

# A QuantCrit Investigation of Society's Educational Debts Due to Racism and Sexism in Chemistry Student Learning

Ben Van Dusen,\* Jayson Nissen, Robert M. Talbot, Hannah Huvad, and Mollee Shultz



Cite This: *J. Chem. Educ.* 2022, 99, 25–34



Read Online

ACCESS |



Metrics & More



Article Recommendations

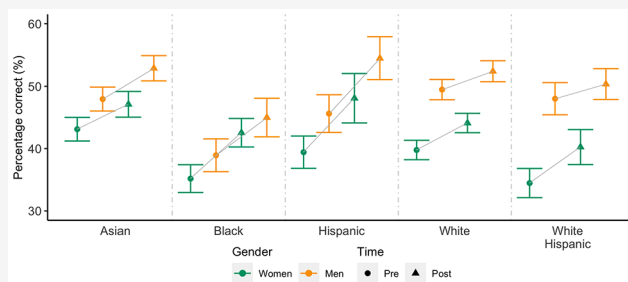


Supporting Information

**ABSTRACT:** The American Chemical Society holds supporting diverse student populations engaging in chemistry as a core value. We analyzed chemical concept inventory scores from 4,612 students across 12 institutions to determine what inequities in content knowledge existed before and after introductory college chemistry courses. We interpreted our findings from a Quantitative Critical (QuantCrit) perspective that framed inequities as educational debts that society owed students due to racism, sexism, or both. Results showed that society owed women and Black men large educational debts before and after instruction. Society's educational debts before instruction were large enough that women and Black men's average scores were lower than White men's average pretest scores even after instruction. Society would have to provide opportunities equivalent to taking the course up to two and a half times to repay the largest educational debts. These findings show the scale of the inequities in the science education systems and highlight the need for reallocating resources and opportunities throughout the K–16 education system to mitigate, prevent, and repay society's educational debts from sexism and racism.

**KEYWORDS:** General Public, Testing/Assessment, Women in Chemistry, Minorities in Chemistry, Chemical Education Research

**FEATURE:** Chemical Education Research



## INTRODUCTION

The American Chemical Society (ACS) holds diversity, equity, inclusion, and respect as core values. Few studies, however, have examined the impacts of systemic sexism and racism on chemistry student outcomes<sup>1</sup> despite the ACS and many other leading bodies in the STEM disciplines calling for an increased effort to improve equity for women and Black, Indigenous, and people of color (BIPOC).<sup>2–6</sup> The lack of fundamental research on inequities in chemistry student outcomes makes it difficult for instructors and researchers to contextualize inequities in their student outcome data.

This study examines the state of inequities in introductory chemistry students' learning from a Quantitative Critical (QuantCrit)<sup>7</sup> perspective. We analyze student pretest and posttest data on the Chemical Concept Inventory (CCI).<sup>8</sup> Mulford and Robinson developed the CCI as a research-based assessment<sup>9</sup> of introductory college student chemistry knowledge. The data came from the Learning About STEM Student Outcomes (LASSO)<sup>10</sup> platform and included data from 4,612 students in 39 introductory chemistry courses at 12 institutions across the United States. This work builds on the existing literature on inequities in chemistry education by measuring the impacts of sexism and racism on chemistry student learning across multiple institutions. We use Ladson-Billings<sup>11,12</sup> concept of "educational debts" to interpret the extent to

which college chemistry courses mitigate, perpetuate, or exacerbate the educational debts society owes to students from groups minoritized<sup>13</sup> by racism and sexism. The results can support chemistry programs in examining the extent to which they repay or add to society's educational debts. To support readers interpretation of our research, we include a selection of terms and definitions for potentially novel statistical modeling and equity-related terms that we use (Table 1).

## RESEARCH QUESTIONS

To better understand the role that sexism and racism play in shaping chemistry student learning, we asked the following questions:

1. To what extent have sexism and racism created educational debts of chemistry knowledge that society

**Special Issue:** Diversity, Equity, Inclusion, and Respect in Chemistry Education Research and Practice

**Received:** April 5, 2021

**Revised:** June 21, 2021

**Published:** July 22, 2021



**Table 1. Definition of Some Statistical Modeling and Equity-Related Terms We Use in the Manuscript**

Term	Definition
Hierarchical linear model	A linear regression model that accounts for the nested nature of data (e.g., students within courses).
Bayesian	A statistical system based on probability, rather than frequency.
Multiple imputation	A principled approach to handling missing data that generates multiple complete data sets based on collected data and combines the results to account for the increased variance.
Uncertainty (standard error)	In Bayesian statistics, the uncertainty provides a probability that a true value falls within a range. For example, Bayesian confidence intervals of 1 standard error have a 61% probability of including the true value. Confidence intervals in frequentist statistics do not mean this but are often misinterpreted as such.
Equality	When things are equal between groups of people. This can be applied to different measures, such as having equal resources, opportunities, or outcomes. In this work, we focus on equal outcomes.
Equity	When a course allocates resources and opportunities according to each person's circumstances to reach equal outcomes.
Educational debt	"Education debt is the foregone schooling resources that we could have (should have) been investing in (primarily) low-income kids, which deficit leads to a variety of social problems (e.g., crime, low productivity, low wages, low labor force participation) that require on-going public investment". <sup>a</sup> This reframes inequities in group outcomes from deficits in student abilities to debts that society owes minoritized students due to racism and sexism. <sup>a,b</sup>
Intersectionality	"Intersectionality means the examination of race, sex, class, national origin, and sexual orientation and how their combinations play out in various settings." <sup>c</sup>
Critical race theory	"[Critical race theory] begins with the notion that racism is 'normal, not aberrant, in American society' (Delgado, 1995, p xiv), and, because it is so enmeshed in the fabric of our social order, it appears both normal and natural to people in this culture." <sup>d</sup>
QuantCrit	A theoretical framework that applies critical race theory to quantitative research.

<sup>a</sup>See ref 11. <sup>b</sup>See ref 12. <sup>c</sup>See ref 14. <sup>d</sup>See ref 15.

**Table 2. Comparative Share of Bachelor and Ph.D. Degrees Earned by Each Social Identifier across Chemistry, Physics, and Biology<sup>a</sup>**

Social Identifier: Sex or Race	U.S. Totals, %	Distribution of People in the United States with These Degrees by Social Identifier, %					
		Earning Chemistry Degrees		Earning Physics Degrees		Earning Biology Degrees	
		Bachelor	Ph.D.	Bachelor	Ph.D.	Bachelor	Ph.D.
Female	50.8	48.5	37.8	19.3	19.3	60.5	52.6
Asian	4.7	14.1	8.9	7.1	7.0	14.8	11.0
Black	12.3	7.0	4.0	3.3	2.3	7.6	4.0
Hispanic	16.3	10.5	6.0	9.0	4.0	12.6	7.9
White	63.9	61.0	72.4	72.2	75.7	57.7	67.6

<sup>a</sup>See ref 17.

owes students before taking introductory college chemistry courses?

- To what extent do introductory college chemistry courses mitigate, perpetuate, or exacerbate the educational debts that society owes students?

## LITERATURE REVIEW

### Equity Research in College Chemistry Education

The representation of women and BIPOC students in chemistry is higher than in other STEM disciplines (see Table 2) but remains problematic in higher education and chemistry professions. White et al.<sup>16</sup> explain that this underrepresentation trickles into the workforce, leading to the percentage of chemistry professionals identifying as Black (6.2%) or Hispanic (7%) being much smaller than in the population of the United States (13.4% and 18.5%). Earning a chemistry degree is a prerequisite to pursuing a chemistry career; therefore, equity in the chemistry workforce starts with equity in college chemistry education. Few studies, however, have investigated equity in undergraduate chemistry education.<sup>1</sup>

In our review of the literature, we found 18 articles that quantitatively measured race order inequities in undergraduate chemistry classrooms.<sup>16,18,27–34,19–26</sup> Studies that examined race classified students as either majority (White and Asian students) or underrepresented minority students (URM;

Black, Hispanic/Latinx, and American Native students). One study<sup>24</sup> looked at the representation of various races within commonly used college-level chemistry textbooks and found that URM students were severely underrepresented. The remaining studies that focused only on race examined student achievement in various undergraduate chemistry courses. Bancroft et al.<sup>18</sup> demonstrated that Black, Latinx, and Native American students depart from STEM majors, including chemistry, at higher rates than their White and Asian peers. Shortlidge<sup>35</sup> illustrated that common active learning techniques do not necessarily eliminate achievement differences between URM students and White students. Thus, although traditional lecture pedagogies disadvantage URM students in chemistry,<sup>16</sup> implementing a different teaching technique may not rid chemistry courses of historical and systemic racism.<sup>1,35</sup>

Studies of gender differences focused on both academic achievement and social-psychological characteristics of chemistry students.<sup>21,27,32,34</sup> Dalgety and Coll<sup>21</sup> and Sunny et al.<sup>32</sup> showed that men had higher self-efficacy for chemistry concepts and chemistry tasks than women. Sunny et al.<sup>32</sup> also showed that men had lower test anxiety levels than women in chemistry, but men and women performed the same on an unspecified chemistry achievement test. Richards-Babb<sup>27</sup> considered a case of women outperforming men in chemistry and examined how replacing in-class quizzes with online homework assignments decreased this difference. The author

hypothesized the online homework improved men's self-reported, poor study habits.

