Evolution, 2024, **XX(XX)**, 1–12
https://doi.org/10.1093/evolut/qpae026
Advance access publication 1 March 2024
Perspective

Building conceptual and methodological bridges between SSE's diversity, equity, and inclusion statement and educational actions in evolutionary biology

Gena C. Sbeglia[1] and Ross H. Nehm[2]

[1]Department of Biology, San Diego State University, San Diego, CA, United States
[2]Department of Ecology and Evolution, Stony Brook University, Stony Brook, NY, United States

Corresponding author: Department of Biology, San Diego State University, 5500 Campanile Drive, San Diego, CA 92182, United States. Email: gsbeglia@sdsu.edu

Abstract

The field of evolutionary biology must bridge the gap between its diversity, equity, and inclusion (DEI) commitments and data-driven educational actions in the nation's undergraduate classrooms and degree programs. In this article, we discuss the urgent need for the adoption of *equity frameworks* and why they are centrally important to data-driven DEI efforts in evolutionary biology. We describe why *equity indicators* (e.g., measures) must be anchored in and aligned with equity frameworks. We introduce a specific equity framework for learning (the enhanced educational debt framework) and illustrate how it may be leveraged to document, interpret, and improve outcomes in evolutionary biology. We apply the equity framework and associated indicators to >3,500 students' first college-level experience with evolutionary biology at a public, 4-year institution in the Northeastern United States to demonstrate how these conceptual tools and empirical perspectives may be used by faculty, departments, and degree programs to better understand their roles in mitigating or perpetuating inequities. We end by discussing how this framework may be applied to a range of evolution concepts and courses in the educational hierarchy and used to help evolutionary biologists better understand the extent to which a core aspect of SSE's diversity statement is being realized.

Keywords: evolution learning, educational debt, equity, frameworks, indicators, undergraduate education

Introduction

Members of historically excluded communities (HECs; e.g., marginalized racial/ethnic groups, females, first-generation students, low-socioeconomic status, etc.) continue to be severely underrepresented in the field of evolutionary biology (Graves, 2019; Mead et al., 2015; NSF & NCSES, 2022; O'Brien et al., 2020; Rushworth et al., 2021; Tseng et al., 2020). The Society for the Study of Evolution (SSE), the world's largest organization of evolutionary biologists,[1] has begun to acknowledge, document, and outline strategies for addressing the educational inequities and systemic biases responsible for patterns of disciplinary underrepresentation. SSE's (2017) diversity statement instantiates a dedication to broadening participation of HECs. The diversity, equity, and inclusion (DEI) challenges facing the community are immense, and many goals, topics, and priorities deserve attention. In this article, we focus on one foundational but underexplored facet of DEI in evolutionary biology: disciplinary understanding and learning (cf. National Research Council [NRC], 2019).

SSE's diversity statement commits to "foster[ing] a broader understanding of evolutionary biology" (SSE, 2017), which appropriately identifies disciplinary understanding as a relevant diversity topic (SSE, 2017) (see Supplementary Section 1). Disciplinary understanding is attained by learning and is a fundamental component of disciplinary interest, degree choice, advancement to upper-division coursework, degree attainment, and career participation. Monitoring learning tells us whether our efforts are moving students forward, whereas more commonly collected static measures such as exam scores and course grades (e.g., Denaro et al., 2022) tell us only where students currently are in their educational journey but not how far we have helped them advance. Monitoring learning is essential for instructors and departments to determine whether undergraduate coursework is helping students achieve the knowledge required for disciplinary growth. Therefore, disciplinary learning is a meaningful equity indicator that is core to building both a more diverse *and* prepared workforce.

In the approximately 6 years since SSE drafted its diversity statement, little discussion has focused on how to measure, interpret, track, and improve evolution learning in a way that broadens participation in the field of evolutionary biology. For example, empirical work on evolution education has historically understudied HECs (Dunk et al., 2019; however, see e.g., Nehm & Schonfeld, 2008), and only a handful of large-scale studies disaggregate evolution learning by demographic variables (i.e., Abraham et al., 2009; Sbeglia & Nehm, 2022). Furthermore, empirical work overwhelmingly focuses on knowledge measures at static time points (e.g., beginning *or* end-of-course knowledge) instead of adopting a longitudinal

[1] According to SSE's communications manager, SSE has 3,416 members as of September 2023, which amounts to more evolutionary biologists than the Society for Integrative and Comparative Biology and the European Society for Evolutionary Biology.

Received July 12, 2023; revisions received October 19, 2023; accepted February 28, 2024

Associate Editor: Scott Taylor; Handling Editor: Jason Wolf

approach. Work is therefore needed to operationalize SSE's learning-related DEI priorities into meaningful, data-driven action. In line with this goal, we propose an *equity framework for learning* that is aligned with theoretically grounded indicators of progress toward equitable learning and a corresponding educational action plan.

We begin by discussing what *equity frameworks* are and why they are of central importance to data-driven equity efforts in evolution education. We then describe why *equity indicators* (e.g., measures) must be anchored in and aligned with equity frameworks. We introduce a specific equity framework—*the enhanced educational debt framework for learning*—which incorporates both *who* is learning and *how much* learning is taking place. Using this two-dimensional framework, we illustrate how it may be leveraged to understand and improve equity-related outcomes in evolutionary biology. We apply the equity framework and associated indicators to >3,500 students' first college-level classroom exposure to evolutionary biology at a large, public, 4-year institution in the northeastern United States. This exercise demonstrates how these conceptual tools and empirical perspectives may be used by faculty, departments, and degree programs to better understand the roles they play in mitigating or perpetuating inequities. We end by proposing that this framework could be productively applied to a range of evolution concepts, courses, and levels in the educational hierarchy in order to help faculty better understand the extent to which one component of SSE's diversity statement is being realized.

Evolutionary biologists must adopt equity frameworks, equity indicators, and measures

Quantitative data can be interpreted through a variety of frameworks. Frameworks are explicit articulations of theory and perspective that provide roadmaps for specifying goals, planning actions, and tracking and interpreting progress. They provide a lens or vantage point from which a system (and change in a system) can be understood (Luft et al., 2022; Sbeglia et al., 2021). For example, consider the lenses through which an exterminator, physician, and evolutionary biologist might interpret the detection of a pesticide-resistant mosquito in Florida. The exterminator may envision the event through the lens of business practices (e.g., ineffective pesticide inventory and customer satisfaction). The physician may envision the event through a public health lens (e.g., malaria transmission and patient protection). The evolutionary biologist may envision the event through the lens of phylogenetics and natural selection (e.g., is the mosquito invasive, or is it the result of selection in local populations). The key point is that the lens through which we interpret the world impacts what we notice and how we make sense of it. Yet such frameworks often remain unrecognized as such or implicit components of epistemic practice. Efforts to understand, monitor, and reform systems therefore benefit from the adoption of explicit *equity frameworks* that articulate and anchor the conceptualization of DEI priorities, the design of reform initiatives, and the interpretation of impacts and outcomes (NRC, 2019).

An *indicator* is a measure or statistic used to track progress toward an objective or monitor the condition of a system (Table 1; Supplementary Table S1). Therefore, indicators of *equity* can be used to empirically track progress towards *equity* goals (NRC, 2019). Equity indicators are operationalized by specific *measures*, which generate the actual data that are collected and tracked through time. Measures generate a size, quantity, or degree of something (Table 1; Supplementary Table S1), but they only hold meaning when they are interpreted through a theoretically grounded equity indicator (Figure 1). In other words, the framework is what connects the measures to the indicators. See Figure 1 for a visualization of the relationships between frameworks, measures, and indicators.

For example, test scores are an empirical *measure* that may be interpreted through a given *framework* to be an *indicator* of knowledge disparities among students. The specific lens through which these measures are interpreted and evaluated (e.g., the system is responsible for the disparity vs. the students are responsible for the disparity) can, in turn, motivate

Table 1. Abbreviated definitions of key terms.

