

UC Irvine

UC Irvine Previously Published Works

Title

A Multi-institutional Cluster Analysis to Identify Groups of Courses with Exemplary Opportunity Gaps for Undergraduate Students in the Biological Sciences.

Permalink

<https://escholarship.org/uc/item/8jf2k8bd>

Journal

CBE : Life Sciences Education, 23(4)

Authors

Denaro, Kameryn

Molinaro, Marco

Fiorini, Stefano

et al.

Publication Date

2024-12-01

DOI

10.1187/cbe.24-02-0051

Peer reviewed

A Multi-institutional Cluster Analysis to Identify Groups of Courses with Exemplary Opportunity Gaps for Undergraduate Students in the Biological Sciences

Kameryn Denaro,^{†*} Marco Molinaro,[‡] Stefano Fiorini,[§] Rebecca L. Matz,^{||} Chris Mead,[¶] Meryl Motika,[#] Nita Tarchinski,[@] Montserrat Valdivia Medinaceli,^{**} W. Carson Byrd,^{††} Benjamin Koester,^{††} Hye Rin Lee,^{§§} Timothy McKay,^{|||} and Brian K Sato^{¶¶}

[†]Division of Teaching Excellence and Innovation, University of California, Irvine, CA 92697;

[‡]Teaching and Learning Transformation Center - Educational Effectiveness, University of Maryland, College Park, MD 20742; [§]Research And Analytics, Institutional Analytics, UITS, Indiana University, Bloomington, IN 47401; ^{||}Center for Academic Innovation, University of Michigan, Ann Arbor, MI 48104; [¶]School of Earth and Space Exploration, Arizona State University, Tempe, AZ 85287; [#]Office of the Vice Chancellor for Student Affairs, University of California, Berkeley, CA 94720; [@]Undergraduate Education, College of Literature, Science, and the Arts, University of Michigan, Ann Arbor, MI 48109; ^{**}Neurobiology, Physiology and Behavior, University of California, Davis, CA 95616; ^{††}Center for the Study of Higher and Postsecondary Education, University of Michigan, Ann Arbor, MI 48109; ^{††}Center for the Study of Higher and Postsecondary Education, University of Michigan, Ann Arbor, MI 48109; ^{§§}College of Education and Human Development, University of Delaware, Newark, DE 19716; ^{|||}Departments of Physics and Astronomy, College of Literature, Science, and the Arts, and School of Education, University of Michigan, Ann Arbor, MI 48109; ^{¶¶}Department of Molecular Biology and Biochemistry and Division of Teaching Excellence and Innovation, University of California, Irvine, CA 92697

ABSTRACT

Examining institutional data from seven cohorts of students intending to major in biology across five research-intensive institutions, this work analyzes opportunity gaps—defined as the difference between the grade received by students from the dominant and nondominant sociodemographic groups in institutions of higher education—at the course-section level across mathematics, physics, biology, and chemistry disciplines. From this analysis, we find that the majority of course sections have large opportunity gaps between female and male students, students who are Black, Latino/a/e/x, or indigenous to the United States and its territories and students who are White or Asian, first-generation and non-first-generation students, and low-income and non-low-income students. This work provides a framework to analyze equity across institutions using robust methodology, including: using multiple approaches to measure grades, quantile regression rankscores which adjust for previous academic performance, and cluster analysis. Recommendations are provided for institutions to identify faculty who have equitable course sections, automate equity analyses, and compare results to other institutions to make a change toward more equitable outcomes.

INTRODUCTION

Inequitable grade outcomes in introductory science, technology, engineering, and mathematics (STEM) courses are well documented in discipline-based education research (Matz *et al.*, 2017; Salehi *et al.*, 2019; Borda *et al.*, 2020; Mead *et al.*, 2020; Castle *et al.*, 2024). In recent years, research across institutions has

Jennifer Knight, *Monitoring Editor*

Submitted Feb 5, 2024; Revised Sep 3, 2024; Accepted Sep 12, 2024

CBE Life Sci Educ December 1, 2024 23:ar53

DOI:10.1187/cbe.24-02-0051

Conflicts of interest: The authors declare no conflicts of interest.

*Address correspondence to: Kameryn Denaro (kdenaro@uci.edu).

© 2024 K. Denaro *et al.* CBE—Life Sciences Education © 2024 The American Society for Cell Biology. This article is distributed by The American Society for Cell Biology under license from the author(s). It is available to the public under an Attribution–Noncommercial–Share Alike 3.0 Unported Creative Commons License (<http://creativecommons.org/licenses/by-nc-sa/3.0>).

“ASCB®” and “The American Society for Cell Biology®” are registered trademarks of The American Society for Cell Biology.

been conducted to determine whether the inequities present at one institution hold true at other institutions (Freeman *et al.*, 2014; Matz *et al.*, 2017; Mead *et al.*, 2020; Hatfield *et al.*, 2022; Fiorini *et al.*, 2023; Fischer *et al.*, 2023). Further, inequitable outcomes for students from historically minoritized populations (students who are Black, Latino/a/x, Pacific Islanders, and indigenous to the United States and its territories; first-generation college students; low-income students; and women) can be seen in exam performance (Eddy *et al.*, 2014), persistence in STEM (Griffith, 2010), and STEM undergraduate, doctorates, and workforce (Allen-Ramdial and Campbell, 2014). And for those students who leave STEM, they experience lower grade outcomes in their first-year STEM courses compared with those who stay in STEM (Chen and Ho, 2012). These inequities manifest themselves at the individual course level and persist through graduation, where the proportion of students of a given subgroup that start and complete in their STEM major are substantially lower for minoritized populations (National Science Foundation, 2021).

The inequities for historically minoritized populations are pervasive in the higher education system (Chen, 2009, 2013; Cataldi *et al.*, 2018; Canning *et al.*, 2019; National Science Foundation, 2021). For example, while the share of STEM degrees awarded to people traditionally excluded due to ethnicity or race (PEERs) (Asai, 2020) over the past 10 years has increased (National Science Foundation, 2021), we still do not know how pervasive student grade equity issues are, and how they vary across institutions. PEERs disproportionately experience negative outcomes across a wide range of metrics including course grades, grade point average (GPAs), graduation rates, and retention in STEM (Ainsworth-Darnell and Downey, 1998; Kao and Thompson, 2003; Bécares and Priest, 2015; Domina *et al.*, 2017; Tatum, 2017; National Science Foundation, 2021), inequities caused by an educational system that is infused with institutional racism (Ladson-Billings, 1995, 2009, 2020; Solórzano and Villalpando, 1998; McCoy and Rodricks, 2015; Patton, 2016; McGee, 2020; Taylor *et al.*, 2023). Similarly, first-generation college students are less likely to earn a bachelor's degree compared with non-first-generation college students (Chen, 2009; Cataldi *et al.*, 2018), as are women compared with men (National Science Foundation, 2021), and low-income students relative to non-low-income students (Chen, 2013).

One barrier to persistence in STEM is one or more negative academic outcomes in individual courses (Chen, 2013) with first-term academic performance serving as a primary predictor of graduation success (Gershenfeld *et al.*, 2016). While active learning approaches have been shown to reduce achievement gaps for minoritized students (Theobald *et al.*, 2020), it has been shown that adoption of student-centered instructional practices vary across discipline and institution (Stains *et al.*, 2018). As such, the student experience is likely to be highly inconsistent from course to course, impacting the degree to which equitable course outcomes occur. Identifying the persistence of opportunity gaps—defined as the difference between the grade received by students from the dominant and nondominant sociodemographic groups—across STEM disciplines and student demographics will greatly help to inform this national discussion.

Theoretical Framework

This research is guided by the critical social and critical race frameworks. Critical theoretical perspectives guide interpretation of grade disparities seen throughout STEM programs. In the current study, we explore opportunity gaps in STEM courses, considering the impact of racism, sexism, and classism on imparting different learning contexts for students. This analysis leverages institutional data to evaluate equity while viewing the results through a critical lens. Careful consideration of student grade outcomes in relation to educational contexts can assist with limiting their misinterpretation. The adoption of a critical perspective supports reframing the individual experience (e.g., an outcome in a course) as stemming from the constraints of the surrounding social and cultural context and that society has produced a university structure which disadvantages people “who are not White, cisgendered, male, heterosexual, able-bodied, wealthy, and Western individuals” (Pearson *et al.*, 2022, p.2). Drawing on critical race theory allows us to attend to the persistent system of racism ingrained in the structures of American society, and in educational organizations such as schools and colleges in particular, that can produce inequitable STEM learning environments (Freeman, 1977, 1988; Bell, 1995; Ladson-Billings, 1998, 2009; Taylor *et al.*, 2009; Patton, 2016; McGee, 2020). Critical social theory and reproduction theories also allow us to focus on issues of power in relation to gender and socioeconomic status as well (Bourdieu and Passeron, 1990; Bourdieu, 1998; Manias and Street, 2000; Leonardo, 2004; Kincheloe and McLaren, 2011).

A consistent thread found across these critical theoretical perspectives is that individual intent and explicit discriminatory acts are not necessary to the maintenance of inequitable learning environments and to perpetuate grade disparities. As a systemic feature of society and organizations, everyday policies and practices can reproduce racial, gender, and socioeconomic disparities without identifiable racist, sexist, or elitist individuals (Acker, 1990; Bourdieu and Passeron, 1990; Bourdieu, 1998; Ray, 2019). These organizationally- and disciplinarily-embedded features shape the differential experiences and outcomes of students and researchers in STEM fields (Asai, 2020; McGee, 2020; Fischer *et al.*, 2023). Common, normalized features of STEM classrooms reflect the maintenance of inequitable educational contexts students must navigate. What the assumed common educational trajectories of students are when they enter the classroom, what instructors and departments need to provide for them to succeed, and why students may not perform well are rooted in the systemic inequities of postsecondary institutions and the instructor's beliefs about success and failure that reinforce racial, gender, and socioeconomic inequalities in education.

Our goal is to consider the grade inequities among biological sciences majors of multiple demographic groups (including race/ethnicity, first-generation status, and low-income status) across multiple STEM courses spanning five research-intensive universities. Examining differences in opportunity gaps across disciplines, institutions, and course characteristics, can assist future researchers to begin to tease out the impacts of the structures and systemic barriers that these courses are situated in that may be causing persistent inequitable outcomes. We

specifically focus on students majoring in biological sciences degree programs as they are typically high-enrollment majors (National Science Foundation, 2021) and more diverse with respect to gender and race/ethnicity compared with other STEM disciplines (Funk, 2021).

RESEARCH QUESTIONS

Institutional data from five public research universities are used to calculate the opportunity gaps in biology, chemistry, mathematics, and physics courses among undergraduate biological sciences students. Specifically, we address the following research questions:

1. How do opportunity gaps vary across STEM courses among biological sciences majors?
2. How similar are these opportunity gaps across institutions?
3. What is the relationship between the observed opportunity gaps and discipline, institution, and course characteristics?

MATERIALS AND METHODS

Context

This study was conducted within the Sloan Equity and Inclusion in STEM Introductory Courses (SEISMIC) collaboration, a multi-institutional research and practice endeavor across 10 public research-intensive universities in the United States that connects individuals across disciplines and institutional roles who are committed to making introductory STEM courses more equitable and inclusive (SEISMIC Overview, SEISMIC Collaboration, n.d.). A central goal of SEISMIC is to examine institutional data across member institutions, leveraging parallel analyses to understand the pervasiveness of inequity across disciplines. Based on interest and access to institutional data, five SEISMIC institutions were included in this study. The institutions each educate about 30,000 undergraduate students on average, one is a Hispanic Serving Institutions (HSI) and two are dual HSIs and Asian American and Native American Pacific Islander-Serving Institutions.