Five studies examined both race and gender discrepancies in undergraduate chemistry.<sup>20,23,28,29,33</sup> Rocabado<sup>28</sup> explained how Black women had lower achievement and emotional satisfaction in chemistry than other students. Lewis and Lewis<sup>25</sup> and Stanich<sup>31</sup> considered several social identifier variables (e.g., race and gender) for "at-risk" chemistry students. They found that adverse experiences in general chemistry courses may be discouraging to URM and other at-risk students.

### The Chemical Concept Inventory

The studies in this review primarily focused on differing academic achievement as influenced by pedagogical techniques, various interventions, or standard chemistry courses, with little attention to instruments used to measure such achievement. Concept inventories are instruments designed to measure students' conceptual understanding and, although not intended to do so, have been used as a proxy for student achievement in the sciences.<sup>36</sup> Chemistry specific concept inventories like the Chemistry Concept Inventory<sup>8</sup> (CCI), the Chemistry Self-Concept Inventory,<sup>37</sup> and the Thermochemistry Concept Inventory<sup>38</sup> have been developed and used to measure chemistry students' learning and achievement.

Mulford and Robinson<sup>8</sup> designed the CCI to use 22 multiple-choice items to measure student alternative conceptions about topics covered in first-semester university chemistry courses. Instructors administer the CCI as both a pretest and posttest to measure shifts in student alternative conceptions after taking a semester of university chemistry. Several additional investigations using both qualitative and quantitative methods support the initial validation argument for the instrument.<sup>39,40</sup> Barbera<sup>40</sup> examined CCI data from over 2,500 students at four institutions using classical test theory and Rasch modeling. Barbera concluded that the assessment and its items were suitable for the large-scale evaluation of students' alternate conceptions. Schwartz and Barbera<sup>41</sup> performed think-aloud interviews with 25 students as they completed the CCI. Their analysis found a mix of support and threats to the response process validity of several questions as some students answered either correctly or incorrectly for the wrong reasons. They concluded that these threats warrant future research examining ways to modify items on the CCI. The articles that referenced Schwartz and Barbera<sup>41</sup> did not follow up on this recommendation, but Schultz et al.<sup>42</sup> considered the CCI to be "...validated through extensive psychometric testing".

## ■ CONCEPTUAL FRAMEWORK

### Quantitative Critical Race Theory (QuantCrit)

Critical race theory (CRT) began in the 1970s<sup>43–45</sup> to address social injustices and racial oppression. CRT emphasizes examining oppressive power structures, challenging the ideas of objectivity, and considering the intersectionality of individual's identities.<sup>45,46</sup> Subsequently, scholars in many fields, including education,<sup>15,47</sup> have used CRT to guide their work.

Critical research has historically used qualitative approaches to investigate people's lived experiences from marginalized groups and the social processes that create racist, sexist, and classist power structures.<sup>48–51</sup> QuantCrit emerged as a quantitative perspective<sup>7</sup> aligned with the core principles of

critical research. Below, we describe four principles of QuantCrit<sup>48</sup> and the ways we strove to embody them in this investigation.

**Principle 1: The Centrality of Oppression.** We assumed that racism and sexism are complex and dynamic processes present throughout society that we must explicitly examine lest our statistical models legitimize existing inequities. Educational inequities come from oppressive power structures that create educational and societal systems that cater to students from dominant groups. As such, we follow Ladson-Billings<sup>11,12</sup> framing of inequities in group performance as educational debts that society owes students due to their continual marginalization.

**Principle 2: Categories Are Neither "Natural" nor Given.** All data are socially constructed and reflect the hegemonic power structures that created them. Our models aggregated students by social identifiers for race and gender. These categories do not represent any natural or scientific truth about students but are social constructs that maintain hegemonic power structures. The dynamic socially negotiated natures of race and gender do not diminish the genuine effects of racism and sexism associated with them. We reflected this in our writing by naming racism and sexism as the causes of educational debts identified by the models.

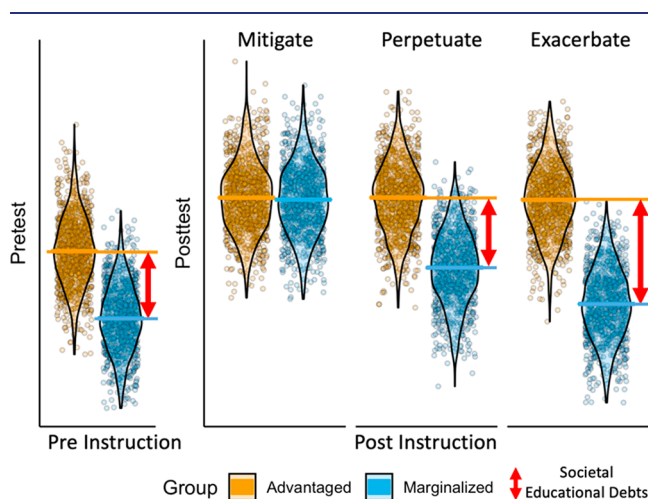
**Principle 3: Data Is Not Neutral and Cannot "Speak for Itself".** Racist and sexist assumptions can shape every stage of collecting, analyzing, and interpreting data.<sup>27</sup> In our analysis, we drew on research and thinking from across fields to use methods that represented the impacts of racism and sexism, knowing that the data and methods were imperfect. For example, we do not use *p*-values in our work. They depend on sample sizes and can lead researchers to dismiss meaningful inequities due to lack of representation in minoritized groups.<sup>28</sup> Instead, we used the standard errors of the estimated scores to inform our confidence in the results.

**Principle 4: The Importance of Intersectionality.** Identity is multifaceted (e.g., race, gender); each aspect dynamically intersects with each other and society's associated oppressive power structures to shape their experience of the world.<sup>46</sup> For example, Black women experience racism differently from Black men and sexism differently from White<sup>52</sup> women. In this analysis, we accounted for the dynamic interactions between sexism and racism by examining the impact of interaction terms for our social identifier variables.

### Operationalizing Equity

Because data cannot speak for itself, we follow the advice of Rodriguez et al.<sup>53</sup> and Stage<sup>7</sup> and operationalized equity to interpret our findings from an antiracist perspective. Kendi<sup>54</sup> defines antiracism as the ideas, beliefs, and policies that hold racial groups as equal. From this antiracist perspective, society's educational debts result from racial discrimination. We operationalized equity as *equality of outcomes* to align with this definition. Equality of outcomes occurs when students from different gender, race, and ethnic groups have the same average achievement at the end of a course and are not owed educational debts on that metric.<sup>53,55–57</sup> This perspective takes a strong social-justice stance. It argues that a just education system must allocate more resources and opportunities to the students owed educational debts to begin repaying them and eliminating inequities.

Figure 1 illustrates our conception of society's educational debts. Students enter a course with different skill or knowledge



**Figure 1.** A visual representation of society's educational debts before and after instruction with three potential outcomes. The figure uses simulated data with the horizontal lines representing mean scores, dots representing individual students, and the violin plot envelopes representing the density of scores. The figure shows how educational debts can be mitigated, perpetuated, or exacerbated. It also shows that, in statistical models, educational debts are measures of average differences between groups, not absolute differences between individuals.

distributions due to oppression, such as systemic racism. Courses can then either mitigate, perpetuate, or exacerbate those educational debts. Figure 1 illustrates the most extreme forms of mitigation (the achievement of equality of outcomes and the full repayment of the educational debt) and exacerbation (the complete denial of education to the minoritized group). Society's educational debts are multifaceted and deeply rooted. The full repayment of one educational debt does not imply the repayment of all educational debts to a minoritized group.

### Positionality

The members of our team each brought unique lived experiences and perspectives to this research. Our identities span genders, races, and areas of disciplinary expertise. To contextualize the research's perspectives, we provide positionality statements for each author in the Supporting Information.

