvaeva, G. Nyunt, C. Waugaman, R. Jackson, Asked more often: Gender differences in faculty workload in research universities and the work interactions that shape them. *Am. Educ. Res. J.* **54**, 1154–1186 (2017).
43. A. Harlow, S. M. Lo, K. Saichaie, B. K. Sato, Characterizing the University of California’s tenure-track teaching position from the faculty and administrator perspectives. *PLoS One.* **15**, e0227633 (2020).
44. A. Harlow, N. Buswell, S. M. Lo, B. K. Sato, paper presented at the Annual Meeting of the American Educational Research Association, Virtual Meeting, April 2021.
45. N. A. Suarez, S. Wang, S. Brydges, S. M., Lo, paper presented at the Annual Meeting of the American Educational Research Association, Virtual Meeting, April 2021.
46. R. Fry, B. Kennedy, C. Funk, “STEM Jobs See Uneven Progress in Increasing Gender, Racial and Ethnic Diversity” (Pew Research Center Science & Society, 2021). [available at <https://www.pewresearch.org/science/2021/04/01/stem-jobs-see-uneven-progress-in-increasing-gender-racial-and-ethnic-diversity/>].
47. T. C. Andrews, P. P. Lemons, It’s personal: Biology instructors prioritize personal evidence over empirical evidence in teaching decisions. *CBE Life Sci. Educ.* **14**, ar7 (2015).
48. A. L. Beach, M. D. Sorcinelli, A. E. Austin, J. K. Rivard, *Faculty Development in the Age of Evidence: Current Practices, Future Imperatives* (Stylus Publishing, LLC, 2016).
49. H. Hubball, J. Collins, D. Pratt, Enhancing reflective teaching practices: Implications for faculty development programs. *Can. J. Higher Ed.* **35** (3), 57-81 (2005).
50. B. Mitchneck, J. L. Smith, M. Latimer, A recipe for change: Creating a more inclusive academy. *Science*, **352** (6282), 148-149 (2016).
51. P. Hardré, M. Cox, Evaluating faculty work: Expectations and standards of faculty performance in research universities. *Res. Pap. Educ.* **24** (4), 383-419 (2009).
52. S. E. Brownell, K. D. Tanner, Barriers to faculty pedagogical change: Lack of training, time, incentives, and... tensions with professional identity? *CBE Life Sci. Educ.* **11**(4), 339-346 (2012).
53. D.J. Asai. Race matters. *Cell.* **181**, 754-757 (2020).
54. Y. Rosseel, lavaan: An R package for Structural Equation Modeling. *J. Stat. Softw.* **48**, 1–36 (2012).
55. D. Bates, M. Mächler, S. Walker, Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* **67**, (2017).

