

Comparing the Efficacy of Fixed-Effects and MAIHDA Models in Predicting Outcomes for Intersectional Social Strata

Ben Van Dusen¹ , Heidi Cian², Jayson Nissen³, Lucy Arellano⁴, and Adrienne D. Woods⁵

Abstract

This investigation examines the efficacy of multilevel analysis of individual heterogeneity and discriminatory accuracy (MAIHDA) over fixed-effects models when performing intersectional studies. The research questions are as follows: (1) What are typical strata representation rates and outcomes on physics research-based assessments? (2) To what extent do MAIHDA models create more accurate predicted strata outcomes than fixed-effects models? and (3) To what extent do MAIHDA models allow the modeling of smaller strata sample sizes? We simulated 3,000 data sets based on real-world data from 5,955 students on the LASSO platform. We found that MAIHDA created more accurate and precise predictions than fixed-effects models. We also found that using MAIHDA could allow researchers to disaggregate their data further, creating smaller group sample sizes while maintaining more accurate findings than fixed-effects models. We recommend using MAIHDA over fixed-effects models for intersectional investigations.

Keywords

critical race theory in education, gender, quantitative research on education, race, STEM

Originating in the 1980s, the concept “intersectionality” highlights “the vexed dynamics of difference and the solidarities of sameness in the context of antidiscrimination and social movement politics” (Cho, Crenshaw, and McCall 2013:787). Following Crenshaw’s (1990) introduction of the term to understand the experiences of Black women, researchers have expanded applications of intersectionality to include diverse identities and disciplines (Carbado et al. 2013; Collins 2019; Collins and Bilge, 2020; Crenshaw, 1989; Harris and Patton 2019; Mena and Bolte 2019; Nash, 2008; Teffera et al. 2018). Collins (2015) advocates for developing intersectionality as a critical social theory to understand and change the existing social order. To use intersectionality in this way,

Collins details four principles of power and social structures: (1) Social strata (e.g., race, class, and gender) act as markers of power that are interdependent and mutually constructed; (2) power relations across the intersections of these power markers create

¹Iowa State University, Ames, IA, USA

²Maine Mathematics and Science Alliance, Augusta, ME, USA

³Nissen Education Research and Design, Slidell, LA, USA

⁴Texas Technology University, Lubbock, TX, USA

⁵SRI, Menlo Park, CA, USA

Corresponding Author:

Ben Van Dusen, Iowa State University, 1918 Stevenson Dr, Ames, IA 50011-2140, USA.

Email: bvd@iastate.edu

complex, interdependent inequalities; (3) individuals' and groups' locations within these intersecting power relations shape their experiences and perspectives; and (4) solving social problems requires intersectional analyses specific to the context of the social problem.

Choo and Ferree's (2010) identification of group-, process-, and systems-centered approaches to studying intersectionality highlights the breadth of its possibilities. Group-centered research centers the voices of multiple marginalized groups, aligning with McCall's (2005) anti-categorical or intracategorical paradigms. Anti-categorical complexity challenges the use of categories that may reinforce inequalities, whereas intracategorical complexity examines specific intersectional strata to understand lived experiences better. Process-centered research focuses on power dynamics and the intersecting oppressions or privileges, often corresponding with McCall's intracategorical and intercategory paradigms. Intercategorical complexity involves analyzing existing strata to explore inequalities and their evolution. System-centered research examines the interactions among multiple forms of oppression, moving beyond a single-system focus to encompass complexities like racism, sexism, and classism (Willis 1981). This approach typically aligns with intercategory complexity, facilitating comparisons across social strata and exploring various social locations, such as STEM fields or teaching methods.

Given the breadth of approaches to studying intersectionality, Cho et al. (2013) highlight the necessity for a dedicated Intersectionality Studies field, and Collins (2019) envisions a vibrant intersectional community that combines grassroots and top-down theories, embracing diverse perspectives, methodologies, and disciplines to overcome theoretical limitations. Both advocate for intersectionality as a vital practice aimed at challenging and reforming unjust systems. Quantitative intersectional research can contribute to these goals by creating pathways for marginalized groups in STEM, through innovative teaching methods, and by redefining participation in STEM fields. Collins's (2015) four principles guide quantitative researchers to develop intersectional models that are "meaningful" rather than merely statistically "significant." The term "meaningful" is crucial here because reliance on statistical significance for model selection can overlook educationally relevant differences (Wasserstein and Lazar 2016) and yield misleading outcomes (Van Dusen and Nissen

2022). Additionally, these principles emphasize the necessity of considering power dynamics beyond individual social strata in models, including instructional methods and school characteristics.

Salem (2018) describes intersectionality as a theory that travels across time, place, and space; it has been adapted and used differently across contexts and communities. For instance, the rise of critical quantitative theories has brought intersectionality into quantitative research (Tabron and Thomas 2023). However, challenges in adapting intersectionality quantitatively have hindered practical and theoretical advances (Bowleg, 2008). Bauer et al. (2021) highlight this, noting a significant gap in theoretical consistency in quantitative intersectionality studies.

Quantitative STEM equity research often groups marginalized populations (e.g., underrepresented minorities [URMs]) and analyzes oppression separately (e.g., racism or sexism). These choices, largely driven by institutional norms rather than theory (e.g., the National Science Foundation's categorization of URMs), aim for statistical significance but can mask true inequities (Shafer, Mahmood, and Stelzer 2021; Wasserstein and Lazar 2016). Incorporating intersectionality could enhance understanding of inequities and inform equitable educational reforms (Van Dusen and Nissen 2020a, 2020b). However, this approach demands analytic methods capable of addressing the complex interplay of social identities, power dynamics, and discipline-specific practices.

In examining intersectionality in feminist literature, McCall (2005) outlines three approaches: anti-categorical, intracategorical, and intercategory. Quantitative research, which inherently categorizes, typically uses intra- or intercategory methods, with the latter being more prevalent (Bauer et al. 2021). The fixed-effects model is frequently used to assess intersectional outcomes, incorporating main effects for each aspect of intersectional social strata (e.g., race, gender) and all of their potential interactions (Evans 2019; Evans, Leckie, and Merlo 2020). Intersectional social strata are the provisionally adopted analytic categories (e.g., the combination of race, gender, and first-generation designation) used to document inequalities. This approach, while offering nuanced insights, becomes increasingly complex and data-intensive with each added identity stratum, leading to challenges in achieving sufficient statistical power for reliable outcomes.

Researchers have proposed a multilevel analysis of individual heterogeneity and discriminatory

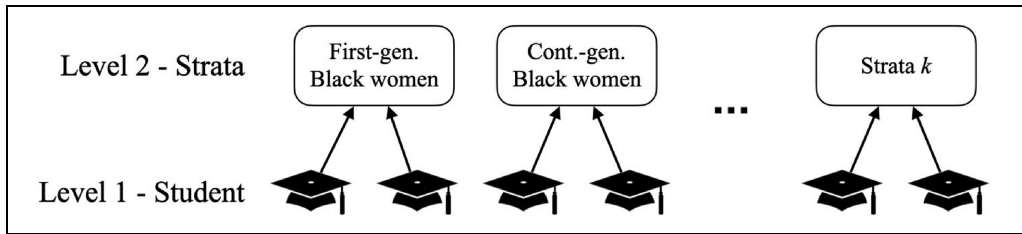


Figure 1. Multilevel analysis of individual heterogeneity and discriminatory accuracy (MAIHDA) nesting students within k strata.

accuracy (MAIHDA; Evans 2015, 2019; Evans et al. 2018; Merlo 2018) to address some of the shortcomings of prior methods. To improve model predictions across strata, MAIHDA nests individuals within their social identities (see Figure 1; Evans et al. 2020; Keller et al. 2023). By combining the main effects with the variance terms for each strata, MAIHDA can create more accurate predictions than adding the primary terms without including interaction terms. MAIHDA offers several potential advantages over fixed-effects models. First, MAIHDA aligns with an intersectional perspective by including all strata in a model regardless of their sample size. Second, by reducing the number of terms in the model, MAIHDA reduces the statistical power requirements. Third, MAIHDA seeks to increase the accuracy of predictions by drawing on the shrinkage, or partial pooling, that nesting within each aspect of a strata provides in a multilevel model (Raudenbush and Bryk 1986). Shrinkage allows multilevel models to make predictions for each strata that are informed by the predictions made for other strata. Shrinkage benefits are likely to be strongest for small- N strata where the small numbers limit the ability of intersectional models to disaggregate outcomes.

MAIHDA offers significant potential for modeling intersecting social identities, but its empirical validation requires further simulation studies. Prior work has used simplified simulations (Bell, Holman, and Jones 2019; Evans et al. 2020; Lizotte et al. 2020) or focused on health outcomes (Mahendran, Lizotte, and Bauer 2022a, 2022b), leaving a gap in demonstrating MAIHDA's effectiveness in educational settings. Specifically, there is a need to show how MAIHDA can enhance prediction accuracy for educational strata, particularly when examining smaller sample sizes. Additionally, simulations have yet to explore MAIHDA's application to education's hierarchical data structure,

such as measurements within students within courses within schools.

In our study, we compare MAIHDA and fixed-effects models to evaluate their performance in analyzing intersectional social strata outcomes. We use the Force Concept Inventory (FCI; Hestenes, Wells, and Swackhamer 1992) for our simulations, set against the backdrop of physics—a field noted for its inequities (Brewer and Sawtelle 2016). Our simulations, reflective of actual strata performance and participation, incorporate students within courses, course-level variation, and varied sample sizes typical of science equity research (Van Dusen and Nissen 2022). Although we simplified course-level variation, we meticulously modeled the common educational scenario of students nested within courses to understand its influence on outcome predictions (Van Dusen and Nissen 2019).

To understand MAIHDA's efficacy for examining outcomes for intersectional social strata, we asked the following three research questions:

Research Question 1: What are typical strata representation rates and outcomes in physics research-based assessments?

Research Question 2: To what extent do MAIHDA models create more accurate predicted strata outcomes than fixed-effects models?

Research Question 3: To what extent do MAIHDA models allow for the modeling of smaller strata sample sizes?

INTERSECTIONALITY IN STEM HIGHER EDUCATION

Most intersectional research in STEM higher education has used qualitative methods to investigate the double bind that racism and sexism pose to

women of color in STEM fields. These studies reveal the covert norms and expectations of these disciplines, such as the emphasis on mastery, competitiveness, and individualism, that disproportionately exclude women of color (Carter et al. 2019; Carlone and Johnson, 2007; Cochran, Boveda, and Prescod-Weinstein 2020; Dawson, 2019; Ireland et al., 2018; McGee 2023; Ong 2023; Traweek 2009; Womack et al. 2023). Such norms not only further marginalize multiply marginalized groups, but they also allow White faculty to profess inclusivity while perpetuating color-evasive racism, undermining the struggles against racism and sexism faced by women of color (Dancy and Hodari 2023; Fries-Britt et al., 2010; 2013; King, Russo-Tait, and Andrews 2023; Robertson et al. 2023).