Data Collection

In this study, we conducted a retrospective cohort study (also known as a historical cohort study) (Sedgwick, 2014). We selected the cohorts by identifying full-time, first-year, nontransfer students who were majoring in the biological sciences at the time of admission into each university over a 7-year period (fall cohorts from 2013 to 2019) and followed each cohort for 2 years. Data for this project came from admissions and registrar records at each institution. For each student, we included their entry term to the university, major at the time of admissions, demographic variables (gender, race/ethnicity, first-generation status, and low-income status) and transcript data (term, course section, grade in course). The demographic variables, defined further in Table 1, are limited to binary categorizations due to institutional data availability and to meet minimum sample size requirements for statistical purposes. Table 2 presents the percentage of students in each demographic category averaged across the course sections for each institution.

Students Majoring in the Biological Sciences. This study specifically focuses on students intending to major in the biological sciences. These students enroll in similar courses across the five participating campuses, including biology, chemistry, mathematics, and physics courses. While the institutions in the SEISMIC collaboration have similar organizational features (i.e., public, large undergraduate student enrollments and “R1” or very high research activity) (McCormick, 2001), the requirements for major declaration vary; for the institutions in this study most require the selection of a major at entry, but one institution does not. For the one institution in this study that does not require declaration of a major at admissions, we defined majoring in the biological sciences as those who took at least one biology course and at least one chemistry, mathematics, physics, or statistics course in their first year because this aligns with biological sciences major course-taking patterns.

Biology, Chemistry, Mathematics, and Physics Course Sections. The dataset includes biology, chemistry, mathematics, and physics course sections that have at least 20 students majoring in biological sciences enrolled. Course sections are the unit of analysis for this study and are defined at the lecture level. For example, an introductory biology course might have 1200 students enrolled in one term, split into four 300-person course sections, and then further divided into smaller discussion sections. In this example, this would equate to four course sections. We use the course section-level data since this is the level at which grades are typically decided, students take exams, and students earn credits toward their degree.

Exclusion Criteria. Course sections worth less than three credit hours were excluded from the analysis. The minimum credit hours of the course sections was used to limit our study to the course sections which the biology students take as the core requirements for their major and exclude courses that are supplemental (i.e., supplemental instruction, seminar courses, etc.). For each course section, if there were fewer than 20 students majoring in biological sciences, the course section was excluded. Further, if there were less than five enrolled students majoring in biological sciences from any subgroup (e.g., less than 5 students who are female, male, PEER, non-PEER, first-generation, non-first-generation, low-income, or non-low-income), the course section was excluded. This minimum sample size per course section was used to ensure that assumptions of the statistical tests were met.

The following students were excluded prior to checking for minimum sample size requirements for each course section: 1) students who withdrew from a course, 2) students taking the course under a pass/no pass grading basis, and 3) transfer students. The first exclusion was made since these students' overall GPA would not be impacted by withdrawing from a course and at most institutions students are only allowed to withdraw from all their courses in a particular term. The second exclusion was made since students intending to major in biological sciences should be taking the core courses for a grade in order for the course to count toward their major requirements. The third exclusion was made as transfer students would not equally be enrolled across the disciplines examined; since the

TABLE 1. Description of demographic variables. Description of the demographic variables and how data were collected across the SEISMIC institutions

Indicator variable	Codes	Notes
Female	1 = female 0 = not female	Sex is self-reported by the students on a binary basis at three of the five institutions. Only two institutions collect information on gender identity in addition to sex.
Persons Excluded because of their Ethnicity or Race (PEER)	1 = students who identify as Black, Latinx, Pacific Islander, and/or indigenous to the United States and its territories 0 = students who identify as White or Asian only	PEER status (Asai, 2020) is used rather than individual race/ethnicity categories to ensure a large enough sample size within each course section.
First-Generation	1 = first-generation 0 = not first-generation	First-generation status includes students who self-reported that neither parent graduated from a 4-year university.
Low-Income	1 = low-income 0 = not low-income	

TABLE 2. Demographic representation of biology students across these SEISMIC institutions. The percent of students in each category are averaged over the respective course sections

	Institution				
	1 %	2 %	3 %	4 %	5 %
Gender					
Female	65	59	69	67	65
Non-female	35	41	31	33	35
Ethnicity					
PEERs	35	17	22	37	10
Non-PEERs	65	83	78	63	90
First-Generation (FG) Status					
FG	28	16	38	50	12
Non-FG	72	84	62	50	88
Low-Income (LI) Status					
LI	43	24	29	36	11
Non-LI	57	76	71	64	88

coursework for some disciplines may have already been completed at a community college. Rather than excluding students with missing demographic information, students with missing demographic information were assumed to be part of the dominant group (i.e., a student who is male, non-PEER, non-first-generation, or non-low-income).

Data Analysis

To calculate the opportunity gaps between two subgroups (females vs. males, PEERs vs. non-PEERs, first-generation vs. non-first-generation, and low-income vs. non-low-income), we took three different approaches. The first two approaches relied solely on the transcript data and paralleled the work of Denaro *et al.* (2021), while the third approach examined course grades while taking into account previous academic performance (including high school GPA, incoming credits from advanced placement exams, international baccalaureate exams, and any college credits earned while in high school) using quantile regression (QR) and normalized regression rankscores (NRR). The three approaches can be summarized

as follows: **Approach #1:** we calculated the difference in the fraction of A and B grades awarded compared with C, D, and F grades awarded in each subgroup ($\Delta\%AB$). **Approach #2:** we calculated the difference in the average grade received by each subgroup on a 4.0 scale (ΔGP). **Approach #3:** we carried out the QR procedure to calculate the NRR for each student in a course section and then calculated the difference in the average rankscores in each subgroup (ΔNRR). Detailed descriptions of the three approaches are provided in the Supplemental Materials (Supplemental Table S1). While, the term “opportunity gap” can describe any of the approaches, we focus our findings for RQ1 and RQ2 on approach #3 and provide results for approach #1 and #2 in the Supplemental Materials. For RQ3, information from all three approaches is used in the cluster analysis, as described below.

Data Sharing. Initial analyses using the student-level data were performed independently by researchers at each institution using shared code and was carried out in R (R Core Team, 2020). The student-level datasets include 292,591 student enrollments. All data identified at the student level were maintained at each home institution and never shared. For every course section, each institution provided the differences in academic performance ($\Delta\%AB$, ΔGP , and ΔNRR) for the four different demographic comparisons: 1) females versus males, 2) PEERs versus non-PEERs, 3) first-generation versus non-first-generation, and 4) low-income versus non-low-income. Only the course section-level data were shared for joint analysis.

QR and NRRs. Linear QR and other quantile methods were developed by Koenker (2000, 2005) and appear in many applications across disciplines (Casady and Cryer, 1976; Portnoy and Koenker, 1997; Eide and Showalter, 1998; He and Shi, 1998; Zhou and Portnoy, 1998; Møller *et al.*, 2008; Daouia *et al.*, 2013; Zhang *et al.*, 2017; Denaro *et al.*, 2021; Xiong and Tian, 2021). QR is particularly useful when the assumption of normality for ordinary least squares regression is violated (i.e., the response variable is non-normal or asymmetric); it is a robust alternative to ordinary least squares regression and does not require normality

among the error terms (Denaro *et al.*, 2021). In our case, the response variable (course grade) tends to be skewed and non-normal. We used QR to estimate the conditional quantiles of the response variable (course grade) given a set of predictor variables (previous academic performance). Unlike traditional regression methods that focus on estimating the conditional mean of the response variable, QR provides a more comprehensive picture of the relationship between the response and predictor variables by estimating multiple quantiles of the response variable. The quantiles range from zero to one and are equally spaced, with the QR model for the 0.50 quantile being the median regression model. After fitting the QR models, we leveraged the NRRs to compare the performance of subgroups of students. The NRRs provide a continuous measure of student performance that accounts for previous academic performance (high school GPA and incoming credits from advanced placement exams, international baccalaureate exams, and college credits earned while in high school). The difference in the NRRs for each subgroup provides a continuous measure of grade equity which adjusts for previous academic performance (note see Supplemental Materials for a discussion of the choice in approaches). The estimates of the quantile-specific regression parameters were found using the *quantreg* package in R (Koenker, 2015). For further details on the QR procedure, see Supplemental Materials. Following the fitting of the QR models, we leverage the opportunity gaps ($\Delta\%AB$, ΔGP , and ΔNRR) to compare the performance of the subgroups of students by using the following clustering algorithms.

Algorithms for Clustering. Cluster analysis is appropriate when there is no response variable of interest and a researcher would like to identify groups of observations (Fisher, 1958; Macqueen, 1967; Hartigan and Wong, 1979a, 1979b; Pollard, 1981; Rousseeuw and Kaufman, 1987; Hastie *et al.*, 2001; Denaro *et al.*, 2021). We conducted hierarchical clustering with complete linkage to group the course sections into statistically homogeneous clusters (Ng and Han, 2002; Kaufman and Rousseeuw, 2005) using the opportunity gaps ($\Delta\%AB$, ΔGP , and ΔNRR) for each pair of subgroups (females vs. males, PEERs vs. non-PEERs, first-generation vs. non-first-generation, and low-income vs. non-low-income). The differences in academic performance ($\Delta\%AB$, ΔGP , and ΔNRR) for all institutions were standardized to have a mean of zero and a standard deviation of one prior to clustering. The *NbClust* package in R was used to carry out the clustering (Charrad *et al.*, 2014). Each clustering algorithm was carried out while varying the cluster size (from $k = 2, \dots, 15$); we note that the number of final clusters was not predetermined. The relevant number of clusters for each clustering algorithm was found by evaluating 26 different internal indices (see Supplemental Table S2 in the Supplemental Materials for a complete list). For further discussion of the indices, see Charrad *et al.* (2014). The internal indices consist of measures of compactness (how similar are objects within the same cluster), separation (how distinct are objects from different clusters), and robustness (how reproducible are the clusters in other datasets). Index citations and whether the specific index should be maximized or minimized are included in the Supplemental Materials (Supplemental Table S2).

Statistical Tests and Data Visualizations. The course section-level dataset contained 3207 unique undergraduate course sections across five public research universities. Data visualizations included box plots providing the minimum, 25th percentile, median, 75th percentile and the maximum (Mcgill *et al.*, 1978) to examine the distributions of opportunity gaps across disciplines and institutions. We note that a value of zero on the boxplots for the opportunity gaps represent no difference in opportunity gaps for a particular demographic characteristic. We created 95% confidence intervals (CI) for the opportunity gaps ($\Delta\%AB$, ΔGP , ΔNRR ; i.e., the differences in academic performance) between each pair of subgroups (females vs. males, PEERs vs. non-PEERs, first-generation vs. non-first-generation, and low-income vs. non-low-income) for all undergraduate course sections (RQ1). We next tested whether the opportunity gaps in the four disciplines were significantly different between demographic groups (RQ1), followed by whether to opportunity gaps across the institutions were significantly different between demographic groups (RQ2) by conducting an analysis of variance (ANOVA) using an overall F-test (Fisher, 1974; Chambers *et al.*, 1992). To unpack the relationship between discipline, institution, and opportunity gaps across a variety of student demographics (RQ3), we examined the differences across the resultant clusters using the χ^2 goodness of fit test (Pearson, 1900; Fleiss *et al.*, 2013). The χ^2 goodness of fit test can be used to compare proportions between three (or more) groups of categorical data; the test compares the observed frequency distribution with the frequency distribution under the null hypothesis (where the null hypothesis is that the counts will be equally distributed across the clusters). In terms of discipline, we tested if the biology, mathematics, chemistry, and physics course sections are equally distributed across the clusters. For each institution, we tested whether the course sections were equally distributed across the clusters. To examine the relationship between course characteristics (student demographics per course section and student enrollments) and opportunity gaps across a variety of student demographics (RQ3), we quantified the differences across the resultant clusters using an overall F-test (Fisher, 1974; Chambers *et al.*, 1992). Table 3 provides the research questions, variables used to address each RQ as well as the corresponding statistical tests and where the results can be found.