Term	Definition
Historically Excluded Community (HEC)	Marginalized groups include racial/ethnic groups, females, first-generation students, low-socioeconomic status, etc.
Between-group difference	An objective description of differences between two or more groups without judgments about their magnitude, significance, or impacts[a]
Disparity	A between-group difference that matters in terms of educational outcomes[a]
Educational inequity	The persistent pattern of between-group disparities in education-related domains such as learning[a]
Indicator	A measure or statistic used to track progress toward objectives or to monitor the health of an economic, environmental, social, or cultural condition over time[a]
Measure (*n*)	A unit used for stating the size, quantity, or degree of something[b]; a tool, instrument, or approach that can generate a size, quantity, or degree of something[b]
Achievement gaps	Differences in achievement (e.g., test scores) among students from different demographic groups[c]
Educational debt	Between-group, education-related disparities produced by decades of historical, economic, sociopolitical, and moral decisions and policies[d]

Note. See Supplementary Table S1 for additional definitions.
[a]NRC 2019.
[b]Oxford University Press.
[c]Coleman et al. (1966).
[d]Ladson-Billings (2006).

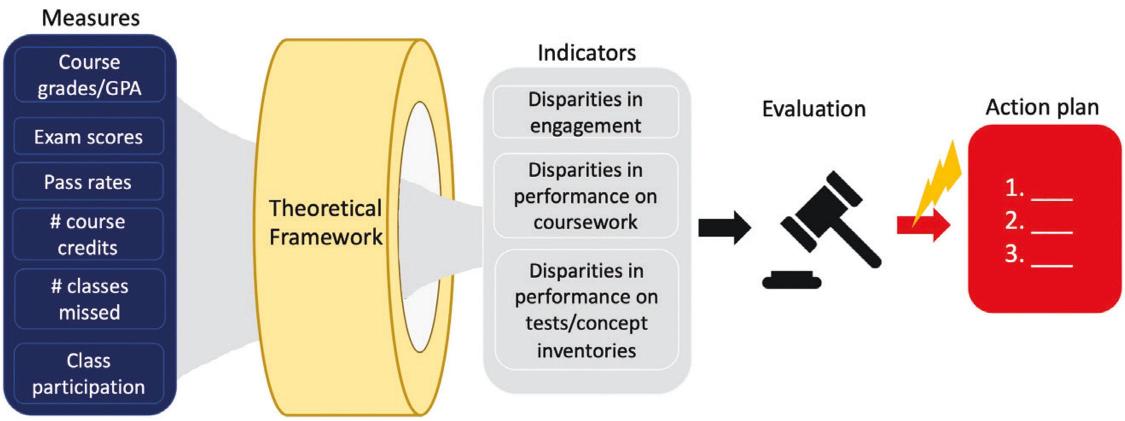

Figure 1. Visualization of the relationship between measures, indicators, and frameworks and the corresponding action plans. The measures and indicators of equity listed here are from the NRC (2019).

an action plan that includes either appropriate (e.g., "fix the system/classroom") or inappropriate actions (e.g., "fix the students") (Shukla et al., 2021). While the approach for the action plan will almost certainly be specific to the institutional context, equity indicators that are situated within evidence-based equity frameworks *need not* vary by context because they serve to set a stable benchmark from which educational equity patterns can be interpreted. This is a key point in NRC (2019). Unfortunately, DEI work is often focused on measures that are not explicitly tied to equity indicators or frameworks, resulting in potentially problematic measures and interpretations (e.g., knowledge disparities can be interpreted as indicating inequities in academic privilege or differences in innate intelligence). As a result, it is critical that evolutionary biologists adopt *equity frameworks* that are conceptualized by theoretically grounded *indicators of equity* and operationalized by empirical *measures*. In the next section, we introduce a specific equity framework that facilitates the interrogation of one specific component of equity that evolutionary biologists are well-suited to tackle: evolution learning in undergraduate classrooms.

The educational debt framework and associated indicators and measures of learning

Learning may be operationalized by measures along multiple dimensions, and we emphasize two categories in need of measurement: *how much* learning is occurring and *who* is successfully doing it. Recent work in other disciplines (e.g., chemistry, physics; see Nissen et al., 2021; Van Dusen & Nissan, 2019; Van Dusen et al., 2022) has anchored measures of *who* is learning within an educational debt framework (Ladson-Billings, 2006). The educational debt framework is a powerful, social-justice-aligned approach for conceptualizing demographic disparities in educational outcomes (Ladson-Billings, 2006). Educational debt acknowledges the structural and institutional barriers to opportunity and achievement that have accrued over time. These barriers have enormous and compounding impacts on nearly every aspect of a person's life, starting from early childhood. In particular, structural barriers to opportunity and access can produce income inequality and residential segregation, leading to disparities in access to resources that support learning, development, and health (Ladson-Billings, 2006; NRC, 2019). Research has shown that students from historically excluded racial, ethnic, socioeconomic, and linguistic communities are more likely to experience poverty and differential treatment within educational systems, as well as insufficient access to healthcare, high-quality schools, books, and experienced teachers (NRC, 2019; e.g., Neuman & Moland, 2019; White et al., 2012; Williams, 1999). The significant and persistent exposure to discrimination influences students' stress response and executive functioning, as well as school readiness, attendance, engagement, and performance (Levy et al., 2016; Merolla & Jackson, 2019; Myers, 2009; NRC, 2019). By adolescence, opportunity differences have accumulated, resulting in stark differences in the overall preparedness of college-bound students that often translate into disparities in education-related outcomes among the nation's undergraduates and STEM workforce (Harris et al., 2020; PCAST, 2012; Salehi et al., 2020, 2021). These patterns of between-group, education-related disparities produced by decades of historical, economic, and sociopolitical policies have been theorized as the "educational debt" owed by society to students (Ladson-Billings, 2006; Shukla et al., 2021; Van Dusen et al., 2022).

Within the context of undergraduate education, students from centered and marginalized backgrounds enter courses with differences in knowledge (and many other variables) due to systemic inequities. These disparities reflect the incoming educational debt owed to marginalized students. Degree programs and their constituent courses may impact these incoming knowledge disparities in one of three ways: they may (a) *mitigate the debt* (i.e., HECs groups have *higher* mean learning gains than centered groups), (b) *perpetuate/maintain the debt* (i.e., both groups have equivalent gains), or (c) *exacerbate the debt* (i.e., HECs have *lower* mean gains than centered groups) (van Dusen et al., 2022). This framework has great potential for enabling data-driven approaches that guide meaningful and effective action within educational systems. However, there are two limitations that could hinder its usability as an empirical framework for courses, departments, and degree programs, which we address through the modifications described below.

First, Van Dusen's framework conceptualizes the educational debt of learning as relating to *who* is learning, but it does not consider the *magnitude* of knowledge change as a salient or potentially variable measure. Rather, there is an implicit assumption that meaningful learning *will* occur. However, prior work in many STEM fields has shown that this assumption is often false, particularly for the topic of evolution (Nehm & Reilly, 2007). Therefore, we propose an