The few quantitative studies in STEM higher education that evoke intersectionality use interaction terms to build intersectional models (Van Dusen and Nissen 2022). These studies have found meaningful differences in content knowledge and beliefs before instruction (Nissen, Horses, and Dusen, 2021; Van Dusen and Nissen 2020b; Van Dusen et al. 2021), opportunities and performance in AP physics and chemistry courses (Krakehl and Kelly 2021; Palermo, Kelly, and Krakehl 2022), and inequities in course failure rates (Van Dusen and Nissen 2020a) that represent the educational debt American society owes to Black, Brown, Indigenous, and poor students (Ladson-Billings 2006) and women in STEM. Yet the common use of *p*-value cutoffs to exclude interaction terms and small sample sizes results in many studies not building intersectional models at all (Stewart et al. 2021; Van Dusen and Nissen 2022).

APPROACHES TO MODELING INTERSECTIONALITY

Fixed-Effects Approach

Bauer (2014) notes that despite their mutual enhancement potential, intersectionality theory and quantitative methods have not fully converged. Health researchers have recognized this gap, acknowledging the complexity of inequities and the need for quantitative analysis to address these issues. In education studies, quantitatively applying intersectionality has provided insights by exploring how intersecting identities (race, gender, socioeconomic status) and their power dynamics

affect academic outcomes (Riegle-Crumb and Grodsky 2010), disciplinary measures (Morris and Perry 2017), and financial planning for college (Quadlin and Conwell 2021). Intersectionality research also includes often overlooked factors, such as body size, alongside race and sex on education (Branigan 2017), showcasing the theory's breadth in examining diverse social strata effects.

In their examination of quantitative intersectional research, Mena and Bolte (2019) found that intersections in health studies are often analyzed using regression models that include interaction terms or through stratified analysis, a practice confirmed by Bauer et al. (2021) across various fields. This approach involves integrating interaction terms (e.g., Gender \times Race) into regression equations to investigate the nuanced effects of intersecting identity strata (Evans 2019). For example, Conwell (2021) utilized cubic regression with intersectional terms, such as race and income, to study their effect on children's math scores. Similarly, Riegle-Crumb and Grodsky (2010) used multivariate regression to explore interactions between ethnicity and various socioeconomic factors, demonstrating the methodology's applicability in understanding complex social dynamics.

Researchers have also recognized that these intersectional approaches can address questions involving nested data and data that evaluate intervention treatments. For instance, Morris and Perry (2017) studied the interaction of race and gender on office referrals. They conducted a series of models including an interaction term for race/ethnicity and gender and using a three-level model, with observations (Level 1) nested within students (Level 2) nested within schools (Level 3).

Problems With Fixed-Effects Approaches

Using interaction terms to analyze intersectional social strata, although insightful, faces challenges with scalability and precision as the number of strata increases (Evans et al. 2018). Moreover, these approaches often require large data sets, typically over 1,000 students (Bauer et al. 2021). The necessity for ample sample sizes to adequately represent and disaggregate each strata complicates the examination of intersectionality, especially in studies with diverse programs or treatments. For example, Riegle-Crumb and Grodsky (2010) tackled mathematical achievement disparities by

conducting separate analyses for students in different levels of math courses using intersectional terms. This division dilutes statistical power, amplifies model uncertainty, and complicates the detection of inequities among strata.

MAIHDA

The challenges associated with fixed effects and interaction terms in intersectional studies highlight logistical and conceptual issues. Scott and Siltanen (2017) critiqued the alignment of multiple regression, a prevalent method in intersectional research, with the core principles of intersectionality, emphasizing the importance of contextual sensitivity, open-ended examination of inequities, and acknowledgment of the multifaceted and multilayered nature of inequity. They argued that traditional regression methods fall short of capturing these intricacies, recommending instead a focus on data that contextualize individuals within broader settings. Despite slow adoption, emerging evidence from simulations indicates that MAIHDA (Merlo 2018) could offer a more fitting approach by effectively incorporating these intersectional considerations. MAIHDA uses “[h]ierarchical and multilevel models to study large numbers of interactions and intersectional identities while partitioning the total variance between two levels—the *between-strata* (or between-category) level and the *within-strata* (or within-category) level” (Evans et al. 2018:64).

The idea of using MAIHDA for intersectional analyses originated in Evans’s (2015) dissertation and was further explored by Evans et al. (2018) in the context of health inequities. Their findings indicate MAIHDA’s superiority over traditional fixed-effects methods in handling intersectional terms for several reasons: MAIHDA simplifies the inclusion of additional intersectional terms, improves accuracy for smaller strata, effectively considers the effects on both multiply marginalized and partially privileged individuals, offers more insightful statistics than merely the significance of interaction terms, and facilitates analysis of within-group differences. MAIHDA’s multilevel structure also allows for the incorporation of experimental designs into the analysis, enhancing its capability to dissect complex data. Evans et al. underscored MAIHDA’s efficiency by comparing Bayesian information criterion (BIC) scores between fixed-effects and MAIHDA

models across simulations with strata numbers ranging from 4 to 384; they found MAIHDA models maintained consistent BIC scores despite increasing strata, unlike the exponential increase observed in fixed-effects models.

Independently, Jones, Johnston, and Manley (2016) examined an analogous technique to MAIHDA. They modeled voting rates using fixed effects for primary terms and random effects to account for interaction terms. They found that using random effects improved the level of detail in the model and protected researchers from overinterpreting their data.

Since these initial publications, simulation studies have examined different aspects of MAIHDA’s efficacy and utility. This work evaluates MAIHDA’s performance compared to alternatives in reducing false positives and yielding accurate models with varying sample sizes. Bell et al. (2019), for example, performed a simulation study examining false-positive rates for intersectional terms in fixed-effects and MAIHDA models. They found that MAIHDA models produced fewer false positives than did saturated fixed-effects models.

Lizotte et al. (2020) demonstrated that when given large sample sizes (100,000) and large strata populations (3,125), MAIHDA models and fixed-effects models were both highly accurate. They proposed that Evans (2019) had misinterpreted the MAIHDA model terms. Specifically, they claimed that the fixed effects had been mistaken for the grand means and that the residual terms had been mistaken for the intersectional effects. Evans et al. (2020) rebutted this suggestion. Using a simulation study, they demonstrated that, as with all multilevel models, the fixed effects in MAIHDA models are precision-weighted grand means, which, under some conditions, will be equal to the grand means.

In their comparative simulation studies, Mahendran et al. (2022a, 2022b) evaluated seven modeling techniques, including cross-classification, regression with interactions, MAIHDA, and various decision trees (CART, CTree, CHAID, and random forest), for analyzing intersectional health outcomes. Using logistic regression for binary outcomes, they discovered that although some methods could occasionally reach MAIHDA’s level of accuracy and sensitivity, MAIHDA consistently outperformed all other techniques across different scenarios (Mahendran et al. 2022b). In their follow-up study focused on

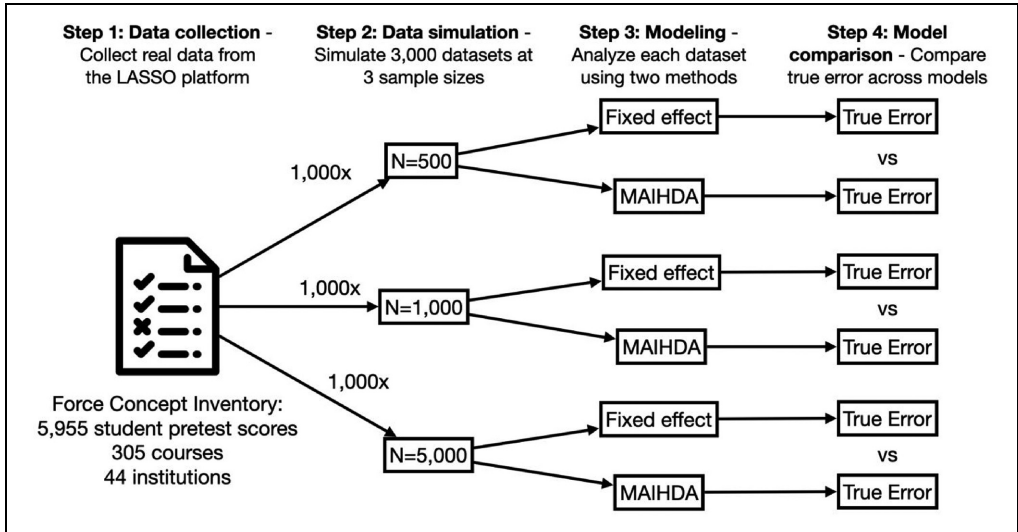


Figure 2. The four steps of our analysis.

continuous outcomes, random forest emerged as the most robust, yet MAIHDA was still favored for its performance across various sample sizes (Mahendran et al. 2022a). Notably, these studies shifted from MAIHDA’s usual Bayesian framework to frequentist models, with researchers arguing that the latter offered comparable results with the advantage of quicker estimation.

Keller et al. (2023) affirmed MAIHDA’s advantages in education research: MAIHDA could efficiently scale to higher dimensional data, maintain model simplicity, and provide precise estimates for strata with few observations, consistent with findings in health research. However, they identified challenges specific to education, such as the difficulty of gathering sufficiently large samples to represent diverse intersectional strata of students. They also highlighted the importance of investigating MAIHDA’s capacity to handle education’s inherent nested data structures, such as students within courses and schools. Evans (2019) conducted a preliminary examination of nesting individuals within schools alongside social strata, but comprehensive studies using cross-classified multilevel MAIHDA models in educational contexts remain unexplored.

In this simulation study, we expand on the literature by creating a realistic model of science student outcomes—in which strata composition varies to mirror likely data sets—and comparing MAIHDA’s accuracy against fixed-effects models

of varying sample sizes. We improved the sophistication of the simulated data from prior studies by building it off of large-scale, real-world education data. We used cross-classified multilevel models that nest students in courses and social strata, a critical feature of large-scale educational studies (DiPrete and Forristal 1994; Niehaus, Campbell, and Inkelas 2014; Raudenbush and Bryk 1986; Van Dusen and Nissen 2019). We also examine how MAIHDA’s improved accuracy can allow researchers to include more strata in their models without sacrificing the accuracy of their predictions.

METHODS

To compare the accuracy and precision of fixed-effects and MAIHDA models across three sample sizes, our analysis proceeds in four steps (see Figure 2). Step 1 uses the FCI data to create a true model of test scores across 20 intersectional identities. Step 2 simulates data for the true model 1,000 times for each of the three sample sizes. Step 3 models the data using both fixed-effects and MAIHDA models. Step 4 compares the true error (predicted – true score) across the fixed-effects and MAIHDA models.