RESULTS

RQ1: How do Opportunity Gaps Vary across STEM Courses among Biological Sciences Majors?

The data for all undergraduate course sections ($n = 3207$) revealed substantial opportunity gaps for students majoring in biological sciences based on a variety of demographic characteristics (Table 4; Supplemental Table S3). These inequities were pervasive and robust to adjustment for previous academic performance. The opportunity gap was largest for PEERs compared with non-PEERs (average $\Delta NRR = -0.37$, 95% CI for ΔNRR : $[-0.38, -0.36]$). PEERs typically receive between 0.36 and 0.38 grade points lower, on average, compared with non-PEERs, even after adjusting for previous academic performance. We also found inequitable outcomes when comparing females with males (95% CI for ΔNRR :

TABLE 3. Data analysis summary. The research questions, description, variables used, statistical test, and where to find the results are presented

Research question	To identify differences in:	Variables used	Statistical test	Results
RQ1: How do opportunity gaps vary across STEM courses among biological sciences majors?	Opportunity gaps across disciplines	$\Delta\%AB$, ΔGP , ΔNRR	Overall F-test	Table 5 and Supplemental Table S4
RQ2: How similar are these opportunity gaps across institutions?	Opportunity gaps across institutions	$\Delta\%AB$, ΔGP , ΔNRR	Overall F-test	Table 6 and Supplemental Table S5
	Discipline-specific opportunity gaps across institutions	$\Delta\%AB$, ΔGP , ΔNRR	Overall F-test	Supplemental Tables S6–S9
RQ3: What is the relationship between the observed opportunity gaps and discipline, institution, and course characteristics?	The distribution of disciplines across the clusters	Discipline	χ^2 goodness of fit test	Table 8
	The distribution of institutions across the clusters	Institution	χ^2 goodness of fit test	Table 8
	The distribution of student demographics across the clusters	Gender, PEER status, first-generation status, low-income status	Overall F-test	Table 8
	The distribution of enrollments across the clusters	Student enrollment	Overall F-test	Table 8

TABLE 4. 95% CIs for the opportunity gap between subgroups as measured by ΔNRR for all undergraduate course sections ($n = 3207$). The interval estimates for the differences in the NRRs (ΔNRR , Approach #3) are adjusted for previous academic performance. The 95% CIs show the range of plausible values for the opportunity gaps; since none of the intervals contain zero, all represent differences in opportunity gaps for each of the subgroups that are significant at the $\alpha = 0.05$ level

Group		Approach #3
1	2	95% CI for ΔNRR
Female	Males	(−0.19, −0.16)
PEERs	Non-PEERs	(−0.38, −0.36)
First-generation	Non-first-generation	(−0.35, −0.32)
Low-income	Non-low-income	(−0.27, −0.24)
		$n = 3207$

[−0.19, −0.16]), first-generation college students to non-first-generation college students (95% CI for ΔNRR : [−0.35, −0.32]), and low-income to non-low-income students (95% CI for ΔNRR : [−0.27, −0.24]).

We next tested whether the opportunity gaps across the four disciplines for each of the demographic comparisons were significantly different (Table 5, Supplemental Figures S1–S5; Supplemental Table S4). Figure 1 presents the results for the opportunity gaps which adjust for previous academic performance (ΔNRR , Approach #3). The opportunity gaps between subgroups of biological sciences majors show consistently large opportunity gaps across all disciplines examined (biology, chemistry, mathematics, physics). The only case in which

there was no statistically measurable difference in opportunity gaps was for first-generation students (first-generation vs. non-first generation ΔNRR ; $F = 1.43$, $p = 0.2316$).

RQ2: How Similar are These Opportunity Gaps across Institutions?

We observed differences in opportunity gaps for females, PEERs, first-generation students and low-income students across the five institutions (Figure 2; Table 6). These differences persisted regardless of analytical approach (see Supplemental Materials, Supplemental Tables S5–S9; Supplemental Figures S6–S10), including accounting for previous academic performance (Female ΔNRR : $F = 11.51$, $p < 0.001$; PEER ΔNRR : $F = 11.26$, $p < 0.001$; First-generation ΔNRR : $F = 17.58$, $p < 0.001$; Low-income ΔNRR : $F = 27.73$, $p < 0.001$). There was no consistent pattern for a single institution showing smaller opportunity gaps than the others across the demographic characteristics (Table 6). For example, institution two had smaller gaps for female students compared with male students while institution five had smaller gaps for low-income students compared with non-low-income students. Moreover, the institution with the smallest gap often differs by analytical approach, which underscores that these gaps are multifaceted.

RQ3: What is the Relationship between the Observed Opportunity Gaps and Discipline, Institution, and Course Characteristics?

Through cluster analysis, we uncovered three distinct and cohesive groupings among the course sections. Index values for the indices and the summary statistics for the best choice

TABLE 5. Discipline summary statistics for opportunity gaps as measured by ΔNRR for all undergraduate course sections ($n = 3207$). Results for ΔNRR (Approach #3) are given for each demographic characteristic examined. The mean and standard error are provided for the opportunity gaps from each discipline. The overall F-test and respective p -values for the test of the difference across disciplines. Significant results are denoted with an asterisk (*)

Approach #3 ΔNRR	Biology Mean (SE)	Chemistry Mean (SE)	Mathematics Mean (SE)	Physics Mean (SE)	F	p -value
Female	-0.15 (0.01)	-0.21 (0.01)	-0.08 (0.02)	-0.28 (0.02)	24.88	<0.001 *
PEER	-0.35 (0.01)	-0.36 (0.01)	-0.41 (0.02)	-0.48 (0.02)	14.72	<0.001 *
First-Generation	-0.33 (0.01)	-0.34 (0.01)	-0.32 (0.02)	-0.36 (0.02)	1.43	0.2316
Low-Income	-0.25 (0.01)	-0.26 (0.01)	-0.22 (0.02)	-0.30 (0.02)	3.04	0.0278 *
Number of course sections	1625	877	334	371		

of cluster size can be found in the Supplementary Materials (Supplemental Tables S10 and S11). These clusters represent fundamental divisions within our dataset and offer insights into the landscape of opportunity gaps by course section

(Figure 3; Table 7; Supplemental Tables S11–S15). The clusters can be characterized as: 1) *large opportunity gaps cluster*, 2) *mixed opportunity gaps cluster*, and 3) *small opportunity gaps cluster*. More than half (54%) of course

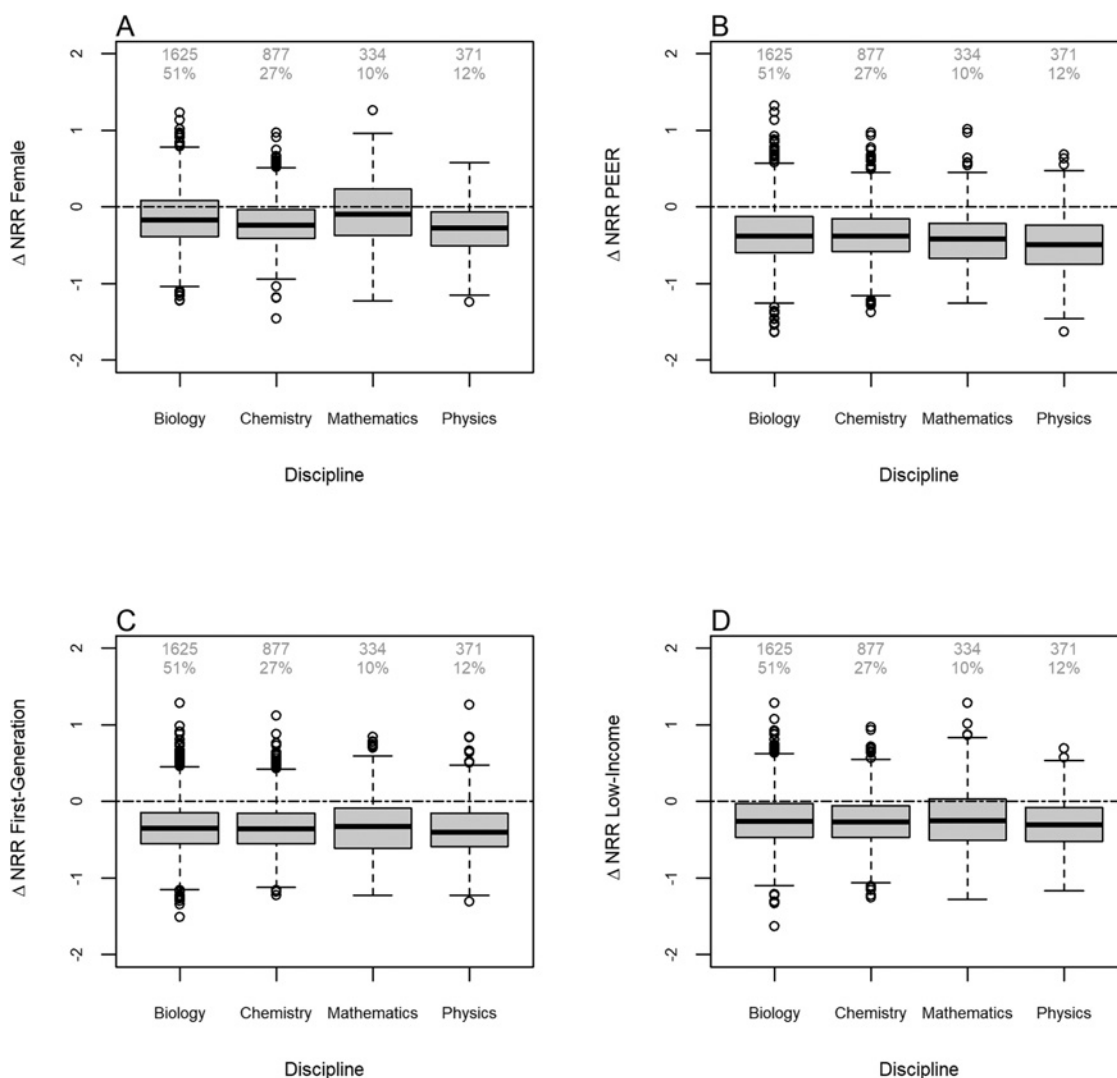


FIGURE 1. Opportunity gaps (ΔNRR) for students intending to major in the biological sciences across disciplines. Panels (A), (B), (C), and (D) present the opportunity gaps for female students, PEERs, first-generation students, and low-income students, respectively, using Approach #3.