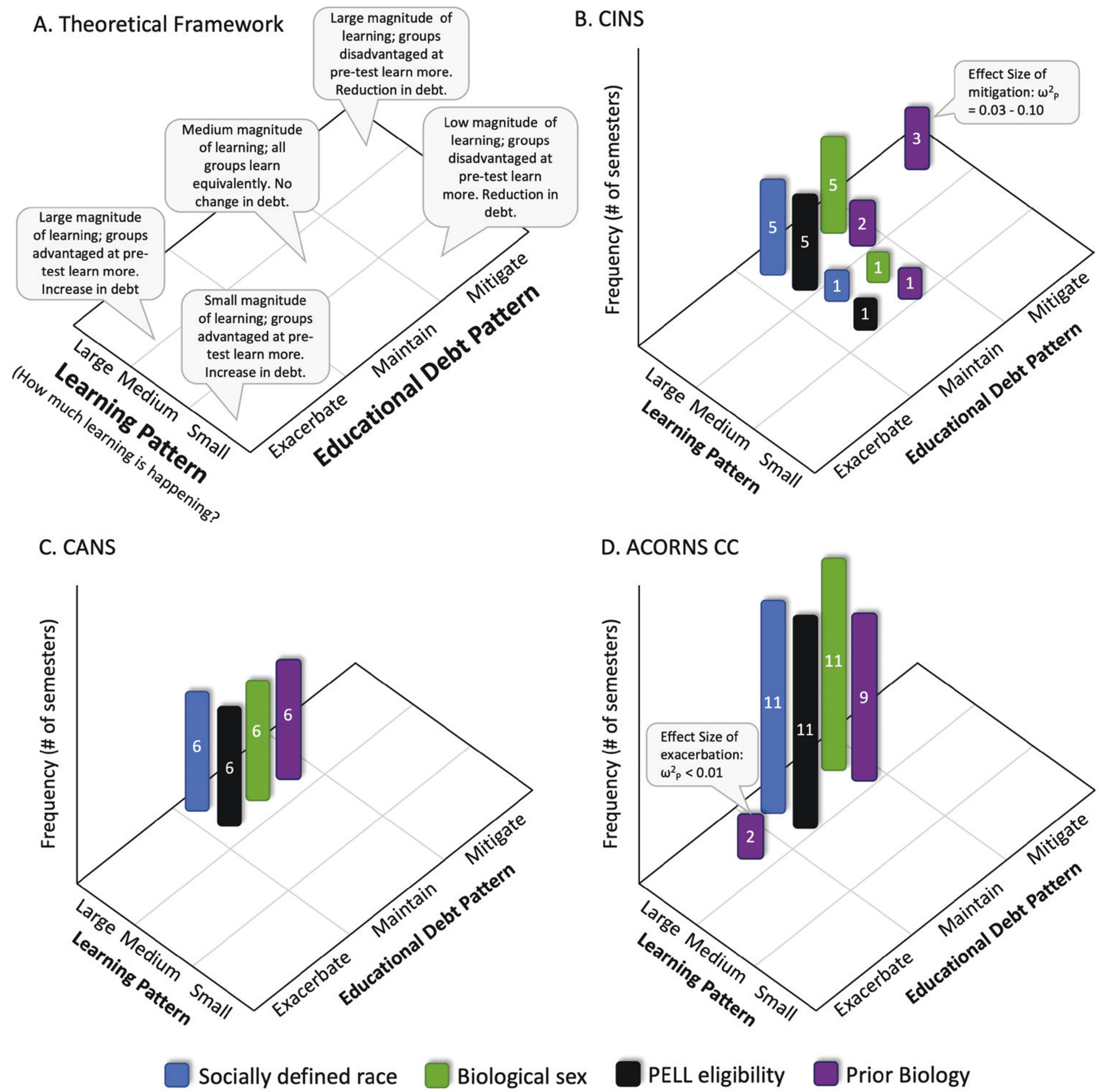

Figure 2. Visualization of the two-dimensional educational debt framework for evolution learning (A). The intersection of learning and educational debt patterns represents the indicator of equity. (B)–(D) The application of the CINS, CANS, and ACORNS results to this framework for the 11 semesters studied. Partial omega squared (ω^2_p): Small = 0.01, medium = 0.06, large = 0.14 (Lakens, 2013). Odds ratio (OR): small = 1.68 (0.59), medium = 3.47 (0.29), large = 6.7 (0.15) (Chen et al., 2010). All effect sizes are significant at $p < 0.001$.

extension of the educational debt framework to incorporate the amount or magnitude of learning as a critical dimension along which patterns of debt repayment can be evaluated. The intersection of these two dimensions (i.e., debt and learning magnitude) represents the *indicator of equity* (see Figure 2A). Specifically, the box in which a course's learning patterns land may be interpreted as an indicator for evaluating how equitable the course is for this particular educational outcome. We argue that a two-dimensional conceptualization of learning—learning debt repayment *and* learning magnitude—is critical for understanding the experiences of diverse learners as they navigate institutional structures.

Second, we propose modifying Van Dusen's terminology to replace the word "perpetuating" with "maintaining." In Van Dusen's framework, "perpetuating" debt describes learning patterns in which all students learn equivalently regardless of backgrounds and incoming debt. Although this pattern indicates that educational debt has not been reduced, it also indicates that it has not increased. The disproportionately high loss of HEC students from STEM pathways, especially between sophomore and junior years (PCAST, 2012; Thiry et al., 2019), raises concerns that many gateway courses may frequently exacerbate the debt, whereas courses undergoing early stages of reform may frequently maintain the debt at precourse levels (although little is known about how equity-related progressions occur during institutional reform). Given this context, using the term "perpetuating" to refer to conditions in which the debt does not change has limitations.

For example, fostering meaningful learning outcomes in introductory STEM courses is effortful and challenging, but often underappreciated work. Labeling this outcome with a term that evokes negative and/or judgmental connotations is unnecessary. In addition, the term may discourage instructors, departments, and institutions from utilizing the framework. Therefore, we propose the more neutral and descriptive term "maintaining," which has also been used by these authors in other work (e.g., Nissen & Van Dusen, 2021).

It is crucial to contrast educational debt frameworks with commonly discussed "achievement gap" frameworks. By their very definition, the latter adopt a deficit perspective, place the burden of "achievement" on marginalized students, and frame the "gap" relative to advantaged demographic group "achievement" (vs. centered; Shukla et al., 2021; e.g., Cohen et al., 2006; Haak et al., 2011; Harackiewicz et al., 2014; Harris et al., 2020; Jordt et al., 2017; Salehi et al & 0.2019; Simmons & Heckler, 2020; Theobald et al., 2020). Achievement gap frameworks are considered problematic because they are not aligned with evidence demonstrating the considerable inequities at play in educational "achievement" (Ladson-Billings, 2006; NRC, 2019; Shukla et al., 2021).

The enhanced debt framework can strengthen empirical evaluation of progress toward equitable outcomes and targets

Four major points have been raised thus far: (a) The SSE has highlighted disciplinary understanding—which is achieved through learning—as a relevant diversity topic; (b) equity frameworks and associated measures are critical for grounding all DEI work within evolutionary biology, (c) reframing outcome disparities away from achievement gaps and towards educational debt frameworks minimizes student deficit perspectives, and (d) our enhanced debt framework offers advantages for uptake and empirical investigations. Next, we illustrate how this enhanced educational debt framework can be applied to classrooms and what information relevant to DEI efforts can be obtained using it.

Specifically, using the enhanced educational debt framework, we ask: (a) What is the magnitude of educational debt among HEC student groups at the start of their undergraduate evolution education? (b) What impact does instruction have on students' educational debt (i.e., is it mitigated, maintained, or exacerbated), and what is the overall magnitude of learning? After answering these questions, we discuss how this framework could be used more broadly by evolution educators to integrate learning (and other critical variables) into DEI plans and monitor educational equity in degree programs and departments.

Methods for operationalizing frameworks and indicators: a case study

Setting and sample

Pre- and postcourse data were collected from 11 semesters of a large (>250 students) gateway biology course at a research university in the northeastern United States. This course was taught by the same two biologists with minimal formal preparation in biology education. The classes represented a middle ground between traditional lecture-based instruction and reformed active learning environments (~10%–36% active learning as measured by the COPUS instrument [Smith et al., 2013]). The amount of class time spent on evolution and genetics instruction did not change meaningfully across semesters (range: ~24%–31%). The course serves a diverse student population in terms of self-reported socially defined race (~18% URM), biological sex (57% female and nonbinary[2]), prior preparation (~33% no prior college biology), socioeconomic status (~39% PELL eligible) (Supplementary Table S2), religiosity (Supplementary Figure S1), and evolution acceptance (Supplementary Figure S2). Although only a minority (37.2%) of the students at the institution are white, like most universities in the United States, the institution could be considered a historically white institution. Student demographic and background data were gathered from a combination of self-report surveys (i.e., socially defined race, prior biology, biological sex) and information from the University's Office of Institutional Research (i.e., PELL eligibility). See Supplementary Section 2 for detailed sample information.