Step 1: Data Collection

To achieve a realistic simulation reflective of actual education research, we utilize simulated

Table 1. Institution Information from Our Data Set.

Total	Type		Size			Highest Degree				Special Designators	
	Public	Private	Small	Medium	Large	AA	BA	MA	PhD	Hispanic-Serving Institution	Minority-Serving Institution
40	25	13	5	15	18	4	5	14	15	11	11

Note: The subcategories do not add up to the total because two institutions were not in the Carnegie Classification of Institutions of Higher Education public 2021 database.

data inspired by real-world data from the Learning About STEM Student Outcomes (LASSO) platform (Van Dusen 2018). LASSO, an online tool for STEM education assessment, offers detailed reports on student performance in core science subjects, providing anonymized student data for research purposes with participant consent. The data set does not capture every institutional context, but it surpasses the representativeness of data commonly used in studies on science higher education (Nissen et al. 2021). The Carnegie Classification of Institutions of Higher Education public 2021 database served as the basis for detailing institution types (see Table 1).

LASSO provided the social identities and pretest scores from 5,955 students in 171 courses at 40 institutions on the FCI. To assess the effectiveness of physics instruction, Hestenes et al. (1992) developed the FCI to probe student understanding of Newtonian forces. Researchers have applied many different quantitative methods to data from the FCI (see e.g., Eaton and Willoughby 2020). One strand of this quantitative research focuses on the fairness of the FCI and its ability to produce unbiased data across different strata. Morley, Nissen, and Van Dusen (2023) found the FCI was measurement invariant across the intersection of 10 racial and gender strata, but evidence indicates that several items on the FCI function differently for men (Traxler et al. 2018), White men in particular (Buncher et al. 2021), than for individuals identifying with other gender or racial strata. Buncher et al. (working paper), however, concluded that the differences in item performance across strata were likely due to differences in construct-relevant background knowledge rather than bias from construct-irrelevant knowledge. Nonetheless, researchers often use the FCI to investigate the effectiveness of instructional techniques (Bruun and Brewe 2013; Caballero et al.

2012; Han et al. 2015; Nissen, Her Many Horses, et al. 2022; Xiao et al. 2020) and equity in courses (Brewe, Kramer, and O'Brien 2009; Good, Maries, and Singh 2019; Van Dusen and Nissen 2020a, 2020b).

Step 2: Data Simulation

To reflect the range of sample sizes commonly encountered in equity research (Van Dusen and Nissen 2022), we generated simulated data sets with 500, 1,000, and 5,000 students, creating 1,000 data sets for each size, totaling 3,000 simulations. In these data sets, each student was placed within a course comprising 50 students; the course average scores had a standard deviation of 10 points, given the FCI score range of 0 to 100. This variability aligns with actual educational data and represents typical variance seen in educational studies (Condon, Lavery, and Engle 2016; Sun and Pan 2014; Van Dusen and Nissen 2019, 2020b). We avoided simulating complex student-classroom demographics, such as having men more represented in higher or lower performing contexts, to maintain the study's focus. Of the initial simulations, 44 data sets at the 500-student sample size were discarded due to having strata with no data, rendering them unsuitable for analysis with either fixed-effects or MAIHDA models. This adjustment led to a final count of 2,956 data sets available for our analysis.

Within each simulated data set, we included strata variables for five racial groups (i.e., Asian = 14 percent, Black = 6 percent, Hispanic = 7 percent, White = 63 percent, and White Hispanic = 10 percent), two gender groups (i.e., women = 36 percent, men = 64 percent), and two college-generation groups (i.e., first-generation [FG] = 36 percent, continuing-generation [CG] = 64 percent), creating 20 intersectional social strata

combinations. Based on the LASSO data, we set each strata's representation rates, true values, and standard deviations. Table 3 shows the true score and proportional representation for each strata and the student-level standard deviations.

The simulation set the proportional representation of each strata in the data set, but how they intersected varied across each simulation. For example, we set 6 percent of students in each simulation as Black, 36 percent as women, and 36 percent as FG. We assigned each axis of social identifiers (i.e., race, gender, and FG/CG designation) independently, meaning we did not ensure that 36 percent of each racial group were women. So, although FG Black women made up 0.76 percent (6 percent \times 36 percent \times 36 percent) of the data on average, in any given simulation, their representation could range from 0 percent (e.g., if none of the Black students were women and FG) to 6 percent (if all the Black students were women and FG). By allowing the proportion of students in each strata to vary across simulations, we can examine the accuracy of predicted outcomes across a broader range of strata sample sizes.

Our simulation also randomly assigned students to courses. For instance, 36 percent of the population were women, but within any given course, the share of women could range from 0 to 100 percent. We set the standard deviation for scores within a strata to be 20 points.

Our student score simulation equation is as follows:

$$\begin{aligned}
 \text{Score}_{ij} = & 47 - 3 \times FG_{ij} - 10 \times \text{women}_{ij} - 6 \\
 & \times \text{Black}_{ij} - 10 \times \text{Hispanic}_{ij} + 2 \\
 & \times \text{White}_{ij} + 3 \times FG_{ij} \times \text{women}_{ij} + 1 \\
 & \times \text{White}_{ij} \times \text{Hispanic}_{ij} - 1 \times FG_{ij} \\
 & \times \text{Black}_{ij} + 2 \times FG_{ij} \times \text{Hispanic}_{ij} + 2 \\
 & \times FG_{ij} \times \text{White}_{ij} - 2 \times \text{women}_{ij} \\
 & \times \text{Black}_{ij} + 1 \times \text{women}_{ij} \times \text{Hispanic}_{ij} - 4 \\
 & \times \text{women}_{ij} \times \text{White}_{ij} - 1 \times FG_{ij} \times \text{White}_{ij} \\
 & \times \text{Hispanic}_{ij} + 5 \times \text{women}_{ij} \times \text{White}_{ij} \\
 & \times \text{Hispanic}_{ij} + 4 \times FG_{ij} \times \text{women}_{ij} \\
 & \times \text{Black}_{ij} + 1 \times FG_{ij} \times \text{women}_{ij} \times \text{White}_{ij} - 1 \\
 & \times FG_{ij} \times \text{women}_{ij} \times \text{White}_{ij} \times \text{Hispanic}_{ij} \\
 & + u_{\text{course}j} + r_i \\
 u_{\text{course}j} \sim & N(0,10) \text{ (Level 2)} \\
 r_i \sim & N(0,20) \text{ (Level 1)}
 \end{aligned}
 \tag{1}$$

Step 3: Modeling

We analyzed each data set using a fixed-effects model and a MAIHDA model. Our fixed-effects intersectional model (see Figure 3) is a frequentist multilevel model that nests students (Level 1) within courses (Level 2). The fixed-effects intersectional model includes an interaction term for each combination of race, gender, and college-generation term:

$$\begin{aligned}
 \text{Score}_{ij} = & \beta_0 + \beta_1 \text{Black}_{ij} + \beta_2 \text{Hispanic}_{ij} \\
 & + \beta_3 \text{White}_{ij} + \beta_4 \text{Hispanic}_{ij} \times \text{White}_{ij} \\
 & + \text{Woman}_{ij} \times (\beta_5 + \beta_6 \text{Black}_{ij} \\
 & + \beta_7 \text{Hispanic}_{ij} + \beta_8 \text{White}_{ij} \\
 & + \beta_9 \text{Hispanic}_{ij} \times \text{White}_{ij}) \\
 & + FG_{ij} \times (\beta_{10} + \beta_{11} \text{Black}_{ij} \\
 & + \beta_{12} \text{Hispanic}_{ij} + \beta_{13} \text{White}_{ij} \\
 & + \beta_{14} \text{Hispanic}_{ij} \times \text{White}_{ij}) \\
 & + FG_{ij} \times \text{Woman}_{ij} \times (\beta_{15} + \beta_{16} \text{Black}_{ij} \\
 & + \beta_{17} \text{Hispanic}_{ij} + \beta_{18} \text{White}_{ij} \\
 & + \beta_{19} \text{Hispanic}_{ij} \times \text{White}_{ij}) \\
 & + u_{\text{course}j} + e_{ij} \\
 u_{\text{course}j} \sim & N(0, \sigma_{u_{\text{course}}}^2) \text{ (Level 2)} \\
 e_{ij} \sim & N(0, \sigma_e^2) \text{ (Level 1)}
 \end{aligned}
 \tag{2}$$

Our MAIHDA model (see Figure 4) is a Bayesian cross-classified multilevel model that nests students (Level 1) within courses (Level 2a) and strata (Level 2b):

$$\begin{aligned}
 \text{Score}_{i(j_1, j_2)} = & \beta_0 + \beta_1 \text{Black}_{(j_1, j_2)} \\
 & + \beta_2 \text{Hispanic}_{(j_1, j_2)} + \beta_3 \text{White}_{(j_1, j_2)} \\
 & + \beta_4 \text{Woman}_{(j_1, j_2)} + \beta_5 FG_{(j_1, j_2)} \\
 & + u_{\text{course}j_1} + u_{\text{strata}j_2} + e_{i(j_1, j_2)} \\
 u_{\text{course}j_1} \sim & N(0, \sigma_{u_{\text{course}}}^2) \text{ (Level 2a)} \\
 u_{\text{strata}j_2} \sim & N(0, \sigma_{u_{\text{strata}}}^2) \text{ (Level 2b)} \\
 e_{i(j_1, j_2)} \sim & N(0, \sigma_e^2) \text{ (Level 1)}
 \end{aligned}
 \tag{3}$$

We follow Fielding and Goldstein's (2006) notation for multilevel models and cross-classified multilevel model equations. For example, the subscripts in $\text{Score}_{i(j_1, j_2)}$ refer to the score for the i th student in the j_1 th course and the j_2 th strata. j_1

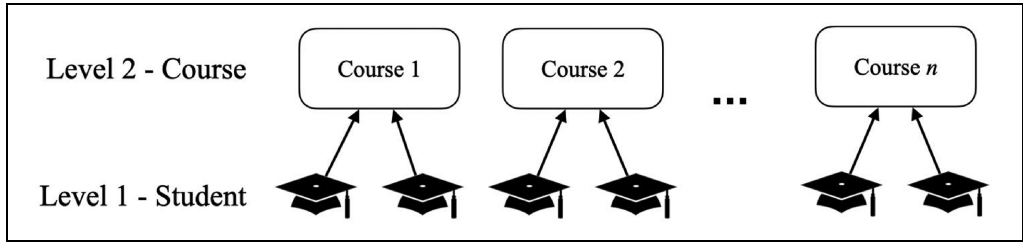


Figure 3. The multilevel structure of our fixed-effects model with students (Level 1) nested within courses (Level 2).