## METHODS

### Data Collection, Cleaning, and Imputation

We accessed student and course data through the Learning About STEM Student Outcomes (LASSO) platform.<sup>10</sup> The LASSO platform automates the administration, scoring, and analysis of research-based assessments online. The LASSO platform builds a database of anonymized student and course data from students who consented to share their data. Most educators administered the CCI as a pretest during the first week of class and as a posttest during the last week of class. The analyzed data came from 4,612 students in 39 first-semester introductory college chemistry courses. The courses were taught at 12 institutions, including one two-year college, two private universities, and three minority-serving institutions. Of the courses, 32 used the Learning Assistant Model<sup>58,59</sup>

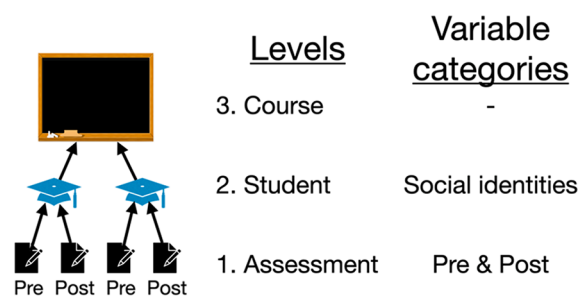
(L.A. model) to implement collaborative instruction. All 39 courses reported using collaborative learning; 16 of the courses completed a survey on their teaching practices implemented in the Spring of 2019, and 15 of these courses used the L.A. model. In response to a question regarding if students worked together in small groups, instructors of two courses responded multiple times per class, 11 responded nearly every class, and 3 responded weekly. In response to a question regarding if instructors used lecture, 1 responded multiple times per class, 13 responded nearly every class, 1 responded weekly, and 1 responded never. In response to a question regarding if students solved problems during class, 6 responded nearly every class and 10 responded weekly.

To clean the data, we removed the pretest or posttest score if the student took less than 5 min on the assessment. We removed any courses with less than either nine pretests or nine posttests. After cleaning the data, we used hierarchical multiple imputation (HMI) with the `hmi`<sup>60</sup> and `mice`<sup>61</sup> packages in RStudio V. 1.2.5042 to impute missing data. We only imputed values for missing pretest and posttest CCI scores. HMI provided a principled method for handling missing data that maximized statistical power and minimized bias while accounting for the hierarchical structure of the data.<sup>62–65</sup>

The rate of missing data for this data set was 17% on the pretest and 34% on the posttest, which fall within the range of missing data on concept inventories.<sup>65</sup> The imputation model included a dependent variable for the posttest and accounted for the pretest score and social identifier variables and nested the students within courses. We included the disaggregated descriptive statistics and violin plots of student scores in the Supporting Information.

### Model Building

We developed models of student's chemistry conceptual knowledge on the pretest and posttest, described by  $score_{ijk}$  in the final model. The models were 3-level hierarchical linear models (Figure 2) with assessment data (CCI scores) in the



**Figure 2.** Three-level structure of the data and model with the variable categories included in each level.

first level ( $i$ ), student data in the second level ( $j$ ), and course data in the third level ( $k$ ). Using hierarchical linear models accounted for the nested nature of the data:<sup>41,42</sup> 7% of the variance occurred at the third level (course), 23% at the second level (student), and 70% at the first level (assessment), indicating that a 3-level model was appropriate for modeling the data.<sup>66</sup> We used Bayesian analysis to run the models and pool the results for the imputed data sets using the `rstan`<sup>67</sup> and `brms`<sup>68</sup> packages in R. The model parameters were fit using the penalized least-squares method, four chains, 1,000 iteration burn-in, and 2,000 total iterations.

The data set included social identity data for gender, race, and ethnicity. The social identity questionnaire for the LASSO platform has changed over the data collection. The current questionnaire is included in the [Supporting Information](#) and included 7 gender options, 6 ethnicity options, and 17 race options. Each included options for writing in another identity or preferring not to answer. Students were able to select multiple responses for the gender, ethnicity, and race questions.

Previous social identity surveys on LASSO have conflated gender and sex, which may have shaped how some students responded to these questions. We grouped students who identified as male or man and those who identified as female or woman. We use the terms men and women.

To determine what social identity variables to include in the models, we first used a general principle to only investigate scores for populations with at least 20 students total.<sup>69</sup> Following this principle meant that we did not include variables for transgender, Hawaiian or Pacific Islander, or Native American or Alaskan Native in our models. Because removing the students with these identities could have biased the course-level results and because some students did not include a gender or race, we combined these students into two categories: gender other and race other. Our model's final variables (shown below) included  $woman_{jk}$ ,  $gender\_other_{jk}$ ,  $Black_{jk}$ ,  $Asian_{jk}$ ,  $Hispanic_{jk}$ ,  $White_{jk}$ , and  $race\_other_{jk}$ . We included interactions between variables whenever a population had more than 20 students but not for the race other and gender other groups. The model included interaction terms between  $Hispanic_{jk}$  and  $White_{jk}$  and between  $woman_{jk}$  and each of the model's racial groups.

The models also included several variables not related to social identity. We included the variable  $posttest_{ijk}$  to identify shifts from pretest to posttest and the variable  $retake_{jk}$  to identify whether a student has previously taken the course.  $posttest_{ijk}$  was interacted with all the social identity terms other than  $gender\_other_{jk}$  and  $race\_other_{jk}$ . These terms were not interacted with  $posttest_{ijk}$  as we are not looking to support any claims about them.

To ensure model parsimony, we identified our final model as the model with the lowest Akaike information criterion corrected<sup>70,71</sup> (AICc) calculated by the dredge function in the MuMin package<sup>72</sup> in R.

final model: level-1 equations (assessment-level)

$$score_{ijk} = \pi_{0jk} + \pi_{1jk}posttest_{ijk} + e_{ijk}$$

level-2 equations (student-level)

$$\begin{aligned} \pi_{0jk} = & \beta_{00k} + \beta_{01k}Black_{jk} + \beta_{02k}Hispanic_{jk} + \beta_{03k}White_{jk} \\ & + \beta_{04k}Hispanic_{jk} \times White_{jk} + Woman_{jk} \\ & \times (\beta_{05k} + \beta_{06k}Black_{jk} + \beta_{07k}Hispanic_{jk} \\ & + \beta_{08k}White_{jk} + \beta_{09k}Hispanic_{jk} \times White_{jk}) \\ & + \beta_{0(10)k}retake_{jk} + \beta_{0(11)k}gender\_other_{jk} \\ & + \beta_{0(12)k}race\_other_{jk} + r_{0jk} \end{aligned}$$

$$\begin{aligned} \pi_{1jk} = & \beta_{10k} + \beta_{11k}Black_{jk} + \beta_{12k}Hispanic_{jk} + \beta_{13k}White_{jk} \\ & + \beta_{14k}Hispanic_{jk} \times White_{jk} + Woman_{jk} \\ & \times (\beta_{15k} + \beta_{16k}Black_{jk} + \beta_{17k}Hispanic_{jk} \\ & + \beta_{18k}White_{jk} + \beta_{19k}Hispanic_{jk} \times White_{jk}) \end{aligned}$$

level-3 equations (course-level)

$$\beta_{00k} = \gamma_{00k} + u_{00k}$$

$$\beta_{0(1-12)k} = \gamma_{0(1-12)k}$$

$$\beta_{1(0-9)k} = \gamma_{1(0-9)k}$$

Woltam<sup>73</sup> provides a detailed description of HLM equations, which we will cover briefly here. The subscripts for  $Score_{ijk}$  refer to the  $i$ th assessment in the  $j$ th student in the  $k$ th course. In the level-1 equation, the  $\pi_{0jk}$  term represents the score before instruction for the  $j$ th student in the  $k$ th course. The  $\pi_{1jk}$  term represents the shift in scores from before to after instruction for the  $j$ th student in the  $k$ th course. The  $e_{ijk}$  term represents the assessment-level error associated with the  $i$ th assessment in the  $j$ th student in the  $k$ th course. The  $e_{ijk}$  term is analogous to the  $\varepsilon$  term in standard linear regressions, which is often omitted in representations of the equations. It represents the difference between the predicted and actual values. In the level-2 equation, the  $\beta_{00k}$  term represents the intercept for scores before instruction in the  $k$ th course. The  $\beta_{10k}$  represents the intercept for the shift in scores from before to after instruction in the  $k$ th course. The  $\beta_{0(1-12)k,1(1-12)k}$  terms represent the impact of gender and race variables on scores before and after instruction. The  $r_{0jk}$  term represents the student-level error associated with the  $j$ th student in the  $k$ th course and allows the intercept to vary across each student. In the level-3 equation, the  $\gamma_{00k}$  term is the intercept for the  $k$ th course. The  $\gamma_{0(1-12)k,1(1-12)k}$  terms represent the slopes (e.g., the regression coefficient) for each variable for the  $k$ th course. The  $u_{00k}$  term represents the course-level error associated with  $k$ th course and allows the intercept to vary across each course. The model is a fixed slope model since the slopes,  $\pi_{1jk}$ , and  $\beta_{0(1-12)k,1(1-12)k}$  equations do not include  $r$  or  $u$  variables.