Instruments

This study employed robustly validated instruments to measure evolution understanding in over 3,500 unique students. The CINS (Anderson et al., 2002) and CANS (Kalinowski et al., 2016) are multiple-choice instruments that measure evolution knowledge and are completed by 2691 students (seven semesters, 76% participation rate) and 2671 students (six semesters, 82% participation rate), respectively. The ACORNS (Nehm et al., 2012; Opfer et al., 2012) is a constructed-response instrument that measures evolution knowledge (i.e., normative ideas: ACORNS CC) and misconceptions (ACORNS MIS) through the scientific practice of explanation (see Supplementary Table S3 for the items). ACORNS responses were scored using the AI-based tool EvoGrader (see Moharreri et al., 2014; Beggrow et al., 2014). The ACORNS was completed by 3203 students (11 semesters, 78% participation rate). Evolution acceptance was treated as a control variable and measured using the I-SEA instrument (Nadelson & Southerland, 2012). See Supplementary Section 3 for detailed information about the instruments.

Analysis

Four evolution understanding outcome variables (i.e., the CINS, CANS, ACORNS CC, ACORNS MIS) were collected during the first and last 2 weeks of the semester. Each observation in the model was characterized by 4 student-level variables: prior biology courses, self-reported biological sex, self-reported socially defined race, PELL eligibility, and pretest I-SEA score. See Supplementary Section 4 for additional details about these variables.

To determine how much evolution learning occurred throughout the course, multiple hierarchical linear models (for CINS, CANS, and ACORNS CC) and logistic models (for ACORNS MIS) were run with a measure of evolution understanding modeled as an outcome variable; time, socially defined race, biological sex, PELL eligibility, and pretest evolution acceptance were modeled as predictors. To determine if students of various backgrounds differed in their pretest evolution knowledge and gained similar magnitudes of evolution knowledge, time was modeled as having an interaction effect with each background variable. An insignificant interaction effect would indicate that students of various background

[2] Biological sex was self-reported and some respondents reported it as nonbinary.

variables gained knowledge comparably from pre- to posttest. In this model, the coefficients of the main effect of a particular background variable (e.g., socially defined race or biological sex) could be interpreted as the relationship between that background variable and pretest evolution knowledge. The *magnitude* of learning and the *magnitude* of educational debt were quantified using partial omega squared (ω^2_p) and odds ratios (OR), as appropriate. See Supplementary Section 4 for additional information about the analysis, models, and effect size thresholds.

Case study results

Key finding 1: At course entry, students of all backgrounds had extremely low levels of evolution knowledge and high levels of evolution misconceptions

At course entry, the sample (N ~3,500 unique students) had a mean evolution knowledge score of 58.7% (SD = 4.19) according to the CINS and 46.4% (SD = 5.06) according to the CANS (100% maximum). Furthermore, according to the ACORNS, half of the sample (49%) did not utilize *any* evolutionary core concepts (CC, i.e., variation, heredity, differential survival) in their explanations of evolutionary change, and over a third exhibited at least one of the major misconceptions (MIS) about evolution (e.g., need, use/disuse). See Supplementary Table S4 for a summary of pretest evolution knowledge scores by student background and Supplementary Figures S3–S6 for a visualization of these patterns.

Key finding 2: Multiple HECs were characterized by incoming educational debt in evolution knowledge

Students from HECs (e.g., PELL eligible, female + nonbinary, underrepresented minority [URM], low prior academic preparation) had significantly lower evolution understanding at course entry than more advantaged students (Supplementary Table S5A), which indicates the presence of incoming educational debt. Because the analysis controlled for all other student variables (including evolution acceptance), these incoming disparities in evolution understanding can be *uniquely* attributed to each background variable. Specifically, Black/African American and Hispanic students, female + nonbinary students, PELL-eligible students, and students with lower prior preparation had significantly lower incoming evolution knowledge according to at least one of the three evolution knowledge measures used in this study (i.e., CINS, CANS, ACORNS CC) ($p < 0.001$ and ω^2_p = <0.01–0.04 for all significant cases, see Supplementary Table S5A for details). In real terms, the analyses revealed that Black/African American students and those with no prior biology scored 5%–9% lower on the CINS and CANS at pretest than more advantaged groups. Female + nonbinary students and PELL-eligible students scored ~3%–4% lower. See panels (A)–(C) in Supplementary Figures S3–S6 for a visualization of these patterns of incoming evolution knowledge, and see Supplementary Table S6 for per-semester statistical results.

Key finding 3: "Misconceptions" at course entry were generally comparable across groups

Over 11 semesters, only one group (those with low prior preparation in biology) showed evidence of incoming educational debt for misconceptions (specifically, this group had about one and a half times the likelihood of using evolution misconceptions as compared to those with prior biology experience; Supplementary Table S5A). In contrast, URM students, PELL-eligible students, and female + nonbinary students had statistically similar levels of incoming misconceptions as compared to their more centered peers (see Supplementary Table S5A for details). See panel D in Supplementary Figures S3–S6 for a visualization of this pattern of incoming evolution misconceptions, and see Supplementary Table S7 for per-semester statistical results.

Key finding 4: All groups demonstrated high magnitudes of evolution learning

Overall, unlike the findings from many prior studies, students experienced significant and meaningful gains in evolution knowledge during their first major exposure to evolution in college, with effect sizes ranging from ω^2_p = 0.37–0.48 ($p < 0.001$) for the three knowledge instruments. Students also displayed significant and meaningful declines in misconceptions (ACORNS MIS: $p < 0.001$, OR = 0.17). Magnitudes of learning and misconception loss differed by semester; however, some course iterations had a medium effect on instruction and others had a large effect (Figure 2B–D; see Supplementary Tables S8 and S9 for the per-semester statistical results for patterns of evolution knowledge and misconception change for all instruments).

Key finding 5: Students' first university exposure to evolution did not increase educational debt for any HEC

Disparities in precourse evolution understanding generally did *not* explain magnitudes of evolution learning either in terms of gaining knowledge or losing misconceptions; students of all backgrounds learned equivalently (see Supplementary Table S5B for details). Put another way, the incoming educational debt of students in this sample did not hinder evolution learning. Therefore, educational debt was generally not exacerbated by the course but rather was *maintained* (Figure 2B–D). The exception to this pattern was for groups differing in prior biology coursework: in three of six semesters in which the CINS was administered, students with no prior biology had proportionally *higher* knowledge gains than students with more prior biology (Supplementary Table S6). In these semesters, the incoming educational debt for evolution knowledge was *mitigated* by the course and the magnitude of the debt reduction ranged from small to medium[3] (Figure 2B).

Implications for DEI efforts in evolutionary biology

Diversifying the scientific workforce requires the integration of learning into DEI efforts

Learning environment reform must be viewed as a core DEI priority and should not be eclipsed by other equity indicators (e.g., sense of belonging, microaggressions, implicit bias). Persistent misconceptions and low levels of incoming knowledge are common in undergraduate classrooms across many STEM disciplines, including evolutionary biology (e.g., Alters & Nelson, 2002; Andrews et al., 2011; Coley & Tanner, 2017; McCloskey, 1983; Nehm & Reilly, 2007; Orgill &

[3] Additionally, in 2 of 11 semesters in which the ACORNS was administered, the amount of prior biology was significantly positively associated with evolution learning, but the effect size was negligible (Figure 2D).

Sutherland, 2008; Smith, 2010a, b; this study). Because disciplinary learning is critical to major declaration, advancement to upper division coursework, degree attainment, and career participation, DEI efforts that do not monitor and improve the amount and equity of learning are unlikely to broaden STEM degree and career participation. In other words, a continued focus on static student outcomes like exam scores and course grades will not be sufficient for reaching the field's stated DEI objectives. Simply put, learning *must* be one of the indicators of evidence-based efforts directed at intentional systemic change, and this work benefits from the integration of explicit equity frameworks (NRC, 2019). Although departments and degree programs nationwide have been working toward reforming introductory courses to improve educational outcomes (e.g., Handelsman et al., 2022; Nardo et al., 2022; Nehm et al., 2022), these efforts typically lack robust measures of learning as well as equity frameworks through which to interpret and evaluate these outcomes.