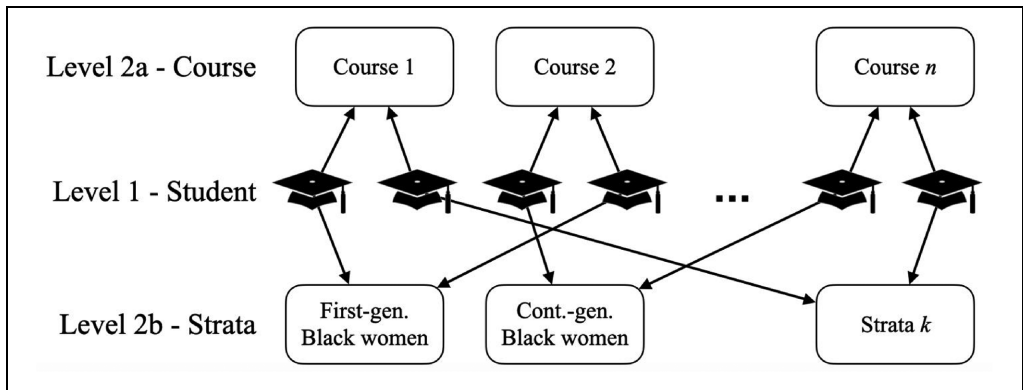


Figure 4. The multilevel structure of our cross-classified MAIHDA model with students (Level 1) nested within courses (Level 2a) and strata (Level 2b).

Note: The strata lie at the intersections of social identity variables included in the model, and the model includes an error term for each strata. MAIHDA = multilevel analysis of individual heterogeneity and discriminatory accuracy.

runs from 1 to n , where n represents the total number of courses. j_2 runs from 1 to k , where k represents the total number of strata. n and k are the number of unique values in the course and strata variables. β_0 represents the score for CG Asian men. e_{ij} represents the student-level error for a specific score. u_{course_j} represents the course-level error for each course. In addition to including student- and course-level error terms ($e_{i(j_1, j_2)}$ and $u_{course_{j_1}}$) in the fixed-effects model, the MAIHDA equation includes a strata error term ($u_{strata_{j_2}}$). None of the variables are centered.

To understand the structure of the variance of our data, we ran an unconditional cross-classified model with no fixed effects. We use this model to calculate the variance partition coefficient (VPC) and the proportional change in variance (PCV). The VPC measures the variance attributed to a level in a multilevel model and is sometimes

called the intraclass correlation coefficient. In most MAIHDA models, the VPC score for strata usually less than 10 percent (Evans et al. 2020). The PCV measures the total between-stratum variation explained by including additive main effects. Our unconditional cross-classified multilevel model is written as follows:

$$\begin{aligned}
 \text{Score}_{i(j_1, j_2)} &= \beta_0 + u_{course_{j_1}} + u_{strata_{j_2}} \\
 &\quad + e_{i(j_1, j_2)} \\
 u_{course_{j_1}} &\sim N(0, \sigma_{u_{course}}^2) \text{ (Level 2a)} \\
 u_{strata_{j_2}} &\sim N(0, \sigma_{u_{strata}}^2) \text{ (Level 2b)} \\
 e_{i(j_1, j_2)} &\sim N(0, \sigma_e^2) \text{ (Level 1)}
 \end{aligned}
 \tag{4}$$

Comparing the variances from the real-world data with our simulated data shows that the share of the variance accounted for by strata (VPC) is

Table 2. The Variance, Variance Partition Coefficient, and Proportional Change in Variance for the Real-World and Simulated Data Sets.

	σ_e^2	$\sigma_{u_course}^2$	$\sigma_{u_strata}^2$	VPC_{course}	VPC_{strata}	PCV_{strata}
Real-world data						
Unconditional CCMM model	423.4	17.4	53.8	3.6%	11.2%	—
MAIHDA model	415.8	17.7	9.2	4.1%	2.5%	78.0%
Simulated data						
Unconditional CCMM model	399.9	109.7	45.7	20%	9%	—
MAIHDA model	399.6	109.6	6.5	21%	2%	79.6%

Note: σ_e^2 = student-level variance.; $\sigma_{u_course}^2$ = course-level variance.; $\sigma_{u_strata}^2$ = strata-level variance; VPC_{course} = course-level variance partition coefficient, converted to a percentage.; VPC_{strata} = strata-level variance partition coefficient, converted to a percentage.; PCV_{strata} = strata-level proportional change in variance, converted to a percentage. The proportional change in variance represents the percentage of the total between-strata variation that is explained by inclusion of primary terms. CCMM = Cross-classified multilevel model; MAIHDA = multilevel analysis of individual heterogeneity and discriminatory accuracy.

very similar in both data sets (see Table 2). Including the fixed-effects terms in the MAIHDA model also accounts for similar shares of the variance due to strata (PCV) for the real-world and simulated data sets. The one place where the two data sets diverge is the share of the variance accounted for by the courses. We reran our simulation with a smaller standard deviation in course scores (4.5 points). This made the VPC_{course} values match the real-world data, but it did not meaningfully affect any of the group predicted outcomes. This shows that our results are robust across a range of variance at the course level. Note, however, that our course-level variance does not have any structure (i.e., there is no correlation between course and strata-level variances).

Our unconditional model shows the VPC_{strata} is 9 percent, within the typical range reported (Evans et al. 2020). The PCV_{strata} in the MAIHDA model is 79.6 percent, which means the primary terms in our MAIHDA model account for 79.6 percent of the variance in the strata level. It also means that 20.4 percent of the variance between strata can be attributed to intersectional effects.

Society’s educational debts owed to marginalized groups were calculated by subtracting the mean score for a strata from those of CG White men.

Step 4: Model Comparison

To compare the precision and accuracy of MAIHDA and fixed-effects models and explore

minimum sample sizes for robust estimation, we calculated the true error in the prediction for each of the 20 strata. The true error is the difference between the predicted and true scores (estimated score – true score). We compared their mean absolute true errors to determine which modeling technique yields more accurate predictions. To determine which modeling technique would allow for the modeling of smaller strata sample sizes, we examined the mean absolute true error for strata with sample sizes of 20 or fewer.

FINDINGS

Research Question 1: What Are Typical Strata Representation Rates and Outcomes on Physics Research-Based Assessments?

LASSO provided data from 5,955 students in 171 courses at 40 institutions on the FCI. These data include 20 intersectional social strata (see Table 3). The share of the sample for a strata ranges from 26 percent (CG White men) to 1 percent (CG Black women, FG Black men, FG Black women, FG Hispanic women, and FG White Hispanic women). The mean score on the assessment for a strata ranges from 49 points (CG White men) to 28 points (CG Hispanic women). Society’s educational debts range from 21 points (CG Hispanic women) to 2 points (CG Asian men).

Table 3. The Share of the Total Sample, True Score, and Student-Level Standard Deviation for Each Intersectional Social Strata as Set by the Simulation.

Intersectional Social Strata	Share of Sample	Score (points)	Society's Educational Debt (points)
CG Asian men	6%	47	2
CG Asian women	3%	37	12
CG Black men	2%	41	8
CG Black women	1%	29	20
CG Hispanic men	3%	37	12
CG Hispanic women	2%	28	21
CG White Hispanic men	4%	40	9
CG White Hispanic women	2%	32	17
CG White men	26%	49	—
CG White women	15%	35	14
FG Asian men	3%	44	5
FG Asian women	2%	37	12
FG Black men	1%	37	12
FG Black women	1%	32	17
FG Hispanic men	2%	36	13
FG Hispanic women	1%	30	19
FG White Hispanic men	2%	40	9
FG White Hispanic women	1%	35	14
FG White men	15%	48	1
FG White women	8%	38	11

Note: CG = continuing-generation; FG = first-generation.

Research Question 2: To What Extent Do MAIHDA Models Create More Accurate Predicted Strata Outcomes Than Fixed-Effects Models?

We compare the true errors of the two modeling methods to determine their accuracy. Table 4 shows the mean absolute true error disaggregated by model, strata, and total sample size. The MAIHDA model has smaller mean absolute true errors in 56 of the 60 cases (i.e., 20 strata in each of the three total sample sizes). In the four exceptions in which the MAIHDA model is less accurate on average, the mean absolute true errors are only marginally larger, ranging from 0.0 to 0.4 points. As the total sample sizes increase, the amount by which MAIHDA outperforms the fixed-effects model decreases. MAIHDA's mean absolute true error is 31 percent (1.7 points) smaller than the fixed-effects model when the total sample size is 500 but only 13 percent (0.2 points) smaller when the sample size is 5,000.

The largest mean absolute true error for the fixed-effects model is for FG Black women (9.3

points), who also have the smallest mean N (3.7). In the MAIHDA model, the mean absolute true error is reduced to 4.9 points. Given that the standard deviation for student scores is 20 points, the shift from fixed-effects to MAIHDA models decreases the mean absolute true error from 0.47 SD to 0.25 SD.

To understand how strata sample sizes affect the accuracy of the models, we examine the true error for each strata prediction for each model. Figure 5 shows the true error for all the predictions versus each model's strata sample size. Both models are similarly accurate when they have larger total and strata sample sizes. The magnitude of the true errors increases for both models with smaller total sample sizes and strata sample sizes. Where the models differ, however, is that the magnitude of the true error increases more quickly in the fixed-effects model as the strata sample sizes decrease. The improvement in performance for MAIHDA over fixed-effects models is most apparent for the strata with the smallest sample size (i.e., FG Black women). The reduction in the mean absolute true error for FG Black women when using MAIHDA ranges from 47 percent

Table 4. Each Strata's Mean Absolute True Error across the Three Total Sample Sizes.