To check the model assumptions, we eschewed doing sensitivity analysis because of its computational difficulty.<sup>74</sup> Visual inspection showed convergence for all variables.

### Model Interpretation

We used the overall gain from pre- to postinstruction to interpret the size of the educational debts. We reasoned that the average shift over one semester provided context for interpreting the educational debts. To account for uncertainty in the model, we used the standard error for each predicted score. An overlap between one standard error bars approximately produces a  $p$ -value of 0.05 for a one-sided  $t$  test. However, we did not use overlap as a binary indicator of significance (for more detail, see Nissen, Her Many Horses, and Van Dusen<sup>75</sup>).

## FINDINGS

The [Findings](#) section focuses on the predicted outcomes for each group by social identifier ([Table 3](#) and [Figure 3](#)). We do not examine the model's coefficients, which we include in the [Supporting Information](#), because this requires combining up to

**Table 3. Predicted Pretest and Posttest Scores from Our Bayesian Model Disaggregated by Gender and Race**

Social Identifiers		Pretest Scores ( <i>N</i> = 4612) <sup>a</sup>		Posttest Scores ( <i>N</i> = 4612) <sup>a</sup>	
Race	Gender	Est Correct Responses, %	SE	Est Correct Responses, %	SE
Asian	Women	43.1	1.9	47.1	2.1
	Men	48.0	1.9	52.9	2.0
Black	Women	35.2	2.2	42.6	2.3
	Men	38.9	2.6	45.0	3.1
Hispanic	Women	39.4	2.6	48.1	4.0
	Men	45.6	3.0	54.5	3.4
White	Women	39.8	1.6	44.1	1.6
	Men	49.5	1.6	52.4	1.7
White Hispanic	Women	34.5	2.3	40.3	2.8
	Men	48.0	2.6	50.4	2.5

<sup>a</sup>The Chemical Concept Inventory (CCI) has 22 multiple-choice questions. See ref 8.

16 coefficients to predict a group's score (e.g., White Hispanic women's posttest scores). We examine society's educational debts from racism, sexism, and then their intersectional effects. To interpret the size of the educational debts, we use four percentage points as the overall average gain from pre- to postinstruction given in the descriptive statistics.

#### Society's Educational Debts Before Instruction

**Racism.** The size of educational debts owed from racism before instruction (Figure 3 and Table 3) varied in size but existed for every group except Asian women, whose average pretest scores were 3.3 percentage points higher than White women's. For women, society owed the largest educational debts from racism to White Hispanic (5.3 percentage points) and Black (4.6 percentage points) students. Both of these educational debts were large compared to both the uncertainty in the measurement and the 4.0 percentage point average gain overall. For men, society owed the largest educational debts from racism to Hispanic (3.9 percentage points) and Black (10.5 percentage points) students. Both of these debts were large compared to the 4.0 percentage point average gain overall, and the educational debt owed to Black students was much larger than the uncertainty in the measurement. The uncertainty was large relative to the educational debts owed to Hispanic men largely because the small number of Hispanic

men in the data which resulted in larger uncertainties for their test scores in the models.

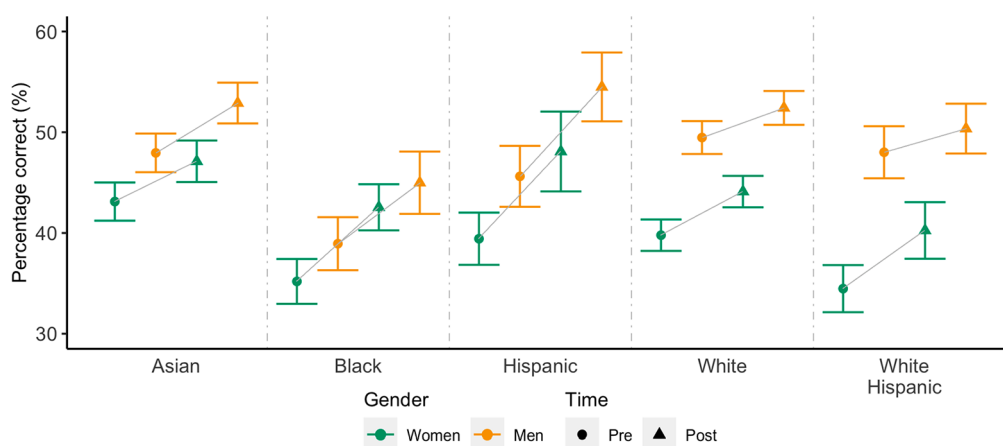
**Sexism.** Evidence of preexisting educational debts from sexism existed for all racial groups (Figure 3). The disparity in average scores between men and women on the pretest was smallest for Black students (3.7 percentage points) and largest for White Hispanic students (13.5 percentage points). All of the society's educational debts from sexism were larger than the uncertainty in the measurement.

**Intersectionality.** The size and source of society's educational debts from sexism and racism before instruction varied across groups. Society owed the largest educational debts to Black and White Hispanic women. The differences in educational debts owed across genders were smaller for Asian and Black students than for Hispanic, White, and White Hispanic students. This indicates that racism and sexism interact. The range of preexisting debts was larger for men (10.6 percentage points) than women (8.6 percentage points). This indicates that the intersectional nature of racism and sexism on chemistry knowledge is less than their sum. Aggregating student data by gender or race would obscure many of these educational debts.

#### Course Impact on Society's Educational Debts

**Racism.** Instruction reduced the size of educational debts owed from racism but did not eliminate them. For women, society's educational debts from racism decreased for White Hispanic students to 3.9 percentage points and Black students to 1.6 percentage points. Both of these debts were large compared to the 4.0 percentage point average gain overall, and the educational debt owed to White Hispanic students was larger than the uncertainty in the measurement. For men, society's educational debts from racism decreased for Black students to 7.4 percentage points and inverted for Hispanic students who outperformed their White peers by 2.1 percentage points. Both of these debts were large compared to the 4.0 percentage point average gain overall, and the educational debt owed to Black students was larger than the uncertainty in the measurement.

**Sexism.** Instruction reduced the size of educational debts owed from sexism for all groups other than Asian women but did not eliminate them. The disparity in average scores between men and women on the posttest was smallest for Black students (2.4 percentage points) and largest for White Hispanic students (10.1 percentage points). Society's educa-



**Figure 3.** Predicted pretest and posttest scores from our Bayesian model disaggregated by gender and race. Error bars are  $\pm 1$  SE.

tional debts from sexism were larger than the uncertainty in the measurement for Asian, White, and White Hispanic students. Society also owed educational debts from sexism to Black and Hispanic women consistent with those for other women. Still, the size of the debts was not as large relative to the uncertainty in the measure.

**Intersectionality.** The size and source of society's educational debts from sexism and racism before instruction varied across groups. Society still owed the largest educational debts to Black and White Hispanic women. The differences in educational debts owed across genders were still smallest for Black students but may have increased for Asian women. The range of preexisting debts was still larger for men (9.5 percentage points) than women (7.8 percentage points). This indicates that the intersectional nature of racism and sexism on chemistry knowledge is still less than their sum. Like the pretest, aggregating student data after instruction by gender or race would obscure many of these educational debts.

## DISCUSSION

Our model showed that White men had small predicted gains (2.4 percentage points). The chemistry courses repaid some of the educational debts, as is evidenced by the larger predicted gains for BIPOC students. Chemistry courses largely repaid educational debts and achieved Equality of Outcomes for Asian, Hispanic, and White Hispanic men. However, while the courses mitigated society's educational debts for Black students and women, they did not achieve Equality of Outcomes for them. These results were similar to Van Dusen and Nissen's<sup>76</sup> findings in physics, showing that society owed the largest educational debts to women and Black students. In physics, however, instruction increased society's educational debts from both racism and sexism.

Society's educational debts before instruction were large enough that women and Black men's average posttest scores *did not* reach the White men's average pretest scores. Women and Black men's average *posttest* scores were 2.4–9.2 percentage points below White men's average *pretest* scores. To contextualize the magnitude of society's educational debts, society would have to provide opportunities equivalent to taking this course two and a half times for White Hispanic women, who had the average learning gains of 5.8 percentage points, to repay the educational debt owed to White Hispanic women before instruction.