In this study, we report that students from HECs had lower incoming evolution knowledge than students from centered groups, which we interpret within our framework to be evidence of educational inequity. Although the effect size of these disparities ranged from small to medium according to well-accepted statistical thresholds (Chen et al., 2010; Lakens, 2013), the functional impact of incoming disparities of this size on sustained disciplinary participation must *not* be underestimated. First, the findings reported in this article are robust; all analyses controlled for the impact of evolution acceptance and all other background variables, so the effect size for any given identity variable is associated with only that variable. Furthermore, these disparities were replicated in *each* of the 11 semesters studied. Second, the educational debt framework posits that disadvantages accumulate to produce patterns such as disparities in degree completion and disciplinary participation. When considering that this article addresses just *one* of the many topics and competencies that are relevant to participation in the field of evolutionary biology, even so-called "small" disparities for a given topic have an enormous potential to negatively impact students. Finally, we contend that there is no acceptable reason for *any* group-level disparities in evolution knowledge by demographic variables (e.g., socially defined race, biological sex, SES, first-generation status). According to policy documents and state education standards (e.g., Brewer & Smith, 2011; NGSS; NGSS Lead States, 2013), evolution is a central organizing principle for the biological sciences, and understanding it is an expected outcome of high school education. If students are not meeting these required state educational standards and these patterns correlate with their identities, we consider this pattern inherently inequitable. There are many documented causes for such patterns, including unequal access to and treatment in K-12 educational institutions (e.g., Graves et al., 2022). Therefore, an argument that patterns of knowledge disparities like those reported here were produced due to students simply choosing not to take a course where evolution was taught is unlikely.

Equity frameworks and indicators facilitate the monitoring of DEI objectives

This article is our response to the urgent need for evolutionary biologists to adopt indicators of equity situated within appropriate equity frameworks in order to understand, monitor, and reform undergraduate classrooms (Shukla et al., 2021; Tashiro & Talanquer, 2021). This work builds upon theoretical and early empirical discussions of educational debt in learning outcomes by integrating two empirical dimensions—debt of learning and magnitude of learning—to generate more nuanced indicators of educational equity in line with NRC (2019) recommendations. Integrating learning magnitude into this framework has the potential to greatly impact the interpretation of learning-based assessment data. For example, the mitigation of incoming educational debt in disciplinary knowledge has been interpreted as an equitable learning outcome (van Dusen et al., 2022). However, with low levels of student learning, debt mitigation would actually do little to diversify STEM majors, degree recipients, and professionals because disciplinary learning is a critical prerequisite to degree and career participation. Therefore, a goal of instructional reform is the mitigation of debt in learning *while also* generating high magnitudes of learning.

Applying this two-dimensional equity framework to classrooms is one small step for educators and departments to generate NRC-aligned measures and indicators to locate courses within the institutional change process (i.e., assessment), make value judgments regarding how equitable the patterns are (i.e., evaluation), and inform next steps (i.e., systemic change planning) (Figure 3). In this way, the framework can guide

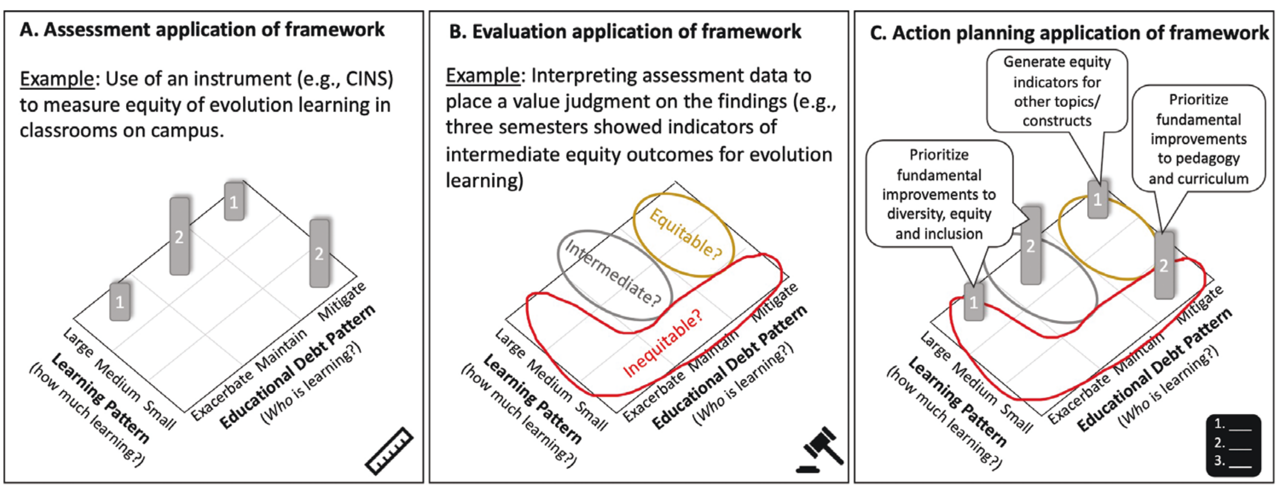

Figure 3. Assessment (A), evaluation (B), and action planning (C) application of the framework.

the measurement, assessment, and evaluation of equity in the classroom as well as inform the development of action plans for reform. Importantly, not all classroom challenges mandate the same solutions. *Small* pre–post gains in understanding, for example, may indicate that the classroom environment is not supporting optimal learning, which may motivate prioritizing fundamental (as opposed to incremental; Cuban, 1999) reforms of pedagogy, assessment, and curriculum (see Supplementary Table S1). In such cases, the pattern of learning-related debt mitigation/exacerbation may not be particularly informative because students are not learning very much in the first place. In contrast, *large* pre–post gains in understanding paired with debt *exacerbation* or *maintenance* may indicate that optimal conditions are present for learning for some students but not for all students, which may motivate prioritizing fundamental reforms related to diversity, equity, and inclusion. As reform efforts are implemented, this framework allows educators to track a course's growth and facilitate a data-driven approach to understanding conditions that promote or inhibit equitable learning outcomes (Nissen et al., 2021).

In this article, we applied our theoretical framework at a research university and studied longitudinal data from 11 semesters (>3,500 unique students) to generate measures and indicators of equity in evolution learning. The results showed that in most semesters, the course *maintained* incoming educational debt and facilitated *large* evolution learning gains. Taken together, these two measures (debt patterns and learning patterns) serve as an equity indicator that is capable of categorizing the degree to which equitable learning is occurring in this course. In particular, they show that students learned evolution at similarly high magnitudes regardless of background and that incoming disparities in evolution knowledge remained stable throughout the semester.

Research suggests that many gateway college courses are *not* generating large magnitudes of learning and are often deepening disparities among students (i.e., the robust finding that HECs are disproportionately "weeded out" of STEM majors; e.g., Hatfield et al., 2022; Nissen et al., 2021; Riegle-Crumb et al., 2019). Prior to the initiation of reform efforts, these problematic patterns were indeed apparent in the gateway course that is the subject of this article. Against this backdrop, learning patterns that indicate intermediate equity outcomes (like those found in this course) should be viewed as meaningful progress toward equitable learning and used to guide the development of a roadmap for continued improvement. In the case of the gateway course highlighted in this article, maintaining debt with high magnitudes of learning suggests optimal learning conditions may be present, but they are not equitably experienced by all students, thus motivating an action plan geared towards DEI-related reforms of the learning environment. These are the types of reforms being implemented in this course since these data were collected (Supplementary Table S10 shows some examples of recent reforms). Ongoing assessment and evaluation work will generate and use these equity indicators to monitor the progress from intermediate to equitable course outcomes. An important point is that without frameworks and data it is challenging to gauge whether progress is being made by courses and departments.