Intersectional Social Strata	True Score	Total Sample Size = 500					Total Sample Size = 1,000					Total Sample Size = 5,000																
		Mean Absolute True Error		Fixed Effects		Delta	Mean Absolute True Error		Fixed Effects		Delta	Mean Absolute True Error		Fixed Effects		Delta												
		Mean N	MAIHDA	Mean N	MAIHDA		Mean N	MAIHDA	Mean N	MAIHDA		Mean N	MAIHDA	Mean N	MAIHDA													
CG Asian men	47	28.3	4.0	3.4	-0.6	56.8	2.8	2.5	-0.3	285.9	1.2	1.2	0.0	CG Asian women	37	15.9	4.7	3.6	-1.1	32.3	3.5	2.6	-0.9	160.8	1.5	1.3	-0.2	
CG Black men	41	11.8	5.4	4.2	-1.2	23.6	3.8	3.0	-0.8	117.4	1.7	1.6	-0.1	CG Black women	29	6.7	6.8	4.3	-2.5	13.1	4.8	3.0	-1.8	65.7	2.2	1.6	-0.6	
CG Hispanic men	37	15.1	5.0	3.7	-1.3	30.2	3.6	2.7	-0.9	151.3	1.5	1.3	-0.2	CG Hispanic women	28	8.5	6.2	3.5	-2.7	17.1	4.7	2.8	-1.9	84.5	1.9	1.4	-0.5	
CG White Hispanic men	40	20.2	4.7	3.5	-1.2	40.5	3.1	2.5	-0.6	200.9	1.4	1.2	-0.2	CG White Hispanic women	32	11.4	5.5	3.9	-1.6	22.8	3.9	2.8	-1.1	113.5	1.7	1.4	-0.3	
CG White men	49	129.0	2.9	2.9	0.0	258.2	2.1	2.1	0.0	1292.5	0.9	0.9	0.0	CG White women	35	72.9	3.1	3.1	0.0	145.1	2.2	2.4	0.2	726.9	1.0	1.1	0.1	
FG Asian men	44	16.2	4.8	3.7	-1.1	31.9	3.4	2.7	-0.7	160.8	1.6	1.4	-0.2	FG Asian women	37	9.1	5.7	3.7	-2.0	18.3	4.3	2.8	-1.5	90.4	2.0	1.5	-0.5	
FG Black men	37	6.7	7.3	4.6	-2.7	13.3	4.7	3.5	-1.2	66.1	2.1	2.0	-0.1	FG Black women	32	3.7	9.3	4.9	-4.4	7.4	6.6	3.4	-3.2	37.1	2.8	2.1	-0.7	
FG Hispanic men	36	8.5	6.1	3.9	-2.2	16.8	4.5	3.3	-1.2	85.1	1.9	1.7	-0.2	FG Hispanic women	30	4.8	8.7	4.0	-4.7	9.6	5.6	2.9	-2.7	47.6	2.4	1.7	-0.7	
FG White Hispanic men	40	11.4	5.7	3.6	-2.1	22.7	3.8	2.6	-1.2	113.2	1.7	1.3	-0.4	FG White Hispanic women	35	6.4	7.3	5.2	-2.1	12.7	4.9	4.1	-0.8	63.8	2.3	2.7	0.4	
FG White men	48	72.5	3.2	3.0	-0.2	146.0	2.3	2.2	-0.1	726.5	1.0	1.0	0.0	FG White women	38	40.9	3.5	3.1	-0.4	81.8	2.6	2.3	-0.3	409.9	1.2	1.1	-0.1	
Total		25.0	5.5	3.8	-1.7	50.0	3.9	2.8	-1.1	250.0	1.7	1.5	-0.2															

Note: The delta mean absolute true error is the MAIHDA mean absolute true error minus the fixed-effects mean absolute true error. Negative delta values indicate the MAIHDA model generally produces less error. CG = continuing-generation; FG = first-generation; MAIHDA = multilevel analysis of individual heterogeneity and discriminatory accuracy.

Table 5. Mean and Absolute Mean of the True Error Values for Each Strata Sample Size Disaggregated by Model Type and Total Sample Size.

Strata Sample Size	Total Sample Size = 500				Total Sample Size = 1,000			
	Traditional		MAIHDA		Traditional		MAIHDA	
	Absolute Mean True Error	Mean True Error	Absolute Mean True Error	Mean True Error	Absolute Mean True Error	Mean True Error	Absolute Mean True Error	Mean True Error
1	15.68	-0.68	5.29	-2.93	13.17	2.73	5.56	-3.58
2	11.16	0.59	4.70	-1.77	11.26	2.69	2.31	-0.53
3	9.54	-0.41	4.82	-1.35	8.22	-1.63	3.09	-1.42
4	8.59	-0.16	4.62	-1.08	8.69	0.28	3.59	-2.34
5	7.55	-0.05	4.41	-0.73	8.27	1.50	3.75	-1.34
6	7.03	-0.43	4.24	-0.70	6.16	-0.26	3.40	-1.88
7	6.24	-0.16	4.07	-0.37	6.27	-0.29	3.50	-1.70
8	6.16	-0.25	4.16	-0.43	5.98	-0.01	3.28	-1.03
9	5.87	-0.14	4.01	-0.26	5.43	0.11	3.36	-1.06
10	5.46	-0.05	3.84	-0.18	5.13	-0.10	3.33	-0.49
11	5.51	-0.35	3.93	-0.51	5.33	0.67	3.56	-0.23
12	5.34	0.02	3.84	-0.03	5.07	0.04	3.44	-0.18
13	4.75	0.24	3.60	0.17	4.81	-0.21	3.33	-0.34
14	4.82	-0.25	3.63	-0.16	4.58	-0.05	3.19	0.02
15	4.73	-0.38	3.62	-0.14	4.56	-0.34	3.08	-0.06
16	4.85	-0.21	3.66	-0.06	4.36	-0.09	2.99	-0.08
17	4.94	-0.25	3.73	-0.06	4.21	0.16	3.02	0.42
18	4.76	0.06	3.66	0.14	4.10	0.67	2.88	0.49
19	5.04	0.48	3.93	0.59	4.21	0.47	2.91	-0.09
20	4.39	-0.19	3.40	-0.01	4.14	0.31	2.80	0.15

Note: MAIHDA = multilevel analysis of individual heterogeneity and discriminatory accuracy.

with a total sample size of 500 to 25 percent with a total sample size of 5,000.

Research Question 3: To What Extent Do MAIHDA Models Allow for Modeling of Smaller Strata Sample Sizes?

To understand the differences in each model’s ability to handle smaller strata sample sizes, we examine the absolute mean and mean of the true error for each strata sample size (see Table 5). Figure 6 shows the mean absolute true error values for strata sample sizes ranging from 1 to 20 across total sample sizes and modeling techniques. We limit the figure to this range, as Simmons, Nelson, and Simonsohn (2016) argue, using 20 as a minimum strata sample size. Figure 6 includes a dashed line as a comparison tool indicating the mean

absolute true error for the fixed-effects model when the strata sample size is 20 and the total sample size is 500 (4.4). For the MAIHDA model, the mean absolute true error does not reach 4.4 until the strata sample sizes are reduced to 5 or fewer.

For models with very small strata sample sizes, shrinkage will decrease the size of the strata’s variance term. This smaller variance term leads to the strata’s predicted outcomes being closer to the prediction produced by only adding the model’s primary coefficients for the strata. In our MAIHDA models, when strata sample sizes fall below 10 individuals, we begin to have a small but consistent negative bias in the mean true error scores (see Table 5). The variance causes this negative bias for the least represented strata (FG Black women) to be positive, on average. In other words, the additive effects of the coefficients for the intercept, FG, Black, and women predict scores lower than the true score for FG Black women. Having

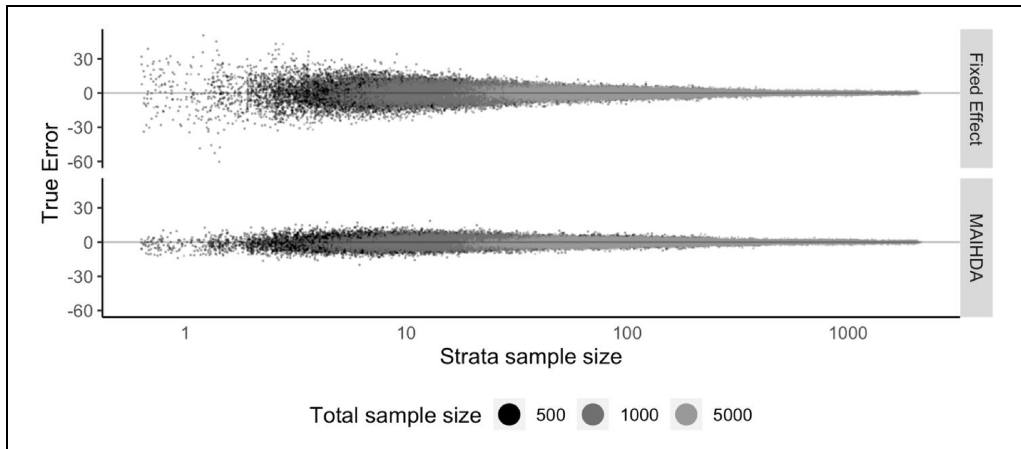


Figure 5. Logarithmic scatter plot of the true error for each predicted strata outcome versus the strata sample size for each model.

Note: The shade of the dots represents the total sample sizes. Predicted values were jittered to enhance visibility.

fewer data for that strata regresses their predicted score closer to the additive score from the model and creates a negative bias.

DISCUSSION

A central goal of intersectional modeling is to disaggregate intersectional social strata outcomes as much as possible to represent the diversity of lived experiences. MAIHDA was developed to support further disaggregation of data across strata than is possible with fixed-effects regression models. To identify whether MAIHDA is superior to fixed-effects models at predicting outcomes for small strata, we compared the accuracy of both modeling techniques on simulated data sets that mirror real-world physics student outcomes.

Our findings show meaningful inequities in physics student scores on the FCI and that MAIHDA creates more accurate models of these inequities than do fixed-effects models (Table 4 and Figure 5). When using the MAIHDA model rather than the fixed-effects model, we see the most significant improvements in accuracy with predictions for small total sample sizes and strata with smaller sample sizes. These improvements support the goal of using MAIHDA to create models with smaller strata, which allows for more strata to reflect the scope of intersectional experiences within a population.

Future Research

Developing and running the MAIHDA and the fixed-effects models took similar effort with our simulated data set. However, one feature of our simulated data that differed from many educational data sets is that it lacked any missing data. In these situations, multiple imputation is often recommended to maximize statistical power and limit the introduction of bias (Nissen, Donatello, and Van Dusen 2019; Rubin 1996). Examining multiple imputed data sets can complicate analyses. Future research should compare the practicality and efficacy of using MAIHDA and fixed-effect models with a real-world data set that includes missing values.

Although we did not examine it in this study, the Bayesian nature of MAIHDA models means their accuracy could be improved through informed priors. Informed priors allow researchers to directly include findings from prior research into their models, similar to what is done in a meta-analysis. Although an investigation may not have the statistical power to include a particular strata in a model, if it can draw on findings across other investigations, it can create reasonably accurate predictions for even smaller strata. Determining the effect of using informed priors with MAIHDA models will require further research.