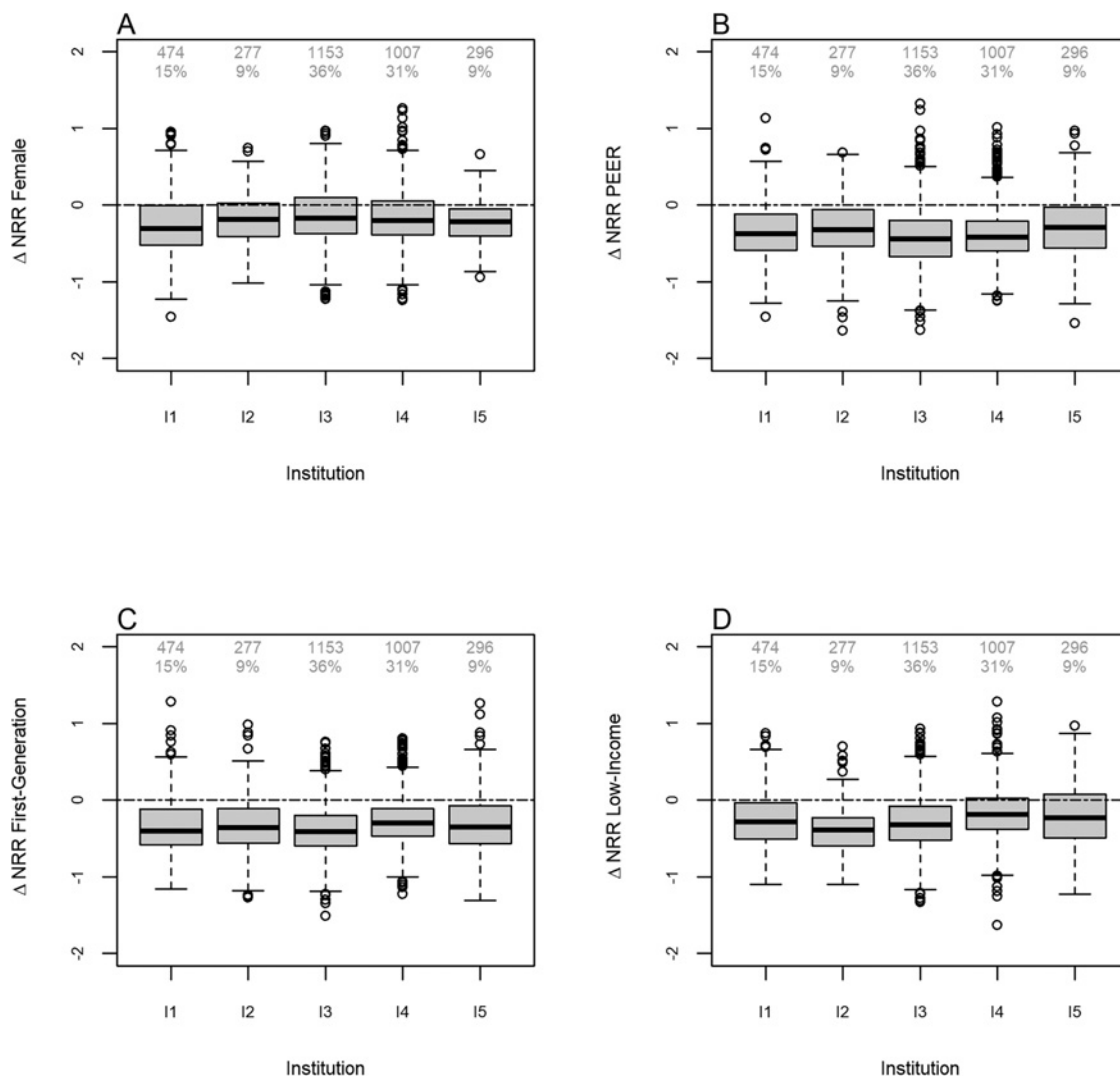


FIGURE 2. Opportunity gaps (ΔNRR) for students intending to major in the biological sciences across institutions. Panels (A), (B), (C), and (D) present the opportunity gaps for female students, PEERs, first-generation students, and low-income students, respectively, using Approach #3.

sections comprised the first cluster, representing course sections that had large opportunity gaps for all four demographic characteristics examined. The second cluster, with 11% of the course sections, had no opportunity gap for female students and small opportunity gaps for low-income students, but

large opportunity gaps for PEERs and first-generation college students similar to the first cluster. Notably, in the second cluster there was a large opportunity gap for male students. The remaining 35% of the course sections were in the third cluster, which included course sections with small opportunity

TABLE 6. Institution summary statistics for opportunity gaps as measured by ΔNRR for all undergraduate course sections ($n = 3207$). Results for ΔNRR (Approach #3) are given for each demographic characteristic examined. The mean and standard error are provided for the opportunity gaps for each institution. The overall F-test and respective p -values for the test of the difference across institutions. Significant results are denoted with an asterisk (*)

Approach #3 ΔNRR	I1	I2	I3	I4	I5	F	p -value
Female	-0.26 (0.02)	-0.19 (0.02)	-0.14 (0.01)	-0.15 (0.01)	-0.22 (0.02)	11.51	<0.001 *
PEER	-0.34 (0.02)	-0.32 (0.02)	-0.41 (0.01)	-0.38 (0.01)	-0.27 (0.02)	11.26	<0.001 *
First-Generation	-0.35 (0.02)	-0.34 (0.02)	-0.39 (0.01)	-0.27 (0.01)	-0.30 (0.03)	17.58	<0.001 *
Low-Income	-0.25 (0.02)	-0.39 (0.02)	-0.30 (0.01)	-0.18 (0.01)	-0.22 (0.02)	27.73	<0.001 *
Number of Course Sections	474	277	1153	1007	296		

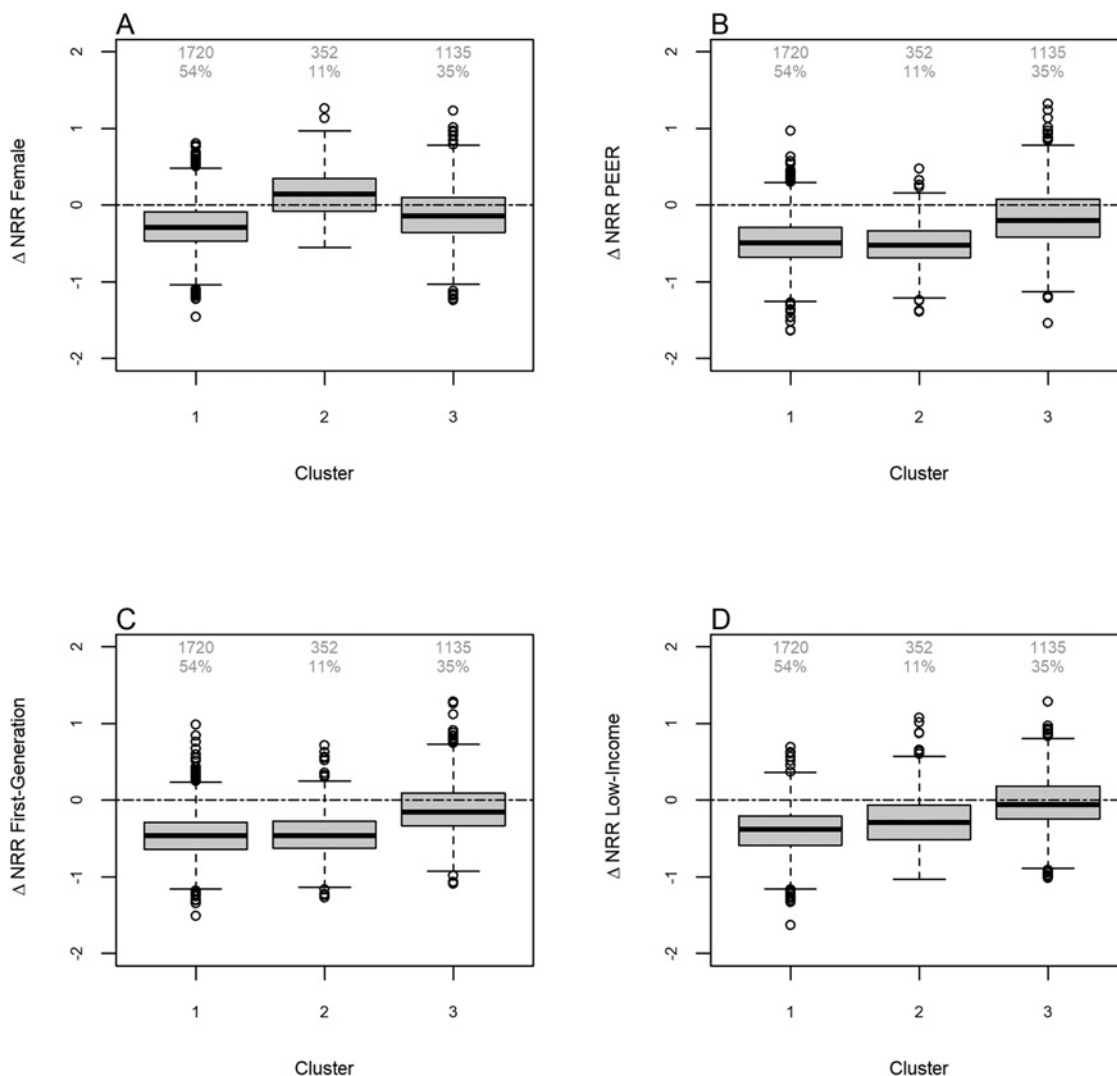


FIGURE 3. Opportunity gaps (ΔNRR) for students intending to major in the biological sciences across clusters. Panels (A), (B), (C), and (D) present the opportunity gaps for female students, PEERs, first-generation students, and low-income students, respectively, using Approach #3.

gaps for all four demographic characteristics. While the clusters were created based on the standardized opportunity gaps, we can further examine the cluster composition by overlaying institution, discipline, student demographics per course section, and student enrollments (Table 8).

Biology course sections were less represented in the *large opportunity gaps cluster* and overrepresented in the *small opportunity gaps cluster* ($\chi^2 = 34.74$, $p < 0.001$). Forty-nine percent of biology course sections were in the *large opportunity gaps cluster* (large gaps on all dimensions; compared with 54% all course sections, 59% non-biology course sections) and 40% were in the *small opportunity gaps cluster* (smaller gaps on all dimensions; compared with 35% all course sections, 31% non-biology course sections). Conversely, chemistry, mathematics, and physics were underrepresented in the *small opportunity gaps cluster* (Chemistry $\chi^2 = 16.71$, $p < 0.001$, Math $\chi^2 = 17.38$, $p < 0.001$, and Physics $\chi^2 = 14.63$, $p < 0.001$). The *mixed opportunity gaps*

cluster included proportionately more mathematics classes (17% compared with 11%).

In terms of institutions, we see that institutions 2 and 3 had higher proportions of course sections falling in the *large opportunity gaps cluster* (I2: $\chi^2 = 25.54$, $p < 0.001$; I3: $\chi^2 = 31.66$, $p < 0.001$) and lower proportions of course sections falling into the *small opportunity gaps cluster* (I2: 24%, compared with 35% for all course sections and 36% for non-I2 course sections; I3: 29%, compared with 39% for non-I3 course sections). The distribution of course sections from institution 1 was no different than expected with similar proportions of course sections falling into each of the clusters compared with the non-I1 course sections (I1: $\chi^2 = 1.68$, $p = 0.432$). Institutions 4 and 5 had lower proportions of course sections in the *large opportunity gaps cluster* and higher proportions of course sections falling in the *small opportunity gaps cluster* (I4: $\chi^2 = 27.33$, $p < 0.001$ and I5: $\chi^2 = 26.97$, $p < 0.001$).