Hispanic students had the largest gains of all BIPOC groups. The gains were large enough to repay society's educational debts from racism owed to Hispanic men. While this indicates that courses meet some Hispanic students' needs, we caution drawing too large conclusions about it for two reasons. First, we do not see a similar trend in White Hispanic student's scores. Second, Hispanic students had the smallest representation in our data set (87 women and 62 men, some of whom only completed a pretest or posttest). While intersectionality tells us that Hispanic and White Hispanic students' experiences can be quite different, it limits the generalizability of the claim about the courses doing a good job serving all Hispanic students.

Society's much larger educational debts owed to Black women in chemistry align with Rocaboda and colleagues'<sup>77</sup> findings for Black women's achievement in chemistry and their experiencing less satisfaction in chemistry courses than their peers. After these chemistry courses, the unpaid educational debts that largely used collaborative instruction for women and

Black men align with Shortlidge and colleagues'<sup>35</sup> findings that collaborative instruction does not always support equity. We were surprised by the large educational debts society owes to women before and after instruction in chemistry because they do not align with women's equitable representation in earning undergraduate chemistry degrees.<sup>78</sup> These debts owed to women align with findings in physics<sup>76</sup> and with the lower representation of women in chemistry PhDs. The size of the educational debts indicated that some women interested in chemistry likely chose not to enroll in a course at all. This filtering effect would remove the students who are owed the largest debts from introductory chemistry courses and our sample, thereby biasing our findings to underreport the size of society's educational debts.

## LIMITATIONS

The data for this study comes from the LASSO platform. While LASSO offers a large database that has been found to include more diverse institutions than the literature,<sup>79</sup> it still underrepresents the institutions that disproportionately serve minoritized students (e.g., historically black colleges and universities, Hispanic-serving institutions, and 2-year colleges). Because LASSO is a platform hosted by the Learning Assistant Alliance,<sup>80</sup> 82% of the courses used Learning Assistants<sup>58,59</sup> to engage students in collaborative learning. Using the LASSO data also introduces potential selection bias, as instructors must know about LASSO and be interested in using a concept inventory in their course. This bias makes it likely that the instructors are more aware of research-based teaching practices and resources than the average instructor. All of the instructors self-reporting collaborative learning in their courses further reflected instructor knowledge of research-based teaching. These course commonalities could limit the generalizability of our findings in more traditional classroom contexts. While the CCI has undergone several rounds of validation research, it was developed at a highly competitive, primarily White-serving institution and has not undergone analysis to test for race or gender bias (e.g., DIF analysis). Expansion of the validation argument to include diverse student groups would strengthen CCI data's ability to support claims about equity in student learning.<sup>81</sup>

## CONCLUSION

Society owes large, consistent educational debts in chemistry knowledge from racism and sexism that demand attention. Current introductory chemistry courses mitigate many educational debts and eliminate educational debts from racism for Asian, Hispanic, and White Hispanic men. The repayment of these educational debts is encouraging. However, the CCI only measures one aspect of educational debt, and other aspects of educational debt undoubtedly remain. For example, Barr and colleagues'<sup>20</sup> found that, for premed students who completed their degrees, poor experiences in chemistry courses explained the large decrease in minoritized students' interest in and pursuit of medical degrees. Qualitative research on the experiences of minoritized students in chemistry courses similar to the work by Fries-Britt, Johnson and Burt,<sup>82</sup> Ong,<sup>83</sup> and McGee and Bentley<sup>84</sup> in STEM majors could inform how to increase the repayment of these educational debts to women and Black men.

Our study's findings of preexisting educational debts from sexism and failure to repay them show that sexism in chemistry

takes a toll on women's educational outcomes and contrasts several other findings. For example, our findings contrast Cheryan and colleagues<sup>78</sup> literature review that found smaller gender differences in early educational experiences and less masculine cultures in chemistry than in engineering, computer science, and physics. Our findings also contrast the near parity of women earning undergraduate degrees in chemistry (49%) while aligning with the inequity of doctoral degrees in chemistry<sup>85</sup> (38%). The parity in earning undergraduate degrees indicates that, despite the large educational debts owed to women, they possess resources and tools<sup>86</sup> to help them overcome the barriers in earning their undergraduate degrees.

As disciplinary researchers work to develop evidence-based, equitable instruction, these findings indicate that collaborative learning does not ensure equality of outcomes. The repayment of many educational debts from racism suggests that collaborative learning has the potential to produce equitable outcomes and provides a starting place for developing equitable instructional strategies. However, if not intentionally structured to do otherwise, collaborative learning can replicate the oppressive power structures in broader society.

The educational debts prior to instruction reveal the inequities in primary and secondary education that minoritized students face. Similarly, inequities after graduation, such as in the pursuit of medical degrees, reflect educational debts from undergraduate education. As college educators, we can suffer from a tunnel vision that focuses on what happens during college or college classes. As citizens and scientists, however, we should take a role in addressing the vast inequities in primary and secondary education. We should attend to our role as educators in addressing the inequities and systemic oppression outside of the classroom and after college.

## ■ ASSOCIATED CONTENT

### SI Supporting Information

The Supporting Information is available at <https://pubs.acs.org/doi/10.1021/acs.jchemeduc.1c00352>.

- Author positionality (PDF)
- Descriptive statistics (PDF)
- Model output (PDF)
- Social identity questionnaire (PDF)

## ■ AUTHOR INFORMATION

### Corresponding Author

**Ben Van Dusen** – School of Education, Iowa State University, Ames, Iowa 50011, United States; [orcid.org/0000-0003-1264-0550](https://orcid.org/0000-0003-1264-0550); Email: [bvd@iastate.edu](mailto:bvd@iastate.edu)

### Authors

**Jayson Nissen** – Nissen Education Research and Design, Corvallis, Oregon 97330, United States

**Robert M. Talbot** – School of Education and Human Development, University of Colorado Denver, Denver, Colorado 80217, United States

**Hannah Huvad** – School of Education and Human Development, University of Colorado Denver, Denver, Colorado 80217, United States

**Mollee Shultz** – Department of Physics, Texas State University, San Marcos, Texas 78666, United States

Complete contact information is available at:

<https://pubs.acs.org/10.1021/acs.jchemeduc.1c00352>

## Notes

The authors declare no competing financial interest.

## ■ ACKNOWLEDGMENTS

This work was funded in part by NSF-IUSE Grant DUE-1928596.

## ■ REFERENCES

- (1) Wilson-Kennedy, Z. S.; Payton-Stewart, F.; Winfield, L. L. Toward intentional diversity, equity, and respect in chemistry research and practice. *J. Chem. Educ.* **2020**, *97*, 2041–2044.
- (2) National Academy of Sciences, National Academy of Engineering, and Institute of Medicine. *Rising above the Gathering Storm, Revisited: Rapidly Approaching Category 5*; The National Academies Press: Washington, DC, 2010; DOI: [10.17226/12999](https://doi.org/10.17226/12999).
- (3) Committee on Prospering in the Global Economy of the 21st Century. *Rising above the Gathering Storm: Energizing and Employing America for a Brighter Economic Future*; National Academies Press: Washington, DC, 2007; DOI: [10.17226/11463](https://doi.org/10.17226/11463).
- (4) National Research Council. *Discipline-Based Education Research: Understanding and Improving Learning in Undergraduate Science and Engineering*; National Academies Press, 2013. DOI: [10.1002/jee.20030](https://doi.org/10.1002/jee.20030).
- (5) National Academies of Sciences. *Minority Serving Institutions: America's Underutilized Resource for Strengthening the STEM Workforce*; National Academies Press, 2019. DOI: [10.17226/25257](https://doi.org/10.17226/25257).
- (6) National Science Foundation. *Women, Minorities, and Persons with Disabilities in Science and Engineering*; 2019.
- (7) Stage, F. K. Answering critical questions using quantitative data. *New directions for institutional research* **2007**, *2007*, 5–16.
- (8) Mulford, D. R.; Robinson, W. An inventory for alternate conceptions among first-semester general chemistry students. *J. Chem. Educ.* **2002**, *79*, 739.
- (9) Madsen, A.; McKagan, S. B.; Martinuk, M. S.; Bell, A.; Sayre, E. C. Research based assessment affordances and constraints: Perceptions of physics faculty. *Physical Review Physics Education Research* **2016**, *12*, 010115 DOI: [10.1103/PhysRevPhysEduRes.12.010115](https://doi.org/10.1103/PhysRevPhysEduRes.12.010115).
- (10) Van Dusen, B. *LASSO: A New Tool to Support Instructors and Researchers*; American Physical Society Forum on Education, Fall, 2018; pp 12–14.
- (11) Ladson-Billings, G. From the Achievement Gap to the Education Debt: Understanding Achievement in U.S. Schools. *Educational Researcher* **2006**, *35*, 3–12.
- (12) Ladson-Billings, G. Pushing past the achievement gap: An essay on the language of deficit. *Journal of Negro Education* **2007**, *76*, 316–323. <https://www.jstor.org/stable/40034574>.
- (13) We use the term minoritized to reflect that students are categorized as minorities through an active social process, rather than a characteristic of the individual.
- (14) Delgado, R.; Stefancic, J. *Critical Race Theory: An Introduction*; New York University Press, 2017.
- (15) Ladson-Billings, G. Just what is critical race theory, and what's it doing in a nice field like education?. *Race Is... Race Isn't*; Routledge, 2019; pp 7–30.
- (16) White, K. N.; Vincent-Layton, K.; Villarreal, B. Equitable and inclusive practices designed to reduce equity gaps in undergraduate chemistry courses. *J. Chem. Educ.* **2021**, *98*, 330.
- (17) National Science Foundation. *Women, Minorities, and Persons with Disabilities in Science and Engineering*; 2017.
- (18) Bancroft, S. F.; Fowler, S. R.; Jalaeian, M.; Patterson, K. Leveling the Field: Flipped Instruction as a Tool for Promoting Equity in General Chemistry. *J. Chem. Educ.* **2020**, *97*, 36.
- (19) Becker, M. L.; Nilsson, M. R. College Chemistry Textbooks Fail on Gender Representation. *J. Chem. Educ.* **2021**, *98*, 1146.