This analysis did not examine how to account for interaction effects between strata and other

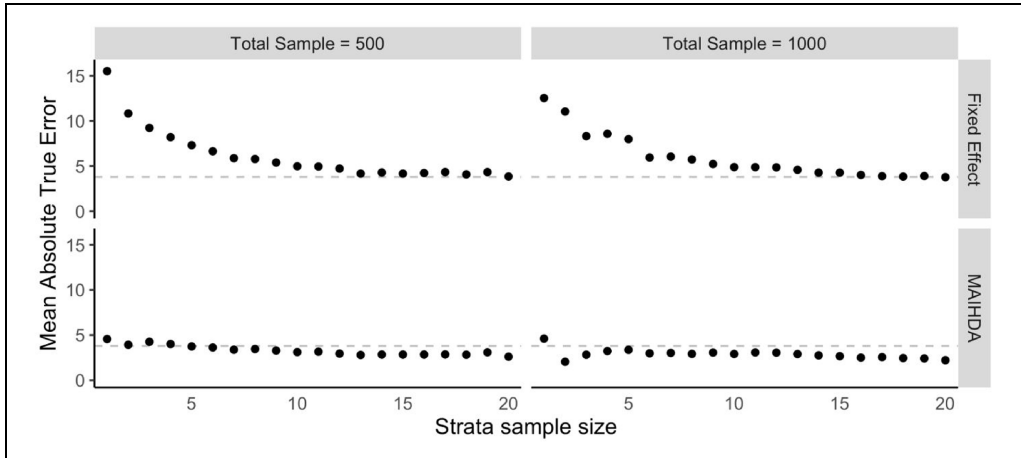


Figure 6. Plot of the mean absolute true error values by strata sample size across total sample size and modeling techniques.

factors. For example, when using a fixed-effects model to determine the differential effect of a classroom intervention across strata, researchers can add fixed interaction terms between the intervention and intersectional social strata variables. In a MAIHDA model, however, the interaction effect is between a fixed variable (e.g., the intervention) and a random variable (e.g., intersectional social strata). It must be included as a -random-effect term. Modeling programs can run such models, but determining the effect of moving these interaction terms from fixed to random will require future research. Furthermore, this study examined continuous outcomes. The efficacy we found likely exists for logistic models of categorical outcomes, but a simulation study is warranted to test it empirically.

This study provides proof that MAIHDA can work with cross-classified multilevel models, but the student–course structures used in our model are reasonably simple. Further research is needed to examine the efficacy of MAIHDA in modeling student outcomes with more complex student, course, and institution structures (e.g., creating subsets of courses that match the makeup of high-, medium-, and low-selectivity institutions). Future simulation studies could also examine the effect of creating cross-classified multilevel models with levels for course and strata versus simply including the strata level.

CONCLUSIONS

The accuracy improvements observed by using MAIHDA over fixed-effects models are strong evidence that MAIHDA can improve the quality of predicted outcomes across intersectional social strata. These results also suggest that researchers may be able to disaggregate their models further, increasing the number of distinct strata represented by accommodating smaller strata sizes. Our examination of the mean absolute true error for small strata sample sizes bore out this idea (Table 5 and Figure 6). One recommended value for minimum strata sample sizes with fixed-effects models is 20 (Simmons et al. 2016). For our fixed-effects models, the mean absolute true error was very similar for total sample sizes of 500 and 1,000 when the strata sample sizes were 20 or fewer. The MAIHDA model maintained an accuracy equal to or better than those observed in the fixed-effects model until the sample sizes were 5 or fewer. These smaller standard error values indicate that researchers could model strata sample sizes below 20 without meaningful compromises in accuracy when using MAIHDA. We refrain from making specific minimum strata sample size recommendations because the context of each research project will determine how much potential error a researcher is willing to entertain to increase the number of strata a model includes.

Beyond being a more effective methodological tool, MAIHDA can change how intersectionality is theorized and reified in quantitative research. Introducing tools to a system can transform cognitive tasks (Hutchins 1995a, 1995b) and lead to novel outcomes (Alvesson and Kärreman, 2007; Roth and Lee 2007). By allowing models to include strata with a sample size of 1, MAIHDA can include all participants as they self-identify without aggregating groups or engaging in data erasure. Opening up this methodological possibility will require researchers to reenvision intersectional theory and its place in quantitative research.

Our findings provide strong evidence of the superiority of using MAIHDA when modeling intersectional outcomes, but MAIHDA is appropriate for only some scenarios. Specifically, MAIHDA requires a sufficient number of strata to offer improved predictions. Evans et al. (2018) recommend having at least 20 strata in a model before using them as a level to nest students to ensure the model has sufficient Level 2 random effects. Researchers have begun to explore MAIHDA models with fewer strata (e.g., Evans 2019; Silva and Evans 2020), but the effect of having fewer strata in a model is currently unknown. Some education equity research uses models that include 20 or more strata (e.g., Jang 2018; Nissen, Van Dusen, and Kukday, 2022; Van Dusen, Nissen, and Johnson 2024), but much of this work does not (Van Dusen and Nissen 2022). With the increased use of novel modeling methods, large-scale data sets (e.g., LASSO), and intersectional perspectives in quantitative research (Wofford and Winkler 2022), the trend may be shifting to models that offer more nuanced depictions of marginalized students' outcomes through the inclusion of more strata.

Although researchers are investigating the implications of using MAIHDA to create intersectional models, the statistical technique of blending fixed and random effects to create predicted outcomes is applicable across many areas of quantitative research. Just as Jones et al. (2016) used an analogous technique to MAIHDA to investigate voting rates, judiciously replacing some fixed effects with random effects can improve model accuracy. When using MAIHDA, researchers have primarily examined replacing interaction terms with random effects (Keller et al. 2023), but the most efficacious replacement of fixed effects will be contextually dependent. Additional research is needed to identify the best method for

selecting fixed-effect terms in MAIHDA or analogous nonintersectional models.

Equity-minded researchers are cautious about treating strata monolithically. Intersectional research that considers variables such as those introduced here somewhat addresses this concern, but more work is needed to explore the potential for disaggregation of data according to more specific racial designations, such as regional Hispanic populations. In this work, we illustrate the potential of MAIHDA using demographic data likely to exist in many surveys that ask standard demographic questions to report on broad categories related to race, gender, and socioeconomic identification (or its proxies, such as college generation). More work is needed to illustrate the potential of MAIHDA—and quantitative intersectionality work generally—using more precise demographic data, such as immigration generation, family country of origin, or country region.

ACKNOWLEDGMENTS

We want to thank the leaders of the summer institute in advanced research methods for Science, Technology, Engineering, and Mathematics education research (Guanglei Hong, Kenneth Frank, Stephen Raudenbush, and Yanyan Sheng) for their support and feedback in the generation of this article.


FUNDING

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was funded, in part, by National Science Foundation Award 1928596.

RESEARCH ETHICS STATEMENT

Because this research does not constitute human subjects research, institutional review board approval was not necessary to carry out this work.

ORCID iD

Ben Van Dusen  <https://orcid.org/0000-0003-1264-0550>

REFERENCES

- Alvesson, Mats, and Dan Kärreman. 2007. "Constructing Mystery: Empirical Matters in Theory Development." *Academy of Management Review* 32(4): 1265–81.

- Bauer, Greta R. 2014. "Incorporating Intersectionality Theory into Population Health Research Methodology: Challenges and the Potential to Advance Health Equity." *Social Science and Medicine* 110:10–17.
- Bauer, Greta R., Siobhan M. Churchill, Mayuri Mahendran, Chantel Walwyn, Daniel Lizotte, and Alma A. Villa-Rueda. 2021. "Intersectionality in Quantitative Research: A Systematic Review of Its Emergence and Applications of Theory and Methods." *SSM-Population Health* 14:100798. doi:10.1016/j.ssmph.2021.100798.
- Bell, Andrew, Daniel Holman, and Kelvyn Jones. 2019. "Using Shrinkage in Multilevel Models to Understand Intersectionality." *Methodology* 15(2):88–96. <https://doi.org/10.1027/1614-2241/a000167>
- Bowleg, Lisa. 2008. "Lesbian + Woman ≠ Black Lesbian Woman: The Methodological Challenges of Qualitative and Quantitative Intersectionality Research." *Sex Roles* 59:312–25.
- Branigan, Amelia R. 2017. "(How) Does Obesity Harm Academic Performance? Stratification at the Intersection of Race, Sex, and Body Size in Elementary and High School." *Sociology of Education* 90(1):25–46.
- Brewer, Eric, Laird Kramer, and George O'Brien. 2009. "Modeling Instruction: Positive Attitudinal Shifts in Introductory Physics Measured with CLASS." *Physical Review Special Topics-Physics Education Research* 5(1):013102. doi:10.1103/PhysRevSTPER.5.013102.
- Brewer, Eric, and Vashti Sawtelle. 2016. "Focused Collection: Gender in Physics." *Physical Review Physics Education Research* 12(2):020001. doi:10.1103/PhysRevPhysEducRes.12.020001.
- Bruun, Jesper, and Eric Brewer. 2013. "Talking and Learning Physics: Predicting Future Grades from Network Measures and Force Concept Inventory Pretest Scores." *Physical Review Special Topics-Physics Education Research* 9(2):020109. doi:10.1103/PhysRevSTPER.9.020109.
- Buncher, John B., Jayson M. Nissen, Ben Van Dusen, and Robert M. Talbot. Working paper. "Is the FCI Biased across the Intersections of Gender and Race?" *Physical Review Physics Education Research*. https://www.researchgate.net/publication/378491611_Is_the_FCI_biased_across_the_intersections_of_gender_and_race
- Buncher, John B., Jayson M. Nissen, Ben Van Dusen, Robert M. Talbot, and Hannah Huvad. 2021. "Bias on the Force Concept Inventory across the Intersection of Gender and Race." Pp. 69–74 in 2021 Physics Education Research Conference Proceedings, edited by M. Bennett, B. Frank, and R. Vieyra. doi:10.1119/perc.2021.pr.Bunche.
- Caballero, Marcos D., Edwin F. Greco, Eric R. Murray, Keith R. Bujak, M. Jackson Marr, Richard Catrambone, Matthew A. Kohlmyer, and Michael F. Schatz. 2012. "Comparing Large Lecture Mechanics Curricula Using the Force Concept Inventory: A Five Thousand Student Study." *American Journal of Physics* 80(7):638–44.
- Carbado, Devon W., Kimberlé Williams Crenshaw, Vickie M. Mays, and Barbara Tomlinson. 2013. "INTERSECTIONALITY: Mapping the Movements of a Theory." *Du Bois Review: Social Science Research on Race* 10(2):303–12.
- Carlone, Heidi B., and Angela Johnson. 2007. "Understanding the Science Experiences of Successful Women of Color: Science Identity as an Analytic Lens." *Journal of Research in Science Teaching* 44(8):1187–218.
- Carter, Deborah Faye, Juanita E. Razo Dueñas, and Rocío Mendoza. 2019. "Critical Examination of the Role of STEM in Propagating and Maintaining Race and Gender Disparities." Pp. 39–97 in *Higher Education: Handbook of Theory and Research*. Vol. 34, edited by M. B. Paulsen, and L. W. Perna. Cham, Switzerland: Springer.
- Cho, Sumi, Kimberlé W. Crenshaw, and Leslie McCall. 2013. "Toward a Field of Intersectionality Studies: Theory, Applications, and Praxis." *Signs* 38(4):785–810.
- Choo, Hae Yeon, and Myra Marx Ferree. 2010. "Practicing Intersectionality in Sociological Research: A Critical Analysis of Inclusions, Interactions, and Institutions in the Study of Inequalities." *Sociological Theory* 28(2):129–49.
- Cochran, Geraldine L., Mildred Boveda, and Chanda Prescod-Weinstein. 2020. "Intersectionality in STEM Education Research." Pp. 257–66 in *Handbook of Research on STEM Education*, edited by C. C. Johnson, M. J. Mohr-Schroeder, T. J. Moore, and L. D. English. Routledge.
- Collins, Patricia Hill. 2015. "Intersectionality's Definitional Dilemmas." *Annual Review of Sociology* 41(1):1–20.
- Collins, Patricia Hill. 2019. *Intersectionality as Critical Social Theory*. Durham, NC: Duke University Press.
- Collins, Patricia Hill, and Sirma Bilge. 2020. *Intersectionality*. New York, NY: John Wiley and Sons.
- Condon, Meghan, Lesley Lavery, and Par Jason Engle. 2016. "Measuring Social Capital: Accounting for Nested Data and Subnetworks within Schools." *Social Indicators Research* 126:1189–207.
- Conwell, Jordan A. 2021. "Diverging Disparities: Race, Parental Income, and Children's Math Scores, 1960 to 2009." *Sociology of Education* 94(1):124–42.
- Crenshaw, Kimberlé W. 1989. "Demarginalizing the Intersection of Race and Sex: A Black Feminist Critique of Antidiscrimination Doctrine, Feminist Theory, and Antiracist Politics." *University of Chicago Legal Forum* 140:139–67.
- Crenshaw, Kimberlé W. 1990. "Mapping the Margins: Intersectionality, Identity Politics, and Violence against Women of Color." *Stanford Law Review* 43(6):1241–99.
- Dancy, Melissa, and Apriel K. Hodari. 2023. "How Well-Intentioned White Male Physicists Maintain