Psychosocial interventions may interact with learning and thus debt mitigation

The evolution instruction in the gateway course in this study was designed using a traditional "knowledge-only" approach, and did not include psychosocial interventions (e.g., values affirmation interventions [e.g., Harackiewicz et al., 2014], evolution acceptance or conflict-based interventions [e.g.; Green & Delgado, 2021]). Our results suggest that this knowledge-only approach may have differentially benefited students with low prior preparation, the only HEC in this study with mitigated debt for evolution knowledge. In fact, prior work has shown that students with lower prior preparation more readily benefit from reform efforts than other HECs (e.g., Tashiro & Talanquer, 2021). Given that evolution knowledge and learning may be intertwined with student identity and beliefs (Bailey et al., 2011; Barnes et al., 2020; Borgerding et al., 2016), reforms in this course are incorporating various psychosocial instructional approaches and interventions (e.g., perceptions of conflict with evolution and religion, course structure promoting belongingness) that may support HEC learning. The present study is foundational to these future efforts because it generates a robust baseline to assess the added value of these different types of interventions to the evolution education outcomes. Unfortunately, most existing intervention work does not have appropriate baselines or comparison groups from which to interpret the findings.

Not all problematic patterns in a classroom are captured by an educational debt approach

The expanded educational debt framework has great potential to help uncover, quantify, and monitor inequitable learning outcomes, but it is important to understand its limitations. First, although many education-related disparities are likely due to systemic inequities, this is not necessarily the case for all empirical patterns of debt in educational settings. Rather, some topics may be more culturally relevant to some groups than others, which could generate knowledge differences that are due to interest and exposure but not inequity. For example, as a group, dog owners may have more knowledge about dog parks than cat owners.

Second, not all problematic patterns are inequitable, meaning that some educational patterns will *not* show educational debt yet still signal an important problem. For example, although evolution misconceptions were concerningly common in our sample, educational debt for misconceptions was typically lacking. In other words, evolution misconceptions at course entry were equivalent among groups; many students had them regardless of socially defined race, biological sex, and SES (corroborating prior work; Nehm & Schonfeld, 2008). However, this does *not* mean that this pattern is desirable. In fact, when left unchallenged, misconceptions can interfere with the building of normative models of biological explanations; explicitly addressing misconceptions about evolution at increasingly high doses has been found to improve evolution learning (Nehm et al., 2022). It is also possible for students to differentially benefit from this and other evidence-based approaches. Attention to learning magnitudes disaggregated by salient student variables is therefore needed even in cases where educational debt at course entry is empirically absent.

The enhanced educational debt framework for evolution learning represents a starting point for broader empirically informed and equity-aligned goals.

We envision the enhanced debt framework as being flexible in several critical ways. First, this framework is suitable for a variety of undergraduate settings (not just introductory courses) and content areas (not just evolution). Indeed, tackling one disciplinary topic in one course will *not* be sufficient to broaden participation in STEM fields. This entire framework needs to be applied in degree programs.

Second, the large/medium/small categories of learning magnitude were determined based on standard effect size cutoffs, but this axis need not be limited to three categories or be categorical at all. Alternate scales and cutoffs along the learning magnitude axis are possible as long as they are based on evidence. In fact, our results show that the magnitudes of evolution learning in the "large" category discussed above varied substantially (see Supplementary Figure S7).

Third, this framework is suitable for intersectional approaches to identity classification. Research has demonstrated that the intersectionality of student identities is critical for understanding and mitigating inequitable outcomes (Hatfield et al., 2022; Keller et al., 2023; Nissen et al., 2021; Young et al., 2022). For example, Nissen et al. (2021) showed that the largest educational debts in a physics course were owed to Black and Hispanic women, and Young et al. (2022) showed that the greater number of marginalized identities a person holds, the worse their grades were in a course (Young et al., 2022). However, the field lacks consensus on how to integrate intersectional identities into analytical frameworks (Keller et al., 2023). Two main approaches have emerged: (a) categorizing participants by the intersection of multiple specific identities (e.g., biological sex and socially defined race) and (b) generating an index that represents the *number* of advantaged or marginalized identities a person holds. Both approaches have strengths and limitations.

In the first approach, specific identity dimensions are preserved and may be modeled as interaction effects in statistical models, but modeling the intersection of more than two identity dimensions is computationally unwieldy and often has substantial sample size limitations (Evans et al., 2018; Keller et al., 2023). Furthermore, modeling intersectionality in this way may make the inclusion of other variables whose intersections might be of interest (e.g., the dosage of active learning) to be untenable. In the second approach, an index represents a person's total number of marginalized or advantaged identities (e.g., Barrier index, Young et al., 2022). It assigns one point for each of three binary identities that are considered to be historically marginalized in higher education: URM status, first-generation status, and low-socioeconomic status. The barrier index ranged from 0 (non-URM, continuing generation, high SES) to 3 (URM, FG, low SES) and was found to correlate with metrics of performance in courses (Young et al., 2022). A limitation is that this approach treats various marginalized identity dimensions as equivalent. Recently, a third approach (called the MAIHDA approach) has been proposed and has attempted to address the limitations of current methodologies (Keller et al., 2023). In this approach, individuals are modeled at the first level of a multilevel regression analysis, and combinations of multiple social identities are modeled at the second level. Although more work is needed regarding how to model intersectional identities, any of these approaches could be integrated into the expanded educational debt framework.

It is important to approach quantitative data from a variety of lenses. Further work examining these findings through alternative lenses could also add crucial insights and would be useful next steps. For example, our study uses numerical scores for the constructs we measured, but the field may want to adopt criterion-based quantitative perspectives (e.g., all students explain evolution without misconceptions). Furthermore, studies using a quantitative critical race theory lens (e.g., Nissen et al., 2021) have increasingly raised questions about statistical approaches and language in education studies (e.g., the practice of setting reference points during statistical analyses, collapsing racial groups, analyzing HEC as discrete instead of intersectional groups, fetishization of gaps). These alternative theoretical lenses could further enhance the educational debt framework highlighted in this article, but they may also provide entirely new and better frameworks. Thus, although the educational debt framework is, in our view, the best theoretical framework that currently exists for bridging the divide between SSE's DEI statement and educational practice, the critical takeaways from this article are that evidence-based frameworks are critical to broadening participation in evolutionary biology and they must include learning.

Conclusion

Over the past few decades, SSE has increased its attention to evolution education (Alters & Nelson, 2002) and, more recently, to equity in the field of evolutionary biology (e.g., Rushworth et al., 2021). Accordingly, SSE's diversity statement explicitly links its DEI objectives with commitments to "foster a broader understanding of evolutionary biology" (SSE, 2017). Although many components of SSE's diversity statement remain in urgent need of operationalization and application in educational settings, this article serves to bridge this one aspect of SSE's diversity statement (i.e., disciplinary learning) with educational actions in classrooms and degree programs. The *equity frameworks* and *equity indicators* for evolution learning introduced in this article (a) align with existing theoretical frameworks to improve learning outcomes in science education, (b) provide a conceptual and empirical bridge between DEI statements and DEI outcomes, and (c) outline an approach that faculty, departments, and degree programs can use to better understand their roles in mitigating or perpetuating inequities in evolution understanding and learning. Applying our extended educational debt framework to >3,500 students' first college-level experience with evolutionary biology provided a practical illustration of how the framework may be used to evaluate course outcomes and develop action plans for improvement. This framework, coupled with additional equity indicators, offers potential application to a diverse array of evolution concepts, courses, and levels in the educational hierarchy. It also facilitates an evidence-based understanding of the extent to which SSE member's courses, departments, and institutions are effectively addressing inequities in our field.

Supplementary material

Supplementary material is available online at *Evolution*.

Data availability

The data underlying the article were granted an exemption to the journal's data sharing policy due to ethical and privacy restrictions; however, data will be shared on reasonable request to the corresponding author. Other data are available as Supplementary Material.

Author contributions

G.C.S. and R.H.N. both contributed to data collection, data organization, paper conceptualization, and manuscript writing. G.C.S. analyzed the data and produced the figures and tables.

Funding

Support for data collection was provided by National Science Foundation grant no. TUES-1322872, and analysis was supported by a Howard Hughes Medical Institute Inclusive Excellence grant. Any opinions, findings, conclusions, or recommendations expressed in this publication are those of the authors and do not necessarily reflect the views of the National Science Foundation or HHMI.

Conflict of interest: The authors report no conflicts of interest.

Acknowledgments

We thank Professor John True for his support and commitment throughout this project.