TABLE 7. Cluster summary statistics for opportunity gaps as measured by $\Delta\%AB$, ΔGP , and ΔNRR . Results for the three analytical approaches are given for each demographic characteristic examined. The mean and standard error are provided for the opportunity gaps from each of the clusters. The overall F-test and respective p-values for the test of the difference across institutions. Significant results are denoted with an asterisk (*)

Approach	Cluster			F	p-value	
	Large opportunity gaps 1	Mixed opportunity gaps 2	Small opportunity gaps 3			
$\Delta\%AB$						
Female	-7.58 (0.29)	8.33 (0.64)	-1.63 (0.36)	282.70	<0.001	*
PEER	-19.16 (0.35)	-20.43 (0.65)	-5.34 (0.40)	383.42	<0.001	*
First-Generation	-18.04 (0.29)	-17.38 (0.73)	-3.31 (0.33)	552.10	<0.001	*
Low-Income	-16.13 (0.30)	-11.35 (0.72)	-1.43 (0.31)	520.23	<0.001	*
ΔGP						
Female	-0.19 (0.01)	0.18 (0.01)	-0.05 (0.01)	316.51	<0.001	*
PEER	-0.48 (0.01)	-0.52 (0.01)	-0.17 (0.01)	452.07	<0.001	*
First-Generation	-0.45 (0.01)	-0.43 (0.02)	-0.12 (0.01)	622.16	<0.001	*
Low-Income	-0.41 (0.01)	-0.28 (0.02)	-0.07 (0.01)	576.73	<0.001	*
ΔNRR						
Female	-0.27 (0.01)	0.15 (0.02)	-0.13 (0.01)	255.48	<0.001	*
PEER	-0.48 (0.01)	-0.53 (0.02)	-0.16 (0.01)	323.22	<0.001	*
First-Generation	-0.46 (0.01)	-0.44 (0.02)	-0.11 (0.01)	461.87	<0.001	*
Low-Income	-0.40 (0.01)	-0.26 (0.02)	-0.03 (0.01)	465.04	<0.001	*
Course sections						
Number	1720	352	1135	<i>n</i> = 3207		
Percent	54	11	35	100		

We also examined the extent to which demographic representation of a course varied across the final clusters. There was not a significant difference in the average proportion of PEERs in course sections between the clusters ($F = 0.55$, $p = 0.579$). Similarly, we did not observe different group representation for females, first-generation students, or low-income students between the clusters.

The last variable we considered in examining the clusters was the average student enrollment per course section for each of the clusters. There was a significant difference in the average student enrollment per course section for the three clusters ($F = 26.9$, $p < 0.001$). The **large opportunity gaps cluster** had the largest average student enrollments (98 students). The **mixed opportunity gaps cluster** had the smallest average student enrollments (76 students). The **small opportunity gaps cluster** had an average of 86 students per course section.

DISCUSSION

While universities often claim that they contribute to increased social mobility and access to future opportunities, we found that the majority of course sections (54%) have large opportunity gaps for historically excluded groups. The reasons could include learning environments, course structures, and instructional practices (Seymour et al., 2019), all driven in large part by pedagogical decisions made by faculty, that are racialized, gendered, and classist as well as the extended impact of systemic biases in preparation and support for college (Dizon et al., 2023). However, we also found that 35% of the course sections in our dataset exhibited no gaps across the demographic characteristics we examined, representing all four disciplines and five institutions in our dataset, showing that

equitable STEM education outcomes, insofar as they are defined herein, are achievable.

Among the disciplines included, biology courses had more equitable results overall. This is in line with prior work highlighting how outcomes tend to be more equitable in biological sciences relative to physical sciences courses (Matz et al., 2017), and may reflect a focused commitment within the discipline to address systemic disparities (Woodin et al., 2010; American Association for the Advancement of Science, 2011, 2015, 2018; Ledbetter, 2012). It could also be reflective of our study methodology, as we focus on course outcomes specifically for biological sciences students. It is possible these student populations as a whole have a greater aptitude or interest for these courses relative to those in the physical sciences, resulting in smaller opportunity gaps, although prior literature has not demonstrated a connection between course interest and equitability of outcomes. Regardless, the presence of significant gaps in a majority of course sections at all of these institutions—and the inconsistent pattern across institutions—underscores the urgent need for all universities to critically examine and reform educational practices and to look for exemplary course sections (at their own institution as well as others).

Across the institutions in this study, we found that the grades received by students who are female, PEERs, first-generation, and low-income were lower than their non-female, non-PEER, non-first-generation, and non-low-income counterparts in the majority of course sections. The opportunity gaps at the course-section level may contribute to retention issues that have been previously identified (Chen and Ho, 2012; Chen, 2013; Dika and D'Amico, 2016; Eddy and Brownell, 2016; Bettencourt et al., 2020; Hatfield et al., 2022;

TABLE 8. Cluster summary statistics for overlaid variables. Summary statistics for each of the overlaid variables (discipline, institution, student demographics per course section, and student enrollments) are presented. The number of course sections and conditional percentages are provided with the respective χ^2 tests for discipline and institution. Note that for the conditional percentages, each row sums to 100. The mean and standard error are provided student demographics per course section and the student enrollments along with the respective overall F-tests

	Cluster			χ^2	p-value	
	Large opportunity gaps 1	Mixed opportunity gaps 2	Small opportunity gaps 3			
	n (%)	n (%)	n (%)			
Discipline						
Biology	790 (49)	188 (12)	647 (40)	34.74	<0.001	*
Chemistry	521 (59)	79 (9)	277 (32)	16.71	<0.001	*
Mathematics	177 (53)	58 (17)	99 (30)	17.38	<0.001	*
Physics	232 (63)	27 (7)	112 (30)	14.63	<0.001	*
Institution						
I1	247 (52)	60 (13)	167 (35)	1.68	0.432	
I2	160 (58)	50 (18)	67 (24)	25.54	<0.001	*
I3	686 (59)	131 (11)	336 (29)	31.66	<0.001	*
I4	494 (49)	92 (9)	421 (42)	27.33	<0.001	*
I5	133 (45)	19 (6)	144 (49)	26.97	<0.001	*
	Cluster			F	p-value	
	Large opportunity gaps 1	Mixed opportunity gaps 2	Small opportunity gaps 3			
	Mean (SE)	Mean (SE)	Mean (SE)			
Demographic representation						
Female	66.87 (0.20)	66.84 (0.47)	66.88 (0.25)	0	0.9969	
PEER	26.94 (0.31)	27.76 (0.73)	27.02 (0.43)	0.55	0.5790	
First Generation	36.11 (0.36)	35.48 (0.82)	36.06 (0.48)	0.25	0.7782	
Low Income	31.10 (0.27)	31.65 (0.66)	30.44 (0.37)	1.79	0.1664	
Class size	97.81 (1.46)	76.46 (2.65)	85.86 (1.71)	26.9	<0.001	*
Course sections						
Number	1720	352	1135	n = 3207		
Percent	54	11	35	100		

Fiorini *et al.*, 2023). The universities accepted each of these students and it is reasonable to argue that, through admission, they have made an implicit promise to educate those students. Even comparing students with similar academic backgrounds, we found that students who are female, PEERs, first-generation, and low-income ended up receiving grades lower than their more privileged counterparts (see also Hatfield *et al.*, 2022). However, there were two institutions with course sections overrepresented in the cluster with small opportunity gap course sections. While all institutions in the study offer professional development activities, institutions 4 and 5 have multiple initiatives on campus related to active learning, including professional development opportunities, relevant classroom infrastructure, and tenure and promotion incentives centered around inclusive teaching. Future research to examine the different initiatives across the campus and the fidelity of implementation is needed to understand the impact of the campus on opportunity gaps.

The largest opportunity gaps identified in this research were between PEERs and non-PEERs. PEER/non-PEER

opportunity gaps were pervasive across disciplines and institutions, suggesting that all disciplines need to examine their course structures and practices to create more equitable opportunities for students from historically excluded ethnic and racial groups. Examples of examining course structures and practices could include syllabi analysis that measures tone, opportunities to make mistakes, flexibility in participation of modes, and grading policies (Eslami *et al.*, 2024). Differential outcomes in college for PEERs may be attributed to the complex interplay of systemic factors resulting in substantial differences across populations regardless of all students being well prepared for college.

Systemic biases within educational systems, including high stakes exams (Au, 2022), limited amounts of student-centered learning opportunities (Cullen *et al.*, 2012), and the prevalence of traditional lecture-based courses (Freeman *et al.*, 2014) can perpetuate inequalities. Addressing these multifaceted issues requires a comprehensive approach that tackles not only educational policies but also the university climate to create a more inclusive and equitable learning

environment for all students. One suggestion from the Joint Working Group on Improving Underrepresented Minorities persistence in STEM is to track and increase awareness of institutional progress toward diversifying STEM (Estrada *et al.*, 2016). Shukla and colleagues also provide a selection of research frameworks that move away from deficit models of student learning (2022). Their work provides examples of how achievement gaps can reinforce racial stereotypes and examples that prevent opportunity gaps, such as course-based undergraduate research experiences (Shukla *et al.*, 2022). Similarly, by identifying course sections with minimal opportunity gaps as we did in this analysis, institutions can learn what successful instructors and departments are doing to foster more equitable outcomes.

One example of effort to reduce inequities in STEM by the SEISMIC collaboration is a project that is providing “Course Equity Reports” and a yearlong curriculum for developing equity-mindedness in faculty and undergraduate students (The SELC Grant, n.d.). Teams of STEM faculty and students, supported by teaching center staff and institutional researchers, regularly review course equity analyses, discuss teaching practices and course structures, and work as a team to make campus-specific recommendations to university leadership. This model is bringing together faculty to make a change to achieve more equitable outcomes.

Other examples of beneficial practices could include interventions to increase student confidence, self-efficacy, and motivation (Koenig *et al.*, 2012; Graham *et al.*, 2013; Musu-Gillette *et al.*, 2015; Gao *et al.*, 2020), the inclusion of diverse examples and perspectives into the curriculum (Tanner, 2013; Schinske *et al.*, 2016; Lygo-Baker *et al.*, 2019), or instructors’ approaches to providing students with opportunities to receive feedback and support, such as office hours or supplemental tutoring (Topping, 1996; Yorke, 2003; Guerrero and Rod, 2013). Instructors can also work to create a more welcoming and supportive classroom culture by establishing clear expectations, promoting respect and inclusivity, and providing opportunities for students to build community and connections with one another (O’Keefe, 2013; Ahn and Davis, 2020; Van Herpen *et al.*, 2020). By examining these course-level remedies, departments and institutions can better understand how to create inclusive learning environments across a university.

Limitations

The data analyzed in this work were limited to the data found in each participating campus’ institutional research records. We rely on legal sex data, but acknowledge that this does not capture all students who may experience implicit or explicit discrimination in the classroom for reasons related to gender identity. Low-income status is defined differently across institutions, and due to data reporting policies at the individual institutions, we were unable to fully standardize low-income status leading to slightly different definitions across institutions. We would have liked to expand analyses to include more characteristics that impact the student experience, such as sense of belonging and other qualitative measures, however that data are not comprehensively collected across institutions.

To create a relatively homogenous dataset, we chose to limit this analysis to STEM courses taken by students intending to major in the biological sciences. There may be non-

STEM courses that are critical for the development of biology majors that we did not include as a result of this choice of scope. While our exclusions may limit the generalizability of the results, the methods remain fully applicable to other majors and course selections. Because we are focused on course sections, we were unable to address intersectional analysis due to the sample size of overlapping identities within each course section. And while we would have liked to consider intersectional identities (e.g., women who are PEERS, PEERS who are first generation, etc.), we would have limited ourselves more to a reduced number of course sections which had all combinations of intersectional identities.