- (20) Barr, D. A.; Gonzalez, M. E.; Wanat, S. F. The Leaky Pipeline: Factors Associated With Early Decline in Interest in Premedical Studies Among Underrepresented Minority Undergraduate Students. *Academic Medicine* **2008**, *83*, 503–511.
- (21) Dalgety, J.; Coll, R. K. Exploring First-Year Science Students' Chemistry Self-Efficacy. *International Journal of Science and Mathematics Education* **2006**, *4*, 97–116.
- (22) Fink, A.; Cahill, M. J.; McDaniel, M. A.; Hoffman, A.; Frey, R. F. Improving general chemistry performance through a growth mindset intervention: selective effects on underrepresented minorities. *Chem. Educ. Res. Pract.* **2018**, *19*, 783–806.
- (23) Harris, R. B.; Mack, M. R.; Bryant, J.; Theobald, E. J.; Freeman, S. Reducing achievement gaps in undergraduate general chemistry could lift underrepresented students into a “hyperpersistent zone.” *Science Advances* **2020**, *6*, eaaz5687.
- (24) King, D.; Domin, D. S. The Representation of People of Color in Undergraduate General Chemistry Textbooks. *J. Chem. Educ.* **2007**, *84*, 342.
- (25) Lewis, S. E.; Lewis, J. E. Seeking effectiveness and equity in a large college chemistry course: An HLM investigation of peer-led guided inquiry. *J. Res. Sci. Teach.* **2008**, *45*, 794–811.
- (26) Rath, K. A.; Peterfreund, A.; Bayliss, F.; Runquist, E.; Simonis, U. Impact of Supplemental Instruction in Entry-Level Chemistry Courses at a Mid-sized Public University. *J. Chem. Educ.* **2012**, *89*, 449.
- (27) Richards-Babb, M.; Jackson, J. K. Research and Practice Gendered responses to online homework use in general chemistry. *Chem. Educ. Res. Pract.* **2011**, *12*, 409–419.
- (28) Rocabado, G. A.; Kilpatrick, N. A.; Mooring, S. R.; Lewis, J. E. Can We Compare Attitude Scores among Diverse Populations? An Exploration of Measurement Invariance Testing to Support Valid Comparisons between Black Female Students and Their Peers in an Organic Chemistry Course. *J. Chem. Educ.* **2019**, *96*, 2371.
- (29) Sass, T. R. *Understanding the STEM Pipeline*; Calder American Institutes for Research, Georgia State University, 2015.
- (30) Shortlidge, E. E.; Rain-Griffith, L.; Shelby, C.; Shusterman, G. P.; Barbera, J. Despite Similar Perceptions and Attitudes, Post-baccalaureate Students Outperform in Introductory Biology and Chemistry Courses. *LSE* **2019**, *18*, ar3.
- (31) Stanich, C. A.; Pelch, M. A.; Theobald, E. J.; Freeman, S. A new approach to supplementary instruction narrows achievement and affect gaps for underrepresented minorities, first-generation students, and women. *Chem. Educ. Res. Pract.* **2018**, *19*, 846–866.
- (32) Sunny, C. E.; Taasooobshirazi, G.; Clark, L.; Marchand, G. Stereotype threat and gender differences in chemistry. *Instructional Science* **2017**, *45*, 157–175.
- (33) Villafane, S. M.; Garcia, C. A.; Lewis, J. E. Research and Practice self-efficacy throughout a semester of college-level preparatory chemistry. *Chem. Educ. Res. Pract.* **2014**, *15*, 114–127.
- (34) Witherspoon, E. B.; Vincent-ruz, P.; Schunn, C. D. When Making the Grade Isn't Enough: The Gendered Nature of Premed Science Course Attrition. *Educ. Res.* **2019**, *48*, 193–204.
- (35) Shortlidge, E. E.; Rain-Griffith, L.; Shelby, C.; Shusterman, G. P.; Barbera, J. Despite similar perceptions and attitudes, post-baccalaureate students outperform in introductory biology and chemistry courses. *CBE Life Sci. Educ.* **2019**, *18*, 18:ar3 DOI: 10.1187/cbe.17-12-0289.
- (36) Sands, D.; Parker, M.; Hedgeland, H.; Jordan, S.; Galloway, R. Using concept inventories to measure understanding. *High. Educ. Pedagog.* **2018**, *3*, 173.
- (37) Bauer, C. F. Beyond “Student Attitudes”: Chemistry Self-Concept Inventory for Assessment of the Affective Component. *J. Chem. Educ.* **2005**, *82*, 1864 DOI: 10.1021/ed082p1864.
- (38) Wren, D.; Barbera, J. Gathering Evidence for Validity during the Design, Development, and Qualitative Evaluation of Thermochemistry Concept Inventory Items. *J. Chem. Educ.* **2013**, *90*, 1590.
- (39) Krause, S.; Birk, J.; Bauer, R.; Jenkins, B.; Pavelich, M. J. Development, testing, and application of a chemistry concept inventory. *34th Annual Frontiers in Education* **2004**, *34*, 103.
- (40) Barbera, J. A psychometric analysis of the chemical concepts inventory. *J. Chem. Educ.* **2013**, *90*, 546–553.
- (41) Schwartz, P.; Barbera, J. Evaluating the content and response process validity of data from the chemical concepts inventory. *J. Chem. Educ.* **2014**, *91*, 630–640.
- (42) Schultz, M.; Lawrie, G. A.; Bailey, C. H.; Bedford, S. B.; Dargaville, T. R.; O'Brien, G.; Tasker, R.; Thompson, C. D.; Williams, M.; Wright, A. H. Evaluation of diagnostic tools that tertiary teachers can apply to profile their students' conceptions. *Int. J. of Science Educ.* **2017**, *39*, 565–586.
- (43) Sleeter, C. E.; Bernal, D. Critical Pedagogy, Critical Race Theory, and Antiracist Education: Implications for Multicultural Education. *Handbook of Research in Multicultural Education*; Jossey-Bass, 2003; pp 240–258.
- (44) West, C. *Critical Race Theory: The Key Writings That Formed the Movement*; The New Press, 1995.
- (45) Ladson-Billings, G. Critical Race Theory—What it is Not! *Handbook of Critical Race Theory in Education* **2014**, 34–47.
- (46) Crenshaw, K. Mapping the Margins: Intersectionality, Identity Politics, and Violence Against Women of Color. *Stanford Law Rev.* **1991**, *43*, 1241–1299.
- (47) Ladson-Billings, G. Race Still Matters: Critical Race Theory in Education. *Routledge International Handbook of Critical Education*; Routledge, 2009; pp 110–122.
- (48) Gillborn, D.; Warmington, P.; Demack, S. QuantCrit: education, policy, ‘Big Data’ and principles for a critical race theory of statistics. *Race Ethnicity and Education* **2018**, *21*, 158–179.
- (49) Garcia, N. M.; López, N.; Vélez, V. N. QuantCrit: rectifying quantitative methods through critical race theory. *Race Ethnicity and Educ.* **2018**, *21*, 149–157.
- (50) López, N.; Erwin, C.; Binder, M.; Chavez, M. J. Making the invisible visible: advancing quantitative methods in higher education using critical race theory and intersectionality. *Race Ethnicity and Educ.* **2018**, *21*, 180–207.
- (51) Jang, S. T. The Implications of Intersectionality on Southeast Asian Female Students' Educational Outcomes in the United States: A Critical Quantitative Intersectionality Analysis. *American Educational Research Journal* **2018**, *55*, 1268–1306.
- (52) In this publication, we capitalize all races, including White, emphasizing that there is no default race and that they are all social constructs with associated sets of cultural practices.
- (53) Rodriguez, I.; Brewé, E.; Sawtelle, V.; Kramer, L. H. Impact of equity models and statistical measures on interpretations of educational reform. *Phys. Rev. Spec. Top. - Phys. Educ. Res.* **2012**, *8*, 1–7.
- (54) Kendi, I. X. *Stamped from the Beginning: The Definitive History of Racist Ideas in America*; Hachette UK, 2016.
- (55) Secada, W. G. Agenda Setting, Enlightened Self-Interest, and Equity in Mathematics Education. *Peabody Journal of Education* **1989**, *66*, 22–56.
- (56) Espinoza, O. Solving the equity–equality conceptual dilemma: a new model for analysis of the educational process. *Educational Research* **2007**, *49*, 343–363.
- (57) Lee, O. Equity Implications Based on the Conceptions of Science Achievement in Major Reform Documents. *Review of educational research* **1999**, *69*, 83–115.
- (58) Otero, V. Nationally scaled model for leveraging course transformation with physics teacher preparation. In *Recruiting and Educating Future Physics Teachers: Case Studies and Effective Practices*; Sandifer, C., Brewé, E., Eds.; American Physical Society, 2015; pp 107–116.
- (59) Barrasso, A. P.; Spilios, K. E. A scoping review of literature assessing the impact of the learning assistant model. *International Journal of STEM Education* **2021**, DOI: 10.1186/s40594-020-00267-8.
- (60) Speidel, M.; Drechsler, J.; Jolani, S. R package hmi: a convenient tool for hierarchical multiple imputation and beyond; No. 16/2018. IAB-Discussion Paper, 2018.
- (61) van Buuren, S.; Groothuis-Oudshoorn, K.; Robitzsch, A.; Vink, G.; Doove, L.; Jolani, S. *Package ‘mice’*; 2015.