Ethical statement

The study was approved by the university's institutional review board (protocol no. 504271) and was classified as *not human subjects research*. The procedures outlined in the present article are in accordance with the ethical standards of the responsible committee on human experimentation (institutional and national) and with the Declaration of Helsinki 1975.

References

Abraham, J. K., Meir, E., Perry, J., Herron, J. C., Maruca, S., & Stal, D. (2009). Addressing undergraduate student misconceptions about natural selection with an interactive simulated laboratory. *Evolution: Education and Outreach*, *2*(3), 393–404. https://doi.org/10.1007/s12052-009-0142-3

Alters, B. J., & Nelson, C. E. (2002). Perspective: Teaching evolution in higher education. *Evolution*, *56*(10), 1891–1901. https://doi.org/10.1111/j.0014-3820.2002.tb00115.x

Anderson, D. L., Fisher, K. M., & Norman, G. J. (2002). Development and evaluation of the conceptual inventory of natural selection. *Journal of Research in Science Teaching*, *39*(10), 952–978. https://doi.org/10.1002/tea.10053

Andrews, T. M., Leonard, M. J., Cogrove, C. A., & Kalinowski, S. T. (2011). Active learning not associated with student learning in a random sample of college biology courses. *CBE—Life Science Education*, *10*, 329–435. https://doi.org/10.1187/cbe.11-07-0061

Bailey, G., Han, J., Wright, D., & Graves, J. (2011). Religiously expressed fatalism and the perceived need for science and scientific process to empower agency. *The International Journal of Science in Society*, *2*(3), 55–88. https://doi.org/10.18848/1836-6236/cgp/v02i03/51265

Barnes, M. E., Supriya, K., Dunlop, H. M., Hendrix, T. M., Sinatra, G. M., & Brownell, S. E. (2020). Relationships between the religious backgrounds and evolution acceptance of black and hispanic biology students. *CBE—Life Sciences Education*, *19*(4), ar59. https://doi.org/10.1187/cbe.19-10-0197

Beggrow, E. P., Ha, M., Nehm, R. H., Pearl, D., & Boone, W. J. (2014). Assessing scientific practices using machine-learning methods: How closely do they match clinical interview performance? *Journal of Science Education and Technology*, *23*(1), 160–182. https://doi.org/10.1007/s10956-013-9461-9

Borgerding, L. A., Deniz, H., & Anderson, E. S. (2016). Evolution acceptance and epistemological beliefs of college biology students. *Journal of Research in Science Teaching*, *54*(4), 493–519. https://doi.org/10.1002/tea.21374

Brewer, C., & Smith, D. (2011). *Vision and change in undergraduate biology education: A call to action*. American Association for the Advancement of Science (AAAS).

Chen, H., Cohen, P., & Chen, S. (2010). How big is a big odds ratio? Interpreting the magnitudes of odds ratios in epidemiological studies. *Communications in Statistics - Simulation and Computation*, *39*(4):860–864. https://doi.org/10.1080/03610911003650383

Cohen, G. L., Garcia, J., Apfel, N., & Masters, A. (2006). Reducing the racial achievement gap: A social-psychological intervention. *Science*, *313*, 1307–1310.

Coleman, J. S., Campbell, E. Q., Hobson, C. J., McPartland, J., Weinfeld, F. D., & York, R. L. (1966). *Equality of educational opportunity*. U.S. Department of Health, Education, and Welfare.

Coley, J. D., & Tanner, K. (2017). Relations between intuitive biological thinking and biological misconceptions in biology majors and nonmajors. *CBE—Life Sciences Education*, *14*(1):1–19. https://doi.org/10.1187/cbe.14-06-0094

Cuban, L. (1999). *How scholars trumped teachers: Change without reform in university curriculum, research, and teaching, 1890–1990*. Teachers College Press.

Denaro, K., Dennin, K., Dennin, M., & Sato, B. (2022). Identifying systemic inequity in higher education and opportunities for improvement. *PLoS One*, *17*(4), e0264059. https://doi.org/10.1371/journal.pone.0264059

Dunk, R. D. P., Barnes, M. E., Reiss, M. J., Alters, B., Asghar, A., Carter, B. E., Cotner, S., Glaze, A. L., Hawley, P. H., Jensen, J. L., Mead, L. S., Nadelson, L. S., Nelson, C. E., Pobiner, B., Scott, E. C., Shtulman, A., Sinatra, G. M., Southerland, S. A., Walter, E. M., … Wiles, J. R. (2019). Evolution education is a complex landscape. *Nature Ecology and Evolution*, *3*(3), 327–329.

Evans, C. R., Williams, D. R., Onnela, J. P., & Subramanian, S. V. (2018). A multilevel approach to modeling health inequalities at the intersection of multiple social identities. *Social Science & Medicine (1982)*, *203*, 64–73. https://doi.org/10.1016/j.socscimed.2017.11.011

Graves, J. (2019). African Americans in evolutionary science: Where we have been, and what's next. *Evolution: Education and Outreach*, *12*(18):1–10. https://doi.org/10.1186/s12052-019-0110-5

Graves, J. L., Kearney, M., Barabinoc, G., & Malcomd, S. (2022). Inequality in science and the case for a new agenda. *Proceedings of the National Academy of Sciences of the United States of America*, *1*(10), 19. https://doi.org/10.1073/pnas.2117831119

Green, K., & Delgado, C. (2021). Crossing cultural borders: Results of an intervention on community college biology students' understanding and acceptance of evolution. *International Journal of Science Education*, *43*(4), 469–496. https://doi.org/10.1080/09500693.2020.1869854

Haak, D. C., HilleRisLambers, J., Pitre, E., & Freeman, S. (2011). Increased structure and active learning reduce the achievement gap in introductory biology. *Science*, *332*(6034), 1213–1216. https://doi.org/10.1126/science.1204820

Handelsman, J., Elgin, S., Estrada, M., Hays, S., Johnson, T., Miller, S., Mingo, V., Shaffer, C., & Williams, J. (2022). Achieving STEM diversity: Fix the classrooms. *Science*, *376*(6597), 1057–1059. https://doi.org/10.1126/science.abn9515

Harackiewicz, J. M., Canning, E. A., Tibbetts, Y., Giffen, C. J., Blair, S. S., Rouse, D. I., & Hyde, J. S. (2014). Closing the social class achievement gap for first-generation students in undergraduate biology. *Journal of Educational Psychology*, *106*(2), 375–389. https://doi.org/10.1037/a0034679

Harris, R. B., Mack, M. R., Bryant, J., Theobald, E. J., & Freeman, S. (2020). Reducing achievement gaps in undergraduate general chemistry could lift underrepresented students into a "hyperpersistent zone." *Science Advances*, *6*(24), 1–9. https://doi.org/10.1126/sciadv.aaz5687

Hatfield, N., Brown, N., & Topaz, C. M. (2022). Do introductory courses disproportionately drive minoritized students out of STEM pathways? *PNAS Nexus*, *1*(4), 1–10.

Jordt, H., Eddy, S. L., Brazil, R., Lau, I., Mann, C., Brownell, S. E., King, K., & Freeman, S. (2017). Values affirmation intervention reduces achievement gap between underrepresented minority and white students in introductory biology classes. *CBE—Life Sciences Education*, *16*(3), ar41. https://doi.org/10.1187/cbe.16-12-0351

Kalinowski, S. T., Leonard, M. J., & Taper, M. L. (2016). Development and validation of the conceptual assessment of natural selection (CANS). *CBE—Life Sciences Education*, *15*(4), ar64. https://doi.org/10.1187/cbe.15-06-0134

Keller, L., Lüdtke, O., Precke, F., & Brunner, M. (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.