We elected to focus on course grades as the outcome of interest, as they are the primary indicator most students receive of their own ability and achievement in a subject area. Furthermore they are essential in determining whether a student is allowed to continue in STEM, stay in college, and enter graduate school. As such, inequitable grades will inevitably lead to inequitable retention in STEM and inequitable enrollment in graduate programs. It is important to note that since our outcome was course grade, we excluded students who dropped out of the course sections because they did not earn a grade. Each institution also has different policies that guide how and when students are able to drop a course, and so rather than treat this as a single population for analysis purposes, we excluded them from the analysis. We acknowledge that better understanding who drops a STEM course and the reasons for that are important areas of research.

While our analysis focused on course grades, we did not take into consideration the types and complexity of assessments undertaken within each section that produced these grades. While some course sections may take advantage of multiple forms of assessment that represent multiple levels of complexity, including higher cognitive levels according to Bloom (Bloom, 1956; Krathwohl, 2002; Adams, 2015), others favor more fundamental recall levels often found in easily scored multiple choice assessments as seen by biology faculty often favoring lower level Bloom’s questions (Larsen *et al.*, 2022). The use of varying types of assessments likely influences the observed opportunity gaps, and thus need to be further considered.

Lastly, we acknowledge that this research was conducted at large, research intensive institutions which do not capture students’ opportunities at a diverse set of institutions. And while this framework for comparing opportunity gaps across different demographic groups could be leveraged at other types of institutions, the results of our study may not be generalizable to non-R1 institutions.

Future Work

The focus of the paper was to apply a robust quantitative approach to measure opportunity gaps across institutions and disciplines. In future work, we would like to integrate qualitative data and follow-up with course sections where we saw smaller opportunity gaps. There are a few areas which we would like to consider for future work and will discuss here briefly. The first area is to integrate qualitative data. Pearson *et al.* (2022) discuss critical approaches for quantitative STEM equity work. Scholars following the guidelines: 1) grapple with historical and present-day reality of racism, 2)

recognize how the practice of naively using statistics can uphold white supremacy, 3) interrogate how social categorizations are varied, contested, and fluid over time, 4) integrate voices of racially marginalized and minoritized individuals through qualitative mixed methods approaches, and 5) embrace research methods to pursue equity goals that align with a social justice agenda (Pearson *et al.*, 2022). Future research could include factors discussed by Ulriksen and colleagues about why students leave STEM and include measures of students identifying with the science community (2015), how students feel about the inclusivity of the learning environment, the degree of competitiveness in STEM courses, and pedagogical choices by faculty. Also, Tinto's model of student retention relies on institutional commitment, academic integration, and social integration (1975) and future work could include measures of these items. And while the goal of this study was to focus on quantitative analysis of opportunity gaps, future research including qualitative analysis could add additional context to our findings.

RECOMMENDATIONS

Recommendations for changing from a state where the majority of course sections had large opportunity gaps to a state where the majority of course sections achieve equitable outcomes are discussed below. According to Lewin's theory of change, there are three steps to a change, unfreezing, moving, and [re]freezing of group standards (Lewin, 1947). If the desired state is to achieve more equitable outcomes, we cannot evaluate equity at one point in time and expect to drive change by focusing on equity. Lewin discusses the necessity of instilling permanency in the change process (Lewin, 1947). Our recommendations are to create awareness for the current state of equitable course sections (unfreezing), provide equity analyses for the faculty (moving), identify faculty who have equitable course sections (moving), compare results with other institutions (moving). All of these efforts could support permanency of equitable student outcomes (refreezing). Below we discuss steps that institutions can take to improve equity. We hope that the recommendations below can help institutions consider how they would like to use data to inform their faculty and stakeholders, which can help them track equity in their courses

By identifying faculty who have equitable course sections, further study of how equity is achieved and what the instructors are doing to promote equity could give insights for other instructors. The institution should have structures in place to help identify these courses, including automated equity analysis through student data dashboards (Verbert *et al.*, 2013; Reinitz, 2022; Williamson and Kizilcec, 2022; Sloan-Lynch and Morse, 2024) and departmental reports that aggregate these findings, as well as appropriate research staff to help stakeholders comprehend these results. There also needs to be relevant support in the form of educational developers who are able to help instructors and programs consider how to leverage these data to generate concrete pedagogical changes. And finally, a change in incentive structure, perhaps driven through the faculty merit and promotion process, is needed; one which rewards faculty who make strides to more inclusive classroom spaces is essential for creating a university structure

that promotes equity and excellence. Finally, it is important that these discussions continue beyond a single institution. By comparing institutional outcomes across colleges and universities, a larger pool of instructors and departments with more equitable outcomes will exist, allowing for the improved identification of beneficial practices and policies.

CONCLUSIONS

By looking at the data from multiple institutions, we were able to identify patterns of differential student grades across thousands of course sections and multiple STEM disciplines. In our work, we found the majority of course sections had large opportunity gaps for females, PEERs, first-generation, and low-income biology students. There was a difference in the representation of disciplines and institutions across the clusters with biology course sections being overrepresented in the most equitable cluster. By identifying course sections that are more equitable, we are able to take the first step in creating more inclusive learning environments.

ACKNOWLEDGMENTS

This work was supported by the SEISMIC project (seismicproject.org), which is funded by the Alfred P. Sloan Foundation. SEISMIC is managed at the University of Michigan for the participating institutions, which include Arizona State University, Indiana University, Michigan State University, Purdue University, University of California Davis, University of California Irvine, University of California Santa Barbara, University of Michigan, University of Minnesota, and University of Pittsburgh. This work was also supported by grant #GT11046 from the Howard Hughes Medical Institute (www.hhmi.org). We would also like to thank Natalia Caporale (University of California, Davis) and Meaghan Pearson (University of Michigan) for their feedback which improved this paper.

REFERENCES

- Acker, J. (1990). HIERARCHIES, JOBS, BODIES: A theory of gendered organizations. *Gender & Society*, 4(2), 139–158. <https://doi.org/10.1177/089124390004002002>
- Adams, N. E. (2015). Bloom's taxonomy of cognitive learning objectives. *Journal of the Medical Library Association: JMLA*, 103(3), 152–153. <https://doi.org/10.3163/1536-5050.103.3.010>
- Ahn, M. Y., & Davis, H. H. (2020). Four domains of students' sense of belonging to university. *Studies in Higher Education*, 45(3), 622–634. <https://doi.org/10.1080/03075079.2018.1564902>
- Ainsworth-Darnell, J. W., & Downey, D. B. (1998). Assessing the oppositional culture explanation for racial/ethnic differences in school performance. *American Sociological Review*, 63(4), 536–553. <https://doi.org/10.2307/2657266>
- Allen-Ramdial, S.-A. A., & Campbell, A. G. (2014). Reimagining the pipeline: Advancing STEM diversity, persistence, and success. *BioScience*, 64(7), 612–618. <https://doi.org/10.1093/biosci/biu076>
- American Association for the Advancement of Science. (2011). Vision and change in undergraduate biology education: A call to action. Retrieved from https://www.aaas.org/sites/default/files/content_files/VC_report.pdf
- American Association for the Advancement of Science. (2015). Vision and change in undergraduate biology education: Chronicling change, inspiring the future.
- American Association for the Advancement of Science. (2018). Vision and change in undergraduate biology education: Unpacking a movement and sharing lessons learned.
- Asai, D. J. (2020). Race matters. *Cell*, 181(4), 754–757. <https://doi.org/10.1016/j.cell.2020.03.044>