- Ignorance of Inequity and Justify Inaction.” *International Journal of STEM Education* 10(1):1–29.
- Dawson, Emily. 2019. *Equity, Exclusion and Everyday Science Learning: The Experiences of Minoritised Groups*. New York, NY: Routledge.
- DiPrete, Thomas A., and Jerry D. Forristal. 1994. “Multilevel Models: Methods and Substance.” *Annual Review of Sociology* 20(1):331–57.
- Eaton, Philip, and Shannon Willoughby. 2020. “Identifying a Preinstruction to Postinstruction Factor Model for the Force Concept Inventory within a Multitrait Item Response Theory Framework.” *Physical Review Physics Education Research* 16(1):010106. doi:10.1103/PhysRevPhysEducRes.16.010106.
- Evans, Clare Rosenfeld. 2015. *Innovative Approaches to Investigating Social Determinants of Health-Social Networks, Environmental Effects and Intersectionality*. Cambridge, MA: Harvard University.
- Evans, Clare Rosenfeld. 2019. “Adding Interactions to Models of Intersectional Health Inequalities: Comparing Multilevel and Conventional Methods.” *Social Science and Medicine* 221:95–105.
- Evans, Clare R., George Leckie, and Juan Merlo. 2020. “Multilevel Versus Single-Level Regression for the Analysis of Multilevel Information: The Case of Quantitative Intersectional Analysis.” *Social Science and Medicine* 245:112499. doi:10.1016/j.socscimed.2019.112499.
- Evans, Clare R., David R. Williams, Jukka-Pekka Onnela, and S. V. Subramanian. 2018. “A Multilevel Approach to Modeling Health Inequalities at the Intersection of Multiple Social Identities.” *Social Science and Medicine* 203:64–73.
- Fielding, Antony, and Harvey Goldstein. 2006. *Cross-Classified and Multiple Membership Structures in Multilevel Models: An Introduction and Review*. London: Department for Education and Skills.
- Fries-Britt, Sharon, Jennifer Johnson, and Brian Burt. 2013. “Black Students in Physics: The Intersection of Academic Ability, Race, Gender, and Class, in Research on African American Education.” Pp. 21–39 in *Living at the Intersections: Social Identities and Black Collegians*, edited by T. L. Strayhorn. Charlotte, NC: Information Age Publishing.
- Fries-Britt, Sharon L., Toyia K. Younger, and Wendell D. Hall. 2010. “Lessons from High-Achieving Students of Color in Physics.” *New Directions for Institutional Research* 148:75–83.
- Good, Melanie, Alexandru Maries, and Chandralekha Singh. 2019. “Impact of Traditional or Evidence-Based Active-Engagement Instruction on Introductory Female and Male Students’ Attitudes and Approaches to Physics Problem Solving.” *Physical Review Physics Education Research* 15(2):020129. doi:10.1103/PhysRevPhysEducRes.15.020129.
- Han, Jing, Lei Bao, Li Chen, Tianfang Cai, Yuan Pi, Shaona Zhou, Yan Tu, and Kathleen Koenig. 2015. “Dividing the Force Concept Inventory into Two Equivalent Half-Length Tests.” *Physical Review Special Topics-Physics Education Research*, 11(1), 010112. doi:10.1103/PhysRevSTPER.11.010112.
- Harris, Jessica C., and Lori D. Patton. 2019. “Un/doing Intersectionality through Higher Education Research.” *The Journal of Higher Education* 90(3):347–72.
- Hestenes, David, Malcolm Wells, and Gregg Swackhamer. 1992. “Force Concept Inventory.” *The Physics Teacher* 30(3):141–58.
- Edwin Hutchins. 1995a. *Cognition in the Wild*. Cambridge, MA: MIT Press.
- Hutchins, Edwin. 1995b. “How a Cockpit Remembers Its Speeds.” *Cognitive Science* 19(3):265–88.
- Ireland, Danyelle T., Kimberley Edelin Freeman, Cynthia E. Winston-Proctor, Kendra D. DeLaine, Stacey McDonald Lowe, and Kamilah M. Woodson. 2018. “(Un)hidden Figures: A Synthesis of Research Examining the Intersectional Experiences of Black Women and Girls in STEM Education.” *Review of Research in Education* 42(1):226–54.
- Jang, Sung Tae. 2018. “The Implications of Intersectionality on Southeast Asian Female Students’ Educational Outcomes in the United States: A Critical Quantitative Intersectionality Analysis.” *American Educational Research Journal* 55(6):1268–306.
- Jones, Kelvyn, Ron Johnston, and David Manley. 2016. “Uncovering Interactions in Multivariate Contingency Tables: A Multi-level Modeling Exploratory Approach.” *Methodological Innovations* 9. doi:10.1177/2059799116672874.
- Keller, Lena, Oliver Luedtke, Franzis Preckel, and Martin Brunner. 2023. “Educational Inequalities at the Intersection of Multiple Social Categories: An Introduction and Systematic Review of the Multilevel Analysis of Individual Heterogeneity and Discriminatory Accuracy (MAIHDA) Approach.” *Educational Psychology Review* 35:31. doi:10.1007/s10648-023-09733-5.
- King, Gretchen P., Tatiane Russo-Tait, and Tessa C. Andrews. 2023. “Evolving Race: STEM Faculty Struggle to Acknowledge Racialized Classroom Events.” *CBE—Life Sciences Education* 22(1): ar14. doi:10.1187/cbe.22-06-0104.
- Krakehl, Robert, and Angela M. Kelly. 2021. “Intersectional Analysis of Advanced Placement Physics Participation and Performance by Gender and Ethnicity.” *Physical Review Physics Education Research* 17(2): 020105. doi:10.1103/PhysRevPhysEducRes.17.020105.
- Ladson-Billings, Gloria. 2006. “From the Achievement Gap to the Education Debt: Understanding Achievement in US Schools.” *Educational Researcher* 35(7):3–12.
- Lizotte, Daniel J., Mayuri Mahendran, Siobhan M. Churchill, and Greta R. Bauer. 2020. “Math versus Meaning in MAIHDA: A Commentary on Multilevel Statistical Models for Quantitative Intersectionality.” *Social Science and Medicine* 245:112500. doi:10.1016/j.socscimed.2019.112500.