Figures and Tables

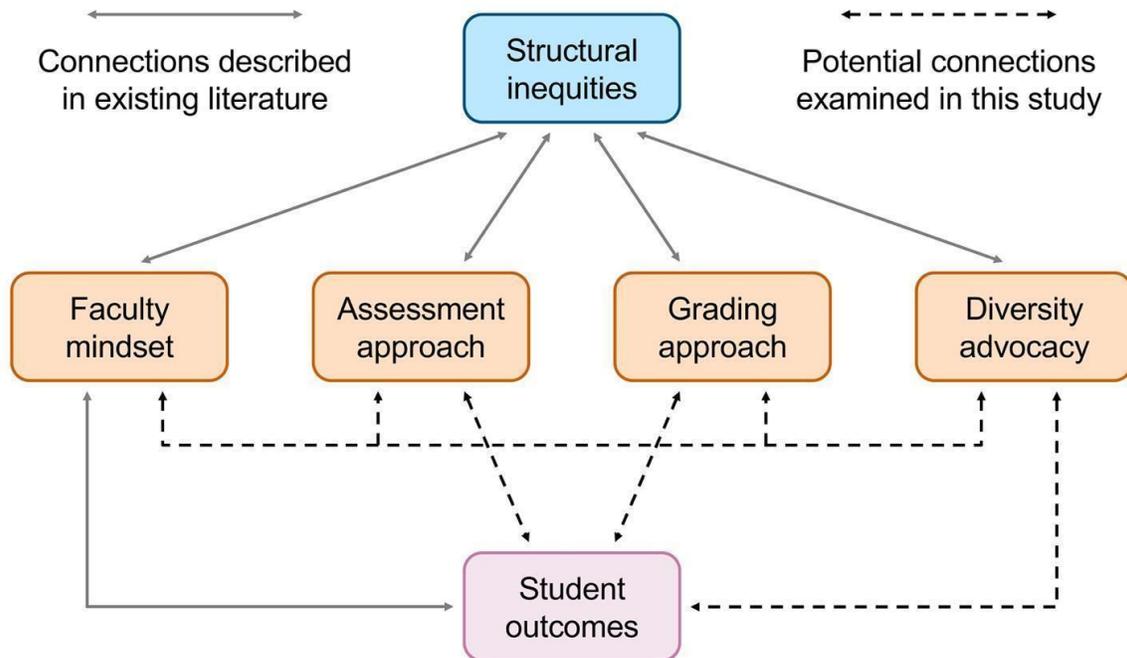


Fig. 1. Summary of potential correlations examined in this study. Faculty mindset beliefs and other factors including approaches to assessment, grading, and diversity are connected with structural inequities in undergraduate STEM education. Faculty mindset beliefs have further been shown to correlate with student outcomes in terms of equity grade gaps. Therefore, we hypothesize that (1) faculty mindset beliefs correlate with approaches to grading, assessment, and diversity and (2) these approaches correlate with equity grade gaps.

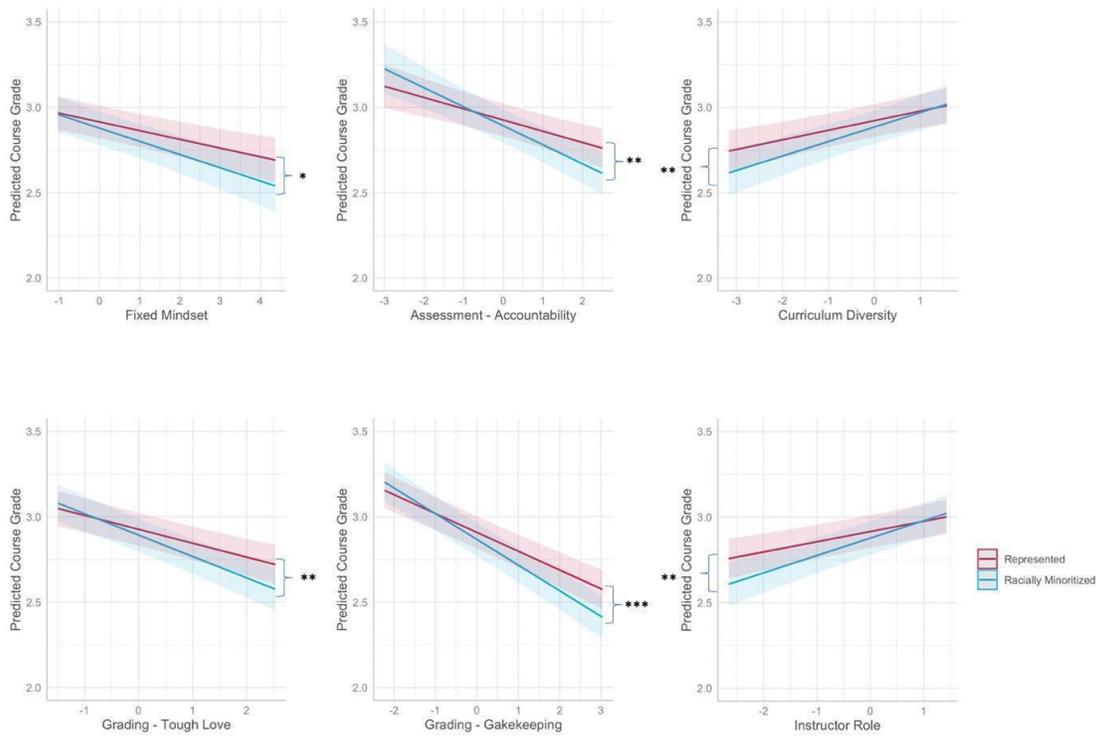


Fig. 2. Predicted values of course grade based on faculty approaches. Cross random effects model results of predicted student grade. Shaded regions indicated +/- 1 standard error. X-axis indicates the number of standard deviations above or below the mean of faculty mindset or approach (0 indicates the mean). Racially minoritized students are defined as Black, Latinx, Native American, Southeast Asian and Pacific Islander.