- Au, W. (2022). *Unequal by Design: High-Stakes Testing and the Standardization of Inequality* (2nd ed.). New York, NY: Routledge. <https://doi.org/10.4324/9781003005179>
- Bécares, L., & Priest, N. (2015). Understanding the influence of race/ethnicity, gender, and class on inequalities in academic and non-academic outcomes among eighth-grade students: Findings from an intersectionality approach. *PLoS One*, *10*(10), e0141363. <https://doi.org/10.1371/journal.pone.0141363>
- Bell, D. A. (1995). Who's afraid of critical race theory?. *U. Ill. L. Rev.*, 893.
- Bettencourt, G. M., Manly, C. A., Kimball, E., & Wells, R. S. (2020). STEM degree completion and first-generation college students: A cumulative disadvantage approach to the outcomes gap. *The Review of Higher Education*, *43*(3), 753–779.
- Bloom, B. S. (Ed.), Engelhart, M. D., Furst, E. J., Hill, W. H., & Krathwohl, D. R. (1956). *Taxonomy of Educational Objectives: The Classification of Educational Goals*, Handbook I: Cognitive domain. London: Longmans. Retrieved on March 15, 2022 from <https://eclass.uoa.gr/modules/document/file.php/PPP242/Benjamin%20S.%20Bloom%20-%20Taxonomy%20of%20Educational%20Objectives%2C%20Handbook%201-%20Cognitive%20Domain-Addison%20Wesley%20Publishing%20Company%20%281956%29.pdf>.
- Borda, E., Schumacher, E., Hanley, D., Geary, E., Warren, S., Ipsen, C., & Stredicke, L. (2020). Initial implementation of active learning strategies in large, lecture STEM courses: Lessons learned from a multi-institutional, interdisciplinary STEM faculty development program. *International Journal of STEM Education*, *7*(1), 4. <https://doi.org/10.1186/s40594-020-0203-2>
- Bourdieu, P. (1998). *The State Nobility: Elite Schools in the Field of Power*. Stanford, CA: Stanford University Press. Retrieved on March 16, 2022, from https://d1wqtxts1x7le7.cloudfront.net/32923408/The_State_Nobility_Elite_Schools_in_the_Field_of_Power-PIERRE_BOURDIEU.pdf.
- Bourdieu, P., & Passeron, J.-C. (1990). *Reproduction in Education, Society and Culture*. Thousand Oaks, CA: SAGE.
- Canning, E. A., Muenks, K., Green, D. J., & Murphy, M. C. (2019). STEM faculty who believe ability is fixed have larger racial achievement gaps and inspire less student motivation in their classes. *Science Advances*, *5*(2), eaau4734. <https://doi.org/10.1126/sciadv.aau4734>
- Casady, R. J., & Cryer, J. D. (1976). Monotone percentile regression. *The Annals of Statistics*, *4*(3), 531–541. <https://doi.org/10.1214/aos/1176343459>
- Cataldi, E. F., Bennett, C. T., & Chen, X.; National Center for Education Statistics (2018). First-Generation Students: College Access, Persistence, and Postbachelor's Outcomes. Stats in Brief. NCES 2018-421. Retrieved March 16, 2022, from <https://eric.ed.gov/?id=ED580935>
- Castle, S. D., Byrd, W. C., Koester, B. P., Pearson, M. I., Bonem, E., Caporale, N., ... & Matz, R. L. (2024). Systemic advantage has a meaningful relationship with grade outcomes in students' early STEM courses at six research universities. *International Journal of STEM Education*, *11*(1), 14. <https://doi.org/10.1186/s40594-024-00474-7>
- Chambers, J. M., Freeny, A. E., & Heiberger, R. M. (1992). Analysis of variance; designed experiments. In: *Statistical Models in S*. New York, NY: Routledge.
- Charrad, M., Ghazzali, N., Boiteau, V., & Niknafs, A. (2014). **NbClust**: An R package for determining the relevant number of clusters in a data set. *Journal of Statistical Software*, *61*(6), 1–36. <https://doi.org/10.18637/jss.v061.i06>
- Chen, X.; National Center for Education Statistics. (2009). Students Who Study Science, Technology, Engineering, and Mathematics (STEM) in Postsecondary Education. Stats in Brief. NCES 2009-161. Retrieved March 17, 2022, from <https://eric.ed.gov/?id=ED506035>
- Chen, X.; National Center for Education Statistics. (2013). STEM Attrition: College Students' Paths into and out of STEM Fields. Statistical Analysis Report. NCES 2014-001. Retrieved March 17, 2022, from <https://eric.ed.gov/?id=ED544470>
- Chen, X., & Ho, P.; National Center for Education Statistics. (2012). STEM in Postsecondary Education: Entrance, Attrition, and Coursetaking among 2003-04 Beginning Postsecondary Students. Web Tables. NCES 2013-152. Retrieved March 17, 2022, from <https://eric.ed.gov/?id=ED566425>
- Cullen, R., Harris, M., & Hill, R. R. (2012). *The Learner-Centered Curriculum: Design and Implementation*. San Francisco, CA: John Wiley & Sons.
- Daouia, A., Gardes, L., & Girard, S. (2013). On kernel smoothing for extremal quantile regression. *Bernoulli*, *19*(5B), 2557–2589. <https://doi.org/10.13150/12-BEJ466>
- Denaro, K., Bailey, B. A., & Conrad, D. J. (2021). Quantifying disease severity of cystic fibrosis using quantile regression methods. *Journal of Data Science*, *18*(1), 148–160. [https://doi.org/10.6339/JDS.202001_18\(1\).0008](https://doi.org/10.6339/JDS.202001_18(1).0008)
- Dika, S. L., & D'Amico, M. M. (2016). Early experiences and integration in the persistence of first-generation college students in STEM and non-STEM majors. *Journal of Research in Science Teaching*, *53*(3), 368–383. <https://doi.org/10.1002/tea.21301>
- Dizon, J. P. M., Salazar, C., Kim, Y. K., & Park, J. J. (2023). Experiences of racial discrimination among STEM majors: The role of faculty. *Journal of Student Affairs Research and Practice*, *60*(5), 653–670. <https://doi.org/10.1080/19496591.2022.2144742>
- Domina, T., Penner, A., & Penner, E. (2017). Categorical inequality: Schools as sorting machines. *Annual Review of Sociology*, *43*, 311–330. <https://doi.org/10.1146/annurev-soc-060116-053354>
- Eddy, S. L., & Brownell, S. E. (2016). Beneath the numbers: A review of gender disparities in undergraduate education across science, technology, engineering, and math disciplines. *Physical Review Physics Education Research*, *12*(2), 020106. <https://doi.org/10.1103/PhysRevPhysEducRes.12.020106>
- Eddy, S. L., Brownell, S. E., & Wenderoth, M. P. (2014). Gender gaps in achievement and participation in multiple introductory biology classrooms. *CBE—Life Sciences Education*, *13*(3), 478–492. <https://doi.org/10.1187/cbe.13-10-0204>
- Eide, E., & Showalter, M. H. (1998). The effect of school quality on student performance: A quantile regression approach. *Economics Letters*, *58*(3), 345–350. [https://doi.org/10.1016/S0165-1765\(97\)00286-3](https://doi.org/10.1016/S0165-1765(97)00286-3)
- Eslami, M., Denaro, K., Collins, P., Sumarsono, J. M., Dennin, M., & Sato, B. (2024). How syllabi relate to outcomes in higher education: A study of syllabi learner-centeredness and grade inequities in STEM. *PLoS One*, *19*(4), e0301331. <https://doi.org/10.1371/journal.pone.0301331>
- Estrada, M., Burnett, M., Campbell, A. G., Campbell, P. B., Denetclaw, W. F., Gutiérrez, C. G., ... & Zavala, M. (2016). Improving underrepresented minority student persistence in STEM. *CBE—Life Sciences Education*, *15*(3), es5. <https://doi.org/10.1187/cbe.16-01-0038>
- Fiorini, S., Tarchinski, N., Pearson, M., Valdivia Medinaceli, M., Matz, R. L., Lucien, J., ... & Byrd, W. C. (2023). Major curricula as structures for disciplinary acculturation that contribute to student minoritization. *Frontiers in Education*, *8*, 1176876. <https://doi.org/10.3389/educ.2023.1176876>
- Fischer, C., Witherspoon, E., Nguyen, H., Feng, Y., Fiorini, S., Vincent-Ruz, P., ... & Schunn, C. (2023). Advanced placement course credit and undergraduate student success in gateway science courses. *Journal of Research in Science Teaching*, *60*(2), 304–329. <https://doi.org/10.1002/tea.21799>
- Fisher, R. A. (1974). *The Design of Experiments* (9th ed.). New York, NY: Hafner Press. Retrieved on March 20, 2024, from <https://home.iitk.ac.in/~shalab/anova/DOE-RAF.pdf>.
- Fisher, W. D. (1958). On grouping for maximum homogeneity. *Journal of the American Statistical Association*, *53*(284), 789–798. <https://doi.org/10.1080/01621459.1958.10501479>
- Fleiss, J. L., Levin, B., & Paik, M. C. (2013). *Statistical Methods for Rates and Proportions*. Hoboken, New Jersey: John Wiley & Sons.
- Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., & Wenderoth, M. P. (2014). Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the National Academy of Sciences*, *111*(23), 8410–8415. <https://doi.org/10.1073/pnas.1319030111>
- Freeman, A. D. (1977). Legitimizing racial discrimination through antidiscrimination law: A critical review of Supreme Court doctrine. *Minnesota Law Review*, *62*, 1049.
- Freeman, A. (1988). Racism, rights and the quest for equality of opportunity: A critical legal essay. *Harv. CR-CLL Rev.*, *23*, 295.
- Fry, R., Kennedy, B., & Funk, C.; Pew Research Center. (2021). STEM jobs see uneven progress in increasing gender, racial and ethnic diversity. Retrieved July 1, 2023, from <https://www.pewresearch.org/social-trends/2021/04/01/stem-jobs-see-uneven-progress-in-increasing-gender-racial-and-ethnic-diversity/>
- Gao, Y., Dicke, A.-L., Safavian, N., & Eccles, J. (2020). Looking into gateway: Expectancy-value profiles predict undergraduates' intent to persist in physics after introductory course. Retrieved July 22, 2023, from <https://eric.ed.gov/?id=ED620382>
- Gershenfeld, S., Ward Hood, D., & Zhan, M. (2016). The role of first-semester GPA in predicting graduation rates of underrepresented students. *Journal*

- of College Student Retention: Research, Theory & Practice, 17(4), 469–488. <https://doi.org/10.1177/1521025115579251>
- Graham, M. J., Frederick, J., Byars-Winston, A., Hunter, A.-B., & Handelsman, J. (2013). Increasing persistence of college students in STEM. *Science*, 341(6153), 1455–1456. <https://doi.org/10.1126/science.1240487>
- Griffith, A. L. (2010). Persistence of women and minorities in STEM field majors: Is it the school that matters? *Economics of Education Review*, 29(6), 911–922. <https://doi.org/10.1016/j.econedurev.2010.06.010>
- Guerrero, M., & Rod, A. B. (2013). Engaging in office hours: A study of student-faculty interaction and academic performance. *Journal of Political Science Education*, 9(4), 403–416. <https://doi.org/10.1080/15512169.2013.835554>
- Hartigan, J., & Wong, M. (1979a). A k-means clustering algorithm. *Applied Statistics*, 28(1), 100–108.
- Hartigan, J., & Wong, M. (1979b). Algorithm AS 136: A K-means clustering algorithm. *Applied Statistics*, 28(1), 100. <https://doi.org/10.2307/2346830>
- Hastie, T., Friedman, J., & Tibshirani, R. (2001). *The Elements of Statistical Learning*. New York, NY: Springer. <https://doi.org/10.1007/978-0-387-21606-5>
- Hatfield, N., Brown, N., & Topaz, C. M. (2022). Do introductory courses disproportionately drive minoritized students out of STEM pathways? *PNAS Nexus*, 1(4), pgac167. <https://doi.org/10.1093/pnasnexus/pgac167>
- He, X., & Shi, P. (1998). Monotone B-spline smoothing. *Journal of the American Statistical Association*, 93(442), 643–650. <https://doi.org/10.1080/01621459.1998.10473717>
- Kao, G., & Thompson, J. S. (2003). Racial and ethnic stratification in educational achievement and attainment. *Annual Review of Sociology*, 29(1), 417–442. <https://doi.org/10.1146/annurev.soc.29.010202.100019>
- Kaufman, L., & Rousseeuw, P. J. (2005). *Finding Groups in Data: An Introduction to Cluster Analysis*. Hoboken, New Jersey: Wiley. <https://doi.org/10.1002/9780470316801>
- Kincheloe, J. L., & McLaren, P. (2011). Rethinking critical theory and qualitative research. In K. Hayes, S. R. Steinberg, K. Tobin (Eds.), *Key Works in Critical Pedagogy. Bold Visions in Educational Research* (Vol 32, pp. 285–326). Rotterdam: Sense Publishers. https://doi.org/10.1007/978-94-6091-397-6_23
- Koenig, K., Schen, M., Edwards, M., & Bao (2012). Addressing STEM retention through a scientific thought and methods course. *Journal of College Science Teaching*, 41(4).
- Koenker, R. (2000). Galton, Edgeworth, Frisch, and prospects for quantile regression in econometrics. *Journal of Econometrics*, 95(2), 347–374. [https://doi.org/10.1016/S0304-4076\(99\)00043-3](https://doi.org/10.1016/S0304-4076(99)00043-3)
- Koenker, R. (2005). *Quantile Regression*. New York, NY: Cambridge University Press.
- Koenker, R. (2015). Quantreg: Quantile regression. Retrieved from <https://doi.org/10.32614/CRAN.package.quantreg>
- Krathwohl, D. R. (2002). A revision of Bloom's taxonomy: An overview. *Theory Into Practice*, 41(4), 212–218. https://doi.org/10.1207/s15430421tip4104_2
- Ladson-Billings, G. (1995). Toward a theory of culturally relevant pedagogy. *American Educational Research Journal*, 32(3), 465–491. <https://doi.org/10.3102/00028312032003465>
- Ladson-Billings, G. (2009). Race still matters: Critical race theory in education. *The Routledge International Handbook of Critical Education*. New York, NY: Routledge. <https://doi.org/10.4324/9780203882993>
- Ladson-Billings, G. (2020). Just what is critical race theory and what's it doing in a nice field like education? *Critical Race Theory in Education*. London: Routledge. <https://doi.org/10.4324/9781003005995>
- Ladson-Billings, G. (1998). Just what is critical race theory and what's it doing in a nice field like education? *International Journal of Qualitative Studies in Education*, 11(1), 7–24.
- Ladson-Billings, G. (2009). Race still matters: Critical race theory in education. In *The Routledge International Handbook of Critical Education* (pp. 110–122). Routledge.
- Larsen, T. M., Endo, B. H., Yee, A. T., Do, T., & Lo, S. M. (2022). Probing internal assumptions of the revised Bloom's taxonomy. *CBE—Life Sciences Education*, 21(4), ar66. <https://doi.org/10.1187/cbe.20-08-0170>
- Ledbetter, M. L. S. (2012). Vision and change in undergraduate biology education: A call to action presentation to faculty for undergraduate neuroscience, July 2011. *Journal of Undergraduate Neuroscience Education*, 11(1), A22–A26.
- Leonardo, Z. (2004). Critical social theory and transformative knowledge: The functions of criticism in quality education. *Educational Researcher*, 33(6), 11–18. <https://doi.org/10.3102/0013189X033006011>
- Lewin, K. (1947). Frontiers in group dynamics: Concept, method and reality in social science; social equilibria and social change. *Human Relations*, 1(1), 5–41. <https://doi.org/10.1177/001872674700100103>
- Lygo-Baker, S., Kinchin, I. M., & Winstone, N. E. (2019). *Engaging Student Voices in Higher Education: Diverse Perspectives and Expectations in Partnership*. Switzerland: Springer. <https://doi.org/10.1007/978-3-030-20824-0>.
- Macqueen, J. (1967). Some methods for classification and analysis of multivariate observations. *Proceedings of 5-Th Berkeley Symposium on Mathematical Statistics and Probability/University of California Press, Theory of Statistics*. (Vol 1, pp. 281–297). Berkeley and Los Angeles, CA: University of California Press. Retrieved on March 15, 2022, from <https://projecteuclid.org/ebooks/berkeley-symposium-on-mathematical-statistics-and-probability/Proceedings-of-the-Fifth-Berkeley-Symposium-on-Mathematical-Statistics-and-probability/Some-methods-for-classification-and-analysis-of-multivariate-observations/bmsmp/1200512992?tab=ChapterArticleLink>.
- Manias, E., & Street, A. (2000). Possibilities for critical social theory and Foucault's work: A toolbox approach. *Nursing Inquiry*, 7(1), 50–60. <https://doi.org/10.1046/j.1440-1800.2000.00048.x>
- Matz, R. L., Koester, B. P., Fiorini, S., Grom, G., Shepard, L., Stangor, C. G., ... & McKay, T. A. (2017). Patterns of gendered performance differences in large introductory courses at five research universities. *AERA Open*, 3(4), 233285841774375. <https://doi.org/10.1177/2332858417743754>
- McCormick, A. (2001). The 2000 Carnegie Classification: Background and description. *The Carnegie Classification of Institutions of Higher ...*. Retrieved March 15, 2022 from, https://www.academia.edu/2947004/The_2000_Carnegie_Classification_Background_and_description
- McCoy, D. L., & Rodricks, D. J. (2015). *Critical Race Theory in Higher Education: 20 Years of Theoretical and Research Innovations: ASHE Higher Education Report, Volume 41, Number 3*. Hoboken, New Jersey: John Wiley & Sons. <https://doi.org/10.1002/aehe.20021>
- McGee, E. O. (2020). Interrogating structural racism in STEM higher education. *Educational Researcher*, 49(9), 633–644. <https://doi.org/10.3102/0013189X20972718>
- McGill, R., Tukey, J. W., & Larsen, W. A. (1978). Variations of box plots. *The American Statistician*, 32(1), 12–16. <https://doi.org/10.1080/00031305.1978.10479236>
- Mead, C., Supriya, K., Zheng, Y., Anbar, A. D., Collins, J. P., LePore, P., & Brownell, S. E. (2020). Online biology degree program broadens access for women, first-generation to college, and low-income students, but grade disparities remain. *PLoS One*, 15(12), e0243916. <https://doi.org/10.1371/journal.pone.0243916>
- Møller, J. K., Nielsen, H. A., & Madsen, H. (2008). Time-adaptive quantile regression. *Computational Statistics & Data Analysis*, 52(3), 1292–1303. <https://doi.org/10.1016/j.csda.2007.06.027>
- Musu-Gillette, L. E., Wigfield, A., Harring, J. R., & Eccles, J. S. (2015). Trajectories of change in students' self-concepts of ability and values in math and college major choice. *Educational Research and Evaluation*, 21(4), 343–370. <https://doi.org/10.1080/13803611.2015.1057161>
- National Science Foundation. (2021). Women, minorities, and persons with disabilities in science and engineering. Retrieved January 5, 2024, from <https://nces.nsf.gov/pubs/nsf21321/>
- Ng, R. T., & Han, J. (2002). CLARANS: A method for clustering objects for spatial data mining. *IEEE Transactions on Knowledge and Data Engineering*, 14(5), 1003–1016. <https://doi.org/10.1109/TKDE.2002.1033770>
- O'Keefe, P. (2013). A sense of belonging: Improving student retention. *College Student Journal*, 47(4), 605–613.
- Patton, L. D. (2016). Disrupting postsecondary prose: Toward a critical race theory of higher education. *Urban Education*, 51(3), 315–342. <https://doi.org/10.1177/0042085915602542>
- Pearson, K. (1900). X. On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 50(302), 157–175. <https://doi.org/10.1080/14786440009463897>
- Pearson, M. I., Castle, S. D., Matz, R. L., Koester, B. P., & Byrd, W. C. (2022). Integrating critical approaches into quantitative STEM equity work. *CBE—Life Sciences Education*, 21(1), es1. <https://doi.org/10.1187/cbe.21-06-0158>