- Mahendran, Mayuri, Daniel Lizotte, and Greta R. Bauer. 2022a. "Describing Intersectional Health Outcomes: An Evaluation of Data Analysis Methods." *Epidemiology* 33(3):395–405. doi:10.1097/EDE.0000000000001466.
- Mahendran, Mayuri, Daniel Lizotte, and Greta R. Bauer. 2022b. "Quantitative Methods for Descriptive Intersectional Analysis with Binary Health Outcomes." *SSM - Population Health* 17:101032. doi:10.1016/j.ssmph.2022.101032.
- McCall, Leslie. 2005. "The Complexity of Intersectionality." *Signs* 30(3):1771–800.
- McGee, Ebony. 2023. "Fear, Fuel, and Fire!: Black STEM Doctoral Students' Career Decisions during the Trump Presidency." *International Journal of Qualitative Studies in Education* 36(5):716–37.
- Mena, Emily, and Gabriele Bolte. 2019. "Intersectionality-Based Quantitative Health Research and Sex/Gender Sensitivity: A Scoping Review." *International Journal for Equity in Health* 18(1):1–11.
- Merlo, Juan. 2018. "Multilevel Analysis of Individual Heterogeneity and Discriminatory Accuracy (MAIHDA) within an Intersectional Framework." *Social Science and Medicine* 203:74–80.
- Morley, Alicen, Jayson M. Nissen, and Ben Van Dusen. 2023. "Measurement Invariance across Race and Gender for the Force Concept Inventory." *Physical Review Physics Education Research* 19(2):020102. doi:10.1103/PhysRevPhysEducRes.19.020102.
- Morris, Edward W., and Brea L. Perry. 2017. "Girls Behaving Badly? Race, Gender, and the Subjective Evaluation in the Discipline of African American Girls." *Sociology of Education* 90(2):127–48.
- Nash, Jennifer C. 2008. "Re-thinking Intersectionality." *Feminist Review* 89(1):1–15.
- Niehaus, Elizabeth, Corbin M. Campbell, and Karen Kurotsuchi Inkelas. 2014. "HLM behind the Curtain: Unveiling Decisions behind the Use and Interpretation of HLM in Higher Education Research." *Research in Higher Education* 55:101–22.
- Nissen, Jayson, Robin Donatello, and Ben Van Dusen. 2019. "Missing Data and Bias in Physics Education Research: A Case for Using Multiple Imputation." *Physical Review Physics Education Research* 15(2):020106. doi:10.1103/PhysRevPhysEducRes.15.020106.
- Nissen, Jayson M., Ian Her Many Horses, and Ben Van Dusen. 2021. "Investigating Society's Educational Debts due to Racism and Sexism in Student Attitudes about Physics Using Quantitative Critical Race Theory." *Physical Review Physics Education Research* 17(1):010116. doi:10.1103/PhysRevPhysEducRes.17.010116.
- Nissen, Jayson M., Ian Her Many Horses, Ben Van Dusen, Manher Jariwala, and Eleanor W. Close. 2021. "Tools for Identifying Courses That Support Development of Expertlike Physics Attitudes." *Physical Review Physics Education Research* 17(1):013103. doi:10.1103/PhysRevPhysEducRes.17.013103.
- Nissen, Jayson, Ian Her Many Horses, Ben Van Dusen, Manher Jariwala, and Eleanor Close. 2022. "Providing Context for Identifying Effective Introductory Mechanics Courses." *The Physics Teacher* 60(3):179–82.
- Nissen, Jayson, Ben Van Dusen, and Sayali Kukday. 2022. "A QuantCrit Investigation of Society's Educational Debts due to Racism, Sexism, and Classism in Biology Student Learning." *BioRxiv*. doi:10.1101/2022.05.05.490808.
- Ong, Maria. 2023. *The Double Bind in Physics Education: Intersectionality, Equity, and Belonging for Women of Color*. Cambridge, MA: Harvard Education Press.
- Palermo, Martin, Angela M. Kelly, and Robert Krakehl. 2022. "Intersectional Analysis of Advanced Placement Chemistry Enrollment and Performance by Gender and Ethnicity." *Journal of Chemical Education* 99(3):1347–357.
- Quadlin, Natasha, and Jordan A. Conwell. 2021. "Race, Gender, and Parental College Savings: Assessing Economic and Academic Factors." *Sociology of Education* 94(1):20–42.
- Raudenbush, Stephen, and Anthony S. Bryk. 1986. "A Hierarchical Model for Studying School Effects." *Sociology of Education* 59(1):1–17.
- Riegle-Crumb, Catherine, and Eric Grodsky. 2010. "Racial-Ethnic Differences at the Intersection of Math Course-Taking and Achievement." *Sociology of Education* 83(3):248–70.
- Robertson, Amy D., Verónica Vélez, W. Tali Hairston, and Eduardo Bonilla-Silva. 2023. "Race-Evasive Frames in Physics and Physics Education: Results from an Interview Study." *Physical Review Physics Education Research* 19(1):010115. doi:10.1103/PhysRevPhysEducRes.19.010115.
- Roth, Wolff-Michael, and Yew-Jin Lee. 2007. "'Vygotsky's Neglected Legacy': Cultural-Historical Activity Theory." *Review of Educational Research* 77(2):186–232.
- Rubin, Donald B. 1996. "Multiple Imputation after 18 + Years." *Journal of the American Statistical Association* 91(434):473–89.
- Salem, Sara. 2018. "Intersectionality and Its Discontents: Intersectionality as Traveling Theory." *European Journal of Women's Studies* 25(4):403–18.
- Scott, Nicholas A., and Janet Siltanen. 2017. "Intersectionality and Quantitative Methods: Assessing Regression from a Feminist Perspective." *International Journal of Social Research Methodology* 20(4):373–85.
- Shafer, Devyn, Maggie S. Mahmood, and Tim Stelzer. 2021. "Impact of Broad Categorization on Statistical Results: How Underrepresented Minority Designation Can Mask the Struggles of Both Asian American and African American Students." *Physical Review*

- Physics Education Research* 17(1):010113. doi:10.1103/PhysRevPhysEducRes.17.010113.
- Silva, Tony, and Clare R. Evans. 2020. "Sexual Identification in the United States at the Intersections of Gender, Race/Ethnicity, Immigration, and Education." *Sex Roles* 83(11–12):722–38. doi:10.1007/s11199-020-01145-x.
- Simmons, Joseph P., Leif D. Nelson, and Uri Simonsohn. 2016. "False-Positive Psychology: Undisclosed Flexibility in Data Collection and Analysis Allows Presenting Anything as Significant." *Psychological Science* 22(11):1359–66.
- Stewart, John, Geraldine L. Cochran, Rachel Henderson, Cabot Zabriskie, Seth DeVore, Paul Miller, Gay Stewart, and Lynnette Michaluk. 2021. "Mediational Effect of Prior Preparation on Performance Differences of Students Underrepresented in Physics." *Physical Review Physics Education Research* 17(1):010107. doi:10.1103/PhysRevPhysEducRes.17.010107.
- Sun, Shuyan, and Wei Pan. 2014. "A Methodological Review of Statistical Methods for Handling Multi-level Non-nested Longitudinal Data in Educational Research." *International Journal of Research and Method in Education* 37(3):285–308.
- Tabron, Lolita A., and Amanda K. Thomas. 2023. "Deeper Than Wordplay: A Systematic Review of Critical Quantitative Approaches in Education Research (2007–2021)." *Review of Educational Research* 93(5):756–86.
- Tefera, Adai A., Jeanne M. Powers, and Gustavo E. Fischman. 2018. "Intersectionality in Education: A Conceptual Aspiration and Research Imperative." *Review of Research in Education* 42(1):vii–xvii.
- Traweek, Sharon. 2009. *Beantimes and Lifetimes*. Cambridge, MA: Harvard University Press.
- Traxler, Adrienne, Rachel Henderson, John Stewart, Gay Stewart, Alexis Papak, and Rebecca Lindell. 2018. "Gender Fairness within the Force Concept Inventory." *Physical Review Physics Education Research* 14(1):010103. doi:10.1103/PhysRevPhysEducRes.14.010103.
- Van Dusen, Ben. 2018. "LASSO: A New Tool to Support Instructors and Researchers." *American Physics Society Forum on Education*. <https://engage.aps.org/fed/resources/newsletters/newsletter-archive/fall-2018>.
- Van Dusen, Ben, and Jayson Nissen. 2019. "Modernizing Use of Regression Models in Physics Education Research: A Review of Hierarchical Linear Modeling." *Physical Review Physics Education Research* 15(2):020108. doi:10.1103/PhysRevPhysEducRes.15.020108.
- Van Dusen, Ben, and Jayson Nissen. 2020a. "Associations between Learning Assistants, Passing Introductory Physics, and Equity: A Quantitative Critical Race Theory Investigation." *Physical Review Physics Education Research* 16(1):010117. doi:10.1103/PhysRevPhysEducRes.16.010117.
- Van Dusen, Ben, and Jayson Nissen. 2020b. "Equity in College Physics Student Learning: A Critical Quantitative Intersectionality Investigation." *Journal of Research in Science Teaching* 57(1):33–57.
- Van Dusen, Ben, and Jayson Nissen. 2022. "How Statistical Model Development Can Obscure Inequities in STEM Student Outcomes." *Journal of Women and Minorities in Science and Engineering* 28(3):27–58.
- Van Dusen, Ben, Jayson Nissen, and Odis Johnson. 2024. "Society's Educational Debts in Biology, Chemistry, and Physics across Race, Gender, and Class." *ArXiv*. doi:10.48550/arXiv.2404.01314.
- Van Dusen, Ben, Jayson Nissen, Robert M. Talbot, Hannah Huvard, and Mollee Shultz. 2021. "A QuantCrit Investigation of Society's Educational Debts due to Racism and Sexism in Chemistry Student Learning." *Journal of Chemical Education* 99(1):25–34.
- Wasserstein, Ronald L., and Nicole A. Lazar. 2016. "The ASA Statement on p-Values: Context, Process, and Purpose." *The American Statistician* 70(2):129–33.
- Willis, Paul. 1981. *Learning to Labor: How Working Class Kids Get Working Class Jobs*. New York, NY: Columbia University Press.
- Wofford, Annie M., and Christa E. Winkler. 2022. "Publication Patterns of Higher Education Research Using Quantitative Criticalism and QuantCrit Perspectives." *Innovative Higher Education* 47:967–88.
- Womack, Veronica Y., Letitia Onyango, Patricia B. Campbell, and Richard McGee. 2023. "'In the Back of My Mind': A Longitudinal Multiple Case Study Analysis of Successful Black Women Biomedical Graduate Students Navigating Gendered Racism." *CBE—Life Sciences Education* 22(3):ar33. doi:10.1187/cbe.22-06-0130.
- Xiao, Yang, Guiqing Xu, Jing Han, Hua Xiao, Jianwen Xiong, and Lei Bao. 2020. "Assessing the Longitudinal Measurement Invariance of the Force Concept Inventory and the Conceptual Survey of Electricity and Magnetism." *Physical Review Physics Education Research* 16(2):020103. doi:10.1103/PhysRevPhysEducRes.16.020103.

Author Biographies

Ben Van Dusen: I identify as a continuing-generation White cisgender, heterosexual man with a hearing impairment. I was raised in low-income households but now earn an upper-middle-class income. People with similar privileges have created and maintained our society's unjust power structures. I believe it is the obligation of those with the privilege to dismantle oppressive systems. My privilege, however, limits my perspective on the lived experiences of marginalized individuals.

Heidi Cian: I am drawn to studying intersectionality and the use of social groups in research to understand how learners of all ages can experience STEM in ways that affirm how they identify with their bodies, communities,

and histories. I am particularly interested in how intersectional research can consider group membership beyond the traditional demographic categories of gender and race and enlighten understanding of how individuals experience STEM opportunities in ways that respect or reject their politics and cultural wealth. I identify as a White cisgender heterosexual woman with centrist political views and a rural, lower socioeconomic upbringing.

Jayson Nissen: 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 motivated me to reflect on and work to dismantle oppressive power structures in science.

Lucy Arellano: As a Xicana from East L.A., raised in a low-income household, the first in my family to be born in the United States, the first to attend college, moving 2,500 miles away from home to enter an undergraduate STEM major at a prestigious university, I utilize

intersectionality and quantitative methods to understand my own journey through higher education. Now, utilizing my privilege as a cisgender heterosexual professor at an R-1 university, I endeavor to spotlight the multiple systems of oppression endured by minoritized student groups in order to transform postsecondary institutions.

Adrienne D. Woods: My work is in service of expanding educational and societal opportunities for vulnerable or historically underserved children, particularly children with disabilities. I typically analyze large, longitudinal data sets using advanced quantitative methods to better understand whether, how, and to what extent early special education intervention and services may remediate or exacerbate existing educational inequities for children with disabilities. I acknowledge my privilege as a nondisabled White cisgender heterosexual woman from a middle-class background. My perspectives on the experiences of members of the disability community are thus limited but informed by a close relative's experiences within the U.S. education system.