- (62) Allison, P. D. *Missing Data*; Sage Publications, 2001.
- (63) Bui, E. R.; Goodson, P.; Neilands, T. B. Out of sight, not out of mind: Strategies for handling missing data. *Am. J. Health Behav.* **2008**, *32*, 83–92.
- (64) Manly, C. A.; Wells, R. S. Reporting the use of multiple imputation for missing data in higher education research. *Res. High. Educ.* **2015**, *56*, 397–409.
- (65) Nissen, J.; Donatello, R.; Van Dusen, B. Missing data and bias in physics education research: A case for using multiple imputation. *Phys. Rev. Phys. Educ. Res.* **15**, **2019**. DOI: 10.1103/PhysRevPhysEduRes.15.020106.
- (66) Raudenbush, S.; Bryk, A. *Hierarchical Linear Models: Applications and Data Analysis Methods*; Sage, 2002.
- (67) Stan Development Team. *RStan: the R interface to Stan*, R package version 2.1; 2016; p 522.
- (68) Bürkner, P. C. brms: An R package for Bayesian multilevel models using Stan. *J. Stat. Softw.* **80**, **2017**. DOI: 10.18637/jss.v080.i01.
- (69) Simmons, J. P.; Nelson, L. D.; Simonsohn, U. False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychol. Sci.* **2011**, *22*, 1359–1366.
- (70) Johnson, J. B.; Omland, K. S. Model selection in ecology and evolution. *Trends Ecol. Evol.* **2004**, *19*, 101–108.
- (71) Burnham, K. P.; Anderson, D. R. Multimodel inference: Understanding AIC and BIC in model selection. *Sociol. Methods Res.* **2004**, *33*, 261–304.
- (72) Barton, K. *MuMIn: Multi-model Inference*; 2009; <http://r-forge.r-project.org/projects/mumin/>.
- (73) Woltman, H.; Feldstain, A.; MacKay, C.; Rocchi, M. An introduction to hierarchical linear modeling. *Tutorials in quantitative methods for psychology* **2012**, *8*, 52–69.
- (74) Gelma, A.; Meng, X.-L. Model checking and model improvement. *Markov Chain Monte Carlo in Practice* **1995**, 207–220.
- (75) Nissen, J. M.; Her Many Horses, I.; Van Dusen, B. Investigating society's educational debts due to racism and sexism in student attitudes about physics using quantitative critical race theory. *Phys. Rev. Phys. Educ. Res.* **2021**, *17*, 10116.
- (76) Van Dusen, B.; Nissen, J. Equity in college physics student learning: A critical quantitative intersectionality investigation. *J. Res. Sci. Teach.* **2020**, *57*, 33–57.
- (77) Rocabado, G. A.; Komperda, R.; Lewis, J. E.; Barbera, J. Research and Practice Addressing diversity and inclusion through group comparisons: a primer on measurement invariance testing. *Chem. Educ. Res. Pract.* **2020**, *21*, 969–988.
- (78) Cheryan, S.; Ziegler, S. A.; Montoya, A.; Jiang, L. Why are some STEM fields more gender balanced than others? *Psychol. Bull.* **2017**, *143*, 1–35, DOI: 10.1037/bul0000052.
- (79) Nissen, J.; Van Dusen, B.; Close, E. W.; Her Many Horse, I.; Jariwala, M. Providing Context for Identifying Effective Introductory Mechanics Courses. *Physics Teacher*, accepted.
- (80) Otero, V.; Langdon, L.; Her Many Horses, I.; Oatley, M.; Van Dusen, B. Learning Assistant Alliance: Social Organizing Tools for Sharing Resources and Building Institutional Networks. In *Envisioning the Future of Undergraduate STEM Education Conference*; 2016; pp 1–7.
- (81) Padilla, A. Quantitative Methods in Multicultural Education Research. *Handbook of Research on Multicultural Education* **2004**, 127–145.
- (82) Fries-Britt, S. L.; Johnson, J.; Burt, B. Black students in physics. In *Living at the Intersections*; Strayhorn, T. L., Ed.; Information Age Publishing, 2013; pp 21–39.
- (83) Ong, M. Body projects of young women of color in physics: Intersections of gender, race, and science. *Social problems* **2005**, *52*, 593–617.
- (84) McGee, E. O.; Bentley, L. The Troubled Success of Black Women in STEM. *Cognition and Instruction* **2017**, *35*, 265–289.
- (85) National Science Foundation. *The State of U.S. Science and Engineering*; 2020; DOI: 10.1097/00001888-199005000-00011.
- (86) Yosso, T. J. Whose culture has capital? A critical race theory discussion of community cultural wealth. *Race Ethnicity and Education* **2005**, *8* (1), 69–91.

## Positionality

Ben Van Dusen's positionality statement: I identify as a White cisgender, heterosexual man. I was raised in a pair of lower-income households, but I now earn an upper-middle-class income. I hold an undergraduate degree in physics and a Ph.D. in education. I was an assistant professor at a teaching-intensive, Hispanic-serving institution but now am an assistant professor at a research-intensive, predominantly White institution. My experiences working with minoritized students, particularly the Latino and Latina students I have had the honor to mentor as Learning Assistants and researchers, have motivated me to use my position and privilege to dismantle oppressive power structures. As someone who seeks to serve as a co-conspirator, it is easy to overlook my privileges. I try to broaden my perspective through feedback from those with more diverse lived experiences than my own.

Jayson Nissen's positionality statement: Identifying as a White, cisgendered, heterosexual man provides me with opportunities denied to others in American society. My experience growing up poor and serving in the all-male submarine service motivate reflecting on and working to dismantle my privilege and oppressive power structures in science. Because I am not a woman or a person of color, I brought a limited perspective to this work on racism and sexism.

Robert (Bud) Talbot's positionality statement: I identify as a white heterosexual man who was raised in a middle class household in a predominately white midwestern university town. My family has a history of working in education and my mother has always served as my professional role model. I have taught secondary science in a variety of contexts: an all American Indian school on the Navajo Reservation, an all white Junior/Senior high school in rural America, and a majority Black high school in suburban Houston, Texas. I currently teach and research at an ethnically and socioeconomically diverse, urban research university. I have largely succeeded and have earned tenure in a system that was designed for people like me- it has been hard for me to fail. All of these experiences and my privilege have greatly shaped my understanding of educational opportunity, and my worldview. These experiences, my work, and

my personal life have motivated me to elevate the voices of others, including those minoritized by our society and systems, above my own.