- Pollard, D. (1981). Strong consistency of k-means clustering. *The Annals of Statistics*, 9(1), 135–140.
- Portnoy, S., & Koenker, R. (1997). The Gaussian hare and the Laplacian tortoise: Computability of squared-error versus absolute-error estimators. *Statistical Science*, 12(4), 279–300. <https://doi.org/10.1214/ss/1030037960>
- Ray, V. (2019). A theory of racialized organizations. *American Sociological Review*, 84(1), 26–53. <https://doi.org/10.1177/0003122418822335>
- Reinitz, B. T. (2022). 2022 EDUCAUSE Horizon Report Data and Analytics Edition (1–54). EDU22. Retrieved January 17, 2024, from <https://www.learntechlib.org/p/221467/>
- Rousseeuw, P., & Kaufman, L. (1987). Clustering by means of medoids. *Proceedings of the Statistical Data Analysis Based on the L1 Norm Conference*, 31.
- Salehi, S., Burkholder, E., Lepage, G. P., Pollock, S., & Wieman, C. (2019). Demographic gaps or preparation gaps?: The large impact of incoming preparation on performance of students in introductory physics. *Physical Review Physics Education Research*, 15(2), 020114. <https://doi.org/10.1103/PhysRevPhysEducRes.15.020114>
- 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. <https://doi.org/10.1187/cbe.16-01-0002>
- Sedgwick, P. (2014). Retrospective cohort studies: Advantages and disadvantages. *BMJ*, 348, g1072. <https://doi.org/10.1136/bmj.g1072>
- SEISMIC Overview. (n.d.). SEISMIC Collaboration. Retrieved March 14, 2022, from <https://www.seismicproject.org/about/overview/>
- SELC grant, SEISMIC Collaboration. (n.d.). SEISMIC Collaboration. https://www.seismicproject.org/seismic-central/the_selc_grant/
- Seymour, E., Hunter, A.-B., & Harper, R. (2019). *Talking About Leaving Revisited: Persistence, Relocation, and Loss in Undergraduate STEM Education*. <https://doi.org/10.1007/978-3-030-25304-2>
- Shukla, S. Y., Theobald, E. J., Abraham, J. K., & Price, R. M. (2022). Reframing educational outcomes: Moving beyond achievement gaps. *CBE—Life Sciences Education*, 21(2), es2. <https://doi.org/10.1187/cbe.21-05-0130>
- Sloan-Lynch, J., & Morse, R. (2024). Equity-forward learning analytics: Designing a dashboard to support marginalized student success. *Proceedings of the 14th Learning Analytics and Knowledge Conference*, 1–11. <https://doi.org/10.1145/3636555.3636844>
- Solórzano, D. G., & Villalpando, O. (1998). *Critical Race Theory, Marginality, and The Experience of Students of Color in Higher Education*, 21. SUNY Press.
- Stains, M., Harshman, J., Barker, M. K., Chasteen, S. V., Cole, R., DeChenne-Peters, S. E., ... & Young, A. M. (2018). Anatomy of STEM teaching in North American universities. *Science*, 359(6383), 1468–1470. <https://doi.org/10.1126/science.aap8892>
- Tanner, K. D. (2013). Structure Matters: Twenty-one teaching strategies to promote student engagement and cultivate classroom equity. *CBE—Life Sciences Education*, 12(3), 322–331. <https://doi.org/10.1187/cbe.13-06-0115>
- Tatum, B. D. (2017). “Why are all the Black kids still sitting together in the cafeteria”: And other conversations about race in the twenty-first century. *Liberal Education*, 103(3–4), 46–56.
- Taylor, E., Gillborn, D., & Ladson-Billings, G. (2023). *Foundations of Critical Race Theory in Education*. New York, NY: Taylor & Francis, Routledge..
- Taylor, E., Gillborn, D., & Ladson-Billings, G. (2009). *Foundations of critical race theory in education*.
- 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. <https://doi.org/10.1073/pnas.1916903117>
- Tinto, V. (1975). Dropout from higher education: A theoretical synthesis of recent research. *Review of Educational Research*, 45(1), 89–125. <https://doi.org/10.3102/00346543045001089>
- Topping, K. J. (1996). The effectiveness of peer tutoring in further and higher education: A typology and review of the literature. *Higher Education*, 32(3), 321–345. <https://doi.org/10.1007/BF00138870>
- Ulriksen, L., Madsen, L. M., & Holmegaard, H. T. (2015). Why do students in STEM higher education programmes drop/opt out? – Explanations offered from research. In E. K. Henriksen, J. Dillon, and J. Ryder (Eds.), *Understanding Student Participation and Choice in Science and Technology Education* (pp. 203–217). New York, NY: Springer. <https://doi.org/10.1007/978-94-007-7793-4>
- Van Herpen, S. G. A., Meeuwisse, M., Hofman, W. H. A., & Severiens, S. E. (2020). A head start in higher education: The effect of a transition intervention on interaction, sense of belonging, and academic performance. *Studies in Higher Education*, 45(4), 862–877. <https://doi.org/10.1080/03075079.2019.1572088>
- Verbert, K., Duval, E., Klerkx, J., Govaerts, S., & Santos, J. L. (2013). Learning analytics dashboard applications. *American Behavioral Scientist*, 57(10), 1500–1509. <https://doi.org/10.1177/0002764213479363>
- Williamson, K., & Kizilcec, R. (2022). A review of learning analytics dashboard research in higher education: Implications for justice, equity, diversity, and inclusion. *LAK22: 12th International Learning Analytics and Knowledge Conference*, 260–270. <https://doi.org/10.1145/3506860.3506900>
- Woodin, T., Carter, V. C., & Fletcher, L. (2010). Vision and change in biology undergraduate education, a call for action—initial responses. *CBE—Life Sciences Education*, 9(2), 71–73. <https://doi.org/10.1187/cbe.10-03-0044>
- Xiong, W., & Tian, M. (2021). Weighted quantile regression theory and its application. *Journal of Data Science*, 17(1), 145–160. [https://doi.org/10.6339/JDS.201901_17\(1\).0007](https://doi.org/10.6339/JDS.201901_17(1).0007)
- Yorke, M. (2003). Formative assessment in higher education: Moves towards theory and the enhancement of pedagogic practice. *Higher Education*, 45(4), 477–501. <https://doi.org/10.1023/A:1023967026413>
- Zhang, L., Wang, H. J., & Zhu, Z. (2017). Composite change point estimation for bent line quantile regression. *Annals of the Institute of Statistical Mathematics*, 69(1), 145–168. <https://doi.org/10.1007/s10463-015-0538-5>
- Zhou, K. Q., & Portnoy, S. L. (1998). Statistical inference on heteroscedastic models based on regression quantiles. *Journal of Nonparametric Statistics*, 9(3), 239–260. <https://doi.org/10.1080/10485259808832745>