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

Table 1. Faculty mindset is associated with instructional approaches

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Assessment Approaches			Diversity Advocacy		Grading Approaches	
Variables	Student Accountability	Student Improvement	Negative Perceptions	Curriculum	Instructor Role	Tough Love	Gatekeeping
Fixed Mindset	0.241*** (0.051)	0.027 (0.071)	0.006 (0.065)	-0.166** (0.062)	-0.168** (0.057)	0.346*** (0.052)	0.284*** (0.064)
Faculty characteristics	x	x	x	x	x	x	x
R-squared	0.223	0.071	0.077	0.144	0.103	0.265	0.152
Observations	216	216	216	216	216	216	216

Each column is a separate regression estimate predicting the seven faculty approaches related to assessment, grading, and diversity advocacy. The specific covariates included in each estimation are: gender, type (research faculty, teaching professors, lecturers), rank, and department. Standard errors are in parentheses.

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

Table 2. Faculty characteristics that predict mindset beliefs and approaches to assessment, grading, and diversity

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Mindset	Assessment Approaches			Diversity Advocacy		Grading Approaches	
Variables	Fixed Mindset	Student Accountability	Student Improvement	Negative Perceptions	Curriculum	Instructor Role	Tough Love	Gatekeeping
Gender								
Women	-0.510**	-0.411**	-0.368*	0.220	0.557***	0.101	-0.524***	-0.293+
	(0.170)	(0.131)	(0.174)	(0.158)	(0.155)	(0.141)	(0.140)	(0.164)
Type								
Research faculty	0.308	0.521**	-0.019	0.047	-0.007	-0.181	0.254	0.306
	(0.255)	(0.196)	(0.261)	(0.238)	(0.232)	(0.212)	(0.210)	(0.245)
Lecturer	0.068	0.669*	-0.002	-0.307	-0.154	-0.319	0.236	0.246
	(0.369)	(0.284)	(0.378)	(0.344)	(0.336)	(0.307)	(0.304)	(0.355)
Rank								
Associate	-0.115	-0.257	-0.417+	-0.411+	-0.207	-0.294	-0.051	0.387
	(0.245)	(0.189)	(0.251)	(0.229)	(0.223)	(0.204)	(0.202)	(0.236)
Full	0.090	-0.242	-0.242	-0.675***	-0.242	-0.365*	0.017	0.160
	(0.209)	(0.161)	(0.214)	(0.195)	(0.190)	(0.174)	(0.172)	(0.201)
Department								
Biological Sciences	0.045	0.032	-0.138	0.065	0.379*	0.257	-0.294+	-0.325+
	(0.202)	(0.155)	(0.207)	(0.189)	(0.184)	(0.168)	(0.166)	(0.195)

Physical Sciences	-0.108	0.324*	0.315+	-0.003	-0.058	-0.129	-0.194	0.153
	(0.170)	(0.131)	(0.174)	(0.159)	(0.155)	(0.141)	(0.140)	(0.164)
R-squared	0.073	0.139	0.070	0.077	0.114	0.107	0.071	0.065
Observations	216	216	216	216	216	216	216	216

Each column is a separate regression estimate predicting mindset beliefs and the seven faculty approaches related to assessment, grading, and diversity advocacy, controlling for faculty gender, type (research faculty, teaching professors, lecturers), rank, and department as covariates. Reference categories are: men, assistant professor, Engineering/ICSCI, teaching professor. The outcome measures are all standardized. Standard errors are in parenthesis.

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