Hannah Huvad's positionality statement: I identify as a Jewish heterosexual woman. As a Jewish woman growing up in Southern California, I understand and know what it feels like to be outside the dominant culture in a classroom. I would often find myself as the "token Jew" in school settings. This experience and my Jewish identity give me a unique lens in which to approach my teaching, mentoring, and research. I have taught at three different Hispanic-Serving Institutes and I mentored low-income high school students through the Trio Upward Bound program. Through these experiences, and my own lived experiences as a Jewish woman in science, I am motivated to continuously check my own privilege and assumptions as well as actively work to dismantle oppressive systems within education.

Mollee Shultz's positionality statement: I identify as a Chinese-American, heterosexual woman. I was adopted and grew up in a White household. I hold a master's degree in mathematics and a Ph.D. in mathematics education. I have worked as adjunct faculty teaching mathematics at three two-year Hispanic-serving institutions. As a female in a male-dominated space with my Chinese heritage, I am familiar with not appearing like others in my community and the extra work it requires. As a mother of two mixed-race children, I am interested in doing work that helps shape STEM fields into inclusive spaces for them and others.

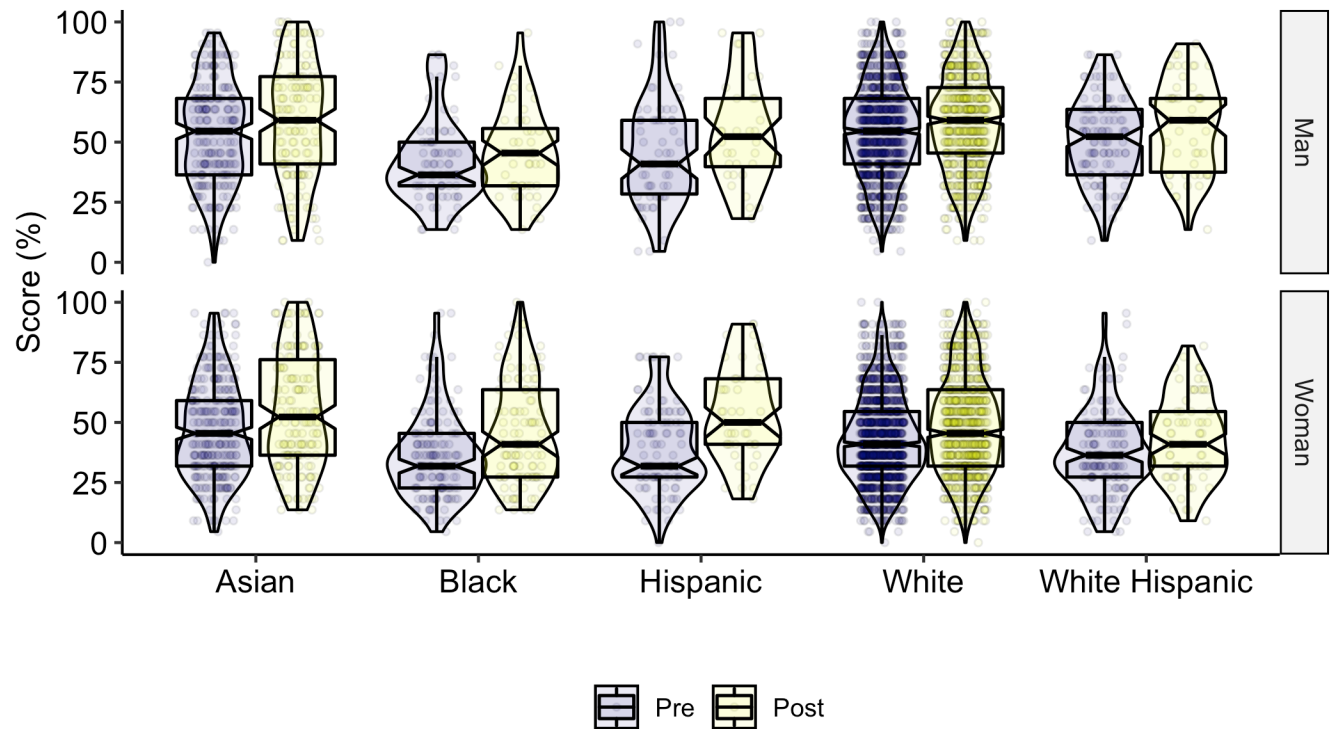
## Descriptive statistics

Table 1 breaks down the descriptive statistics across social identification groups used in the final model. Figure 1 includes violin plots, box plots, and jittered scatter plots for the data before imputation. We included these plots because they provide data transparency for readers. They break down the ‘gap gazing’ perspective that means and standard errors from either descriptive statistics or statistical models can reinforce. In other words, the data show differences in the mean scores between groups (inter-group variance), but the spread of scores within each group (intra-group variance) is much larger, which is shown by the overlap in the distributions across all groups. All the plots in Fig. 1 show differences but not gaps across groups.

Table 1. Descriptive Statistics Disaggregated by Race and Gender for the Multiply Imputed Data

Social Identifiers		Pretest Scores ( $N = 4612$ ) <sup>a</sup>		Posttest Scores ( $N = 4612$ ) <sup>a</sup>		
Race	Gender	N	Mean (%)	S.D.	Mean (%)	S.D.
Asian	Women	335	47.2	19.9	54.7	22.0
	Men	271	53.7	20.3	58.8	22.7
Black	Women	202	36.0	17.5	44.9	20.7
	Men	97	42.2	17.5	48.6	19.8
Hispanic	Women	118	39.3	17.5	51.8	20.4
	Men	76	46.7	21.2	59.1	21.8
White	Women	1715	43.2	18.1	49.3	20.3
	Men	1329	54.1	18.7	58.0	19.8
White	Women	174	38.0	17.2	46.1	20.3
Hispanic	Men	134	51.3	17.2	55.5	20.2

<sup>a</sup>The Chemical Concept Inventory (CCI) has 22 multiple-choice questions. See ref 8.



**Figure 1.** Violin plots, boxplots, and scatter plots of the data before imputation for pretest and posttest CCI scores. The violin plot is a reflected density plot to show the distribution of the data. The notched boxplots show the distribution but focus on the medians with notches to show the 95% confidence intervals. The scatter plot is jittered to randomly distribute the points left and right of the vertical axis for clarity. The scatter plot illustrates the number of data points in each group and adds detail to how the data is distributed, particularly in the tails. The plots show no indication of ceiling or floor effects. These plots show easily identifiable differences across groups but also show these differences are shifts in the distributions of scores and not gaps that separate groups.

## Model Output

Table 1. Bayesian Hierarchical Linear Model Coefficients and Standard Errors for Predicted Average Test Scores

Variable	Coefficient	S.E.
Intercept	47.94	1.93
retake	3.5	0.83
gender_other	2.1	2.83
race_other	4.17	1.6
posttest	4.97	1.67
women	-4.8	1.71
Black	-8.99	2.59
Hispanic	-2.31	2.93
White	1.53	1.39
posttest*women	-0.93	2.25
Hispanic*White	0.83	3.78
posttest*Black	1.16	3.56
posttest*Hispanic	4.05	4.35
posttest*White	-2.01	1.86
women*Black	1.04	3.16
women*Hispanic	-1.38	3.75
women*White	-4.88	1.9
posttest*Hispanic*White	-4.65	5.15
women*Hispanic*White	-2.47	4.97
posttest*women*Black	2.22	4.11
posttest*women*Hispanic	0.84	5.82
posttest*women*White	2.32	2.64
posttest*women*Hispanic*White	1.01	6.92

## Social identity questionnaire

Table 1. Optional Social Identity Questions from the LASSO Survey

Question	Options
How do you describe your gender? (Check all that apply)	Man Woman Transman Transwoman Genderqueer/Gender non-conforming Another identity (please state): Prefer not to answer
What is your ethnicity?	Not of Hispanic/Latinx/Spanish origin Mexican/Mexican American/Chicano/Chicana Puerto Rican Cuban Another Hispanic/Latinx/Spanish origin (please state): Prefer not to answer
What is your race? (Check all that apply)	White Black or African American American Indian or Alaskan Native Chinese Filipino Asian Indian Vietnamese Japanese Korean An Asian race not listed Native Hawaiian Samoan Chamorro A Pacific Islander Race not listed Middle Eastern or North African Another race/other races (please state): Prefer not to answer