Ladson-Billings, G. (2006). From the achievement gap to the education debt: Understanding achievement in U.S. schools. *Educational Researcher*, *35*(7), 3–12. https://doi.org/10.3102/0013189x035007003

Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for *t*-tests and ANOVAs. *Frontiers in Psychology*, *4*, 1–12 https://doi.org/10.3389/fpsyg.2013.00863

Levy, D. J., Heissel, J. A., Richeson, J. A., & Adam, E. K. (2016). Psychological and biological responses to race-based social stress as pathways to disparities in educational outcomes. *The American Psychologist*, *71*(6), 455–473. https://doi.org/10.1037/a0040322

Luft, J. A., Jeong, S., Idsardi, R., & Gardner, G. (2022). Literature reviews, theoretical frameworks, and conceptual Frameworks: An introduction for new biology education researchers. *CBE—Life Sciences Education*, *21*(3), rm33. https://doi.org/10.1187/cbe.21-05-0134

McCloskey, M. (1983). Intuitive physics. *Scientific American*, *248*(4), 122–130. https://doi.org/10.1038/scientificamerican0483-122

Mead, L. S., Clarke, J. B., Forcino, F., & Graves, J. L. (2015). Factors influencing minority student decisions to consider a career in evolutionary biology. *Evolution: Education and Outreach*, *8*(1), 1–11. https://doi.org/10.1186/s12052-015-0034-7

Merolla, D. M., & Jackson, O. (2019). Structural racism as the fundamental cause of the academic achievement gap. *Sociology Compass*, *13*(6), e12696.

Moharreri, K., Ha, M., & Nehm, R. H. (2014). EvoGrader: An online formative assessment tool for automatically evaluating written evolutionary explanations. *Evolution: Education and Outreach*, *7*(1), 1–14.

Myers, H. F. (2009). Ethnicity- and socio-economic status-related stresses in context: An integrative review and conceptual model. *Journal of Behavioral Medicine*, *32*(1), 9–19. https://doi.org/10.1007/s10865-008-9181-4

Nadelson, L. S., & Southerland, S. (2012). A more fine-grained measure of student's acceptance of evolution: Development of the inventory of student evolution acceptance–I-SEA. *International Journal of Science Education*, *34*(11), 1637–1666. https://doi.org/10.1080/09500693.2012.702235

Nardo, J. E., Chapman, N. C., Shi, E. Y., Wieman, C., & Salehi, S. (2022). Perspectives on active learning: Challenges for equitable active learning implementation. *Journal of Chemical Education*, *99*(4), 1691–1699. https://doi.org/10.1021/acs.jchemed.1c01233

National Research Council (NRC). (2019). *Monitoring educational equity*. National Academies Press.

National Science Foundation (NSF). (2022). *Survey of earned doctorates. Research doctorate recipients, by detailed field of doctorate, citizenship status, ethnicity, and race: 2022. Detailed Statistical Table 3-3*. National Center for Science and Engineering Statistics (NCSES). https://ncses.nsf.gov/pubs/nsf23300/data-tables

Nehm, R. H., Beggrow, E. P., Opfer, J. E., & Ha, M. (2012). Reasoning about natural selection: Diagnosing contextual competency using the ACORNS instrument. *The American Biology Teacher*, *74*(2), 92–98. https://doi.org/10.1525/abt.2012.74.2.6

Nehm, R. H., Finch, S., & Sbeglia, G. C. (2022). Is active learning enough? The contributions of misconception-focused instruction and active-learning dosage on student learning of evolution. *BioScience*, *72*(11), 1105–1117.

Nehm, R. H., & Reilly, L. (2007). Biology majors' knowledge and misconceptions of natural selection. *BioScience*, *57*(3), 263–272. https://doi.org/10.1641/b570311

Nehm, R. H., & Schonfeld, I. (2008). Measuring knowledge of natural selection: A comparison of the CINS, and open-response instrument, and oral interview. *Journal of Research in Science Teaching*, *1160*, 1131–1160.

Neuman, S. B., & Moland, N. (2019). Book deserts: The consequences of income segregation on children's access to print. *Urban Education*, *54*(1), 126–147. https://doi.org/10.1177/0042085916654525

NGSS Lead States. (2013). *Next generation science standards: For states by states* (Vol. 2). Achieve Inc., on behalf of the 26 States and Partners.

Nissen, J. M., Horses, I. H. M., & Van Dusen, B. (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.

O'Brien, L. T., Bart, H. L., & Garcia, D. M. (2020). Why are there so few ethnic minorities in ecology and evolutionary biology? Challenges to inclusion and the role of sense of belonging. *Social Psychology of Education*, *23*(2), 449–477. https://doi.org/10.1007/s11218-019-09538-x

Opfer, J. E., Nehm, R. H., & Ha, M. (2012). Cognitive foundations of science assessment design: Knowing what students know about evolution. *Journal of Research in Science Teaching*, *49*(6), 744–777. https://doi.org/10.1002/tea.21028

Orgill, M., & Sutherland, A. (2008). Undergraduate chemistry stuA dents' perceptions of and misconceptions. *Chemistry Education Research and Practice*, *9*, 131–143.

President's Council of Advisors on Science and Technology. (2012). *Engage to excel: Producing one million additional college graduates with degrees in science, technology, engineering and mathematics*. https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/pcast-engage-to-excel-final_2-25-12.pdf

Riegle-Crumb, C., King, B., & Irizarry, Y. (2019). Does STEM stand out? Examining racial/ethnic gaps in persistence across postsecondary fields. *Educational Researcher*, *48*(3), 133–144. https://doi.org/10.3102/0013189x19831006

Rushworth, C. A., Baucom, R. S., Blackman, B. K., Neiman, M., Orive, M. E., Sethuraman, A., Ware, J., & Matute, D. R. (2021). Who are we now? A demographic assessment of three evolution societies. *Evolution*, *75*(2), 208–218. https://doi.org/10.1111/evo.14168

Salehi, S., Berk, S. A., Brunelli, R., Cotner, S., Creech, C., Drake, A. G., Fagbodun, S., Hall, C., Hebert, S., Hewlett, J., James, A. C., Shuster, M., St. Juliana, J. R., Stovall, D. B., Whittington, R., Zhong, M., & Ballen, C. J. (2021). Context matters: Social psychological factors that underlie academic performance across seven institutions. *CBE—Life Sciences Education*, *20*(4), ar68. https://doi.org/10.1187/cbe.21-01-0012

Salehi, S., Burkholder, E., Lepage, G. P., Pollock, S., & Wieman, C. (2019). Demographic gaps or preparation gaps? The large impact of incoming preparation on performance of students in introductory physics. *Physical Review Physics Education Research*, *15*(2), 020114.

Salehi, S., Cotner, S., & Ballen, C. J. (2020). Variation in incoming aca- demic preparation: Consequences for minority and first-generation students. *Frontiers in Education*, *5*, 5. https://doi.org/10.3389/feduc.2020.552364

Sbeglia, G. C., Goodridge, J. A., Gordon, L. H., & Nehm, R. H. (2021). Are faculty changing? How reform frameworks, sampling intensities, and instrument measures impact inferences about student-centered teaching practices. *CBE—Life Sciences Education*, *20*(3), ar39. https://doi.org/10.1187/cbe.20-11-0259

Sbeglia, G. C., & Nehm, R. H. (2022). Measuring evolution learning: Impacts of student participation incentives and test timing. *Evolution: Education and Outreach*, *15*(1):1–15. https://doi.org/10.1186/s12052-022-00166-2

Shukla, S. Y., Theobald, E. J., Abraham, J. K., & Price, R. M. (2021). Reframing educational outcomes: Moving beyond achievement gaps. *CBE—Life Sciences Education*, *21*(2), 2. https://doi.org/10.1187/cbe.21-05-0130

Simmons, A. B., & Heckler, A. F. (2020). Grades, grade component weighting, and demographic disparities in introductory physics. *Physical Review Physics Education Research*, *16*(2), 16. https://doi.org/10.1103/PhysRevPhysEducRes.16.020125

Smith, M. K., Jones, F. H. M., Gilbert, S. L., & Wieman, C. E. (2013). The classroom observation protocol for undergraduate STEM (COPUS): A new instrument to characterize university STEM classroom practices. *CBE—Life Sciences Education*, *12*(4), 618–627. https://doi.org/10.1187/cbe.13-08-0154

Smith, M. U. (2010a). Current status of research in teaching and learning evolution: I. Philosophical/epistemological issues. *Science & Education*, *19*(6–8), 523–538. https://doi.org/10.1007/s11191-009-9215-5

Smith, M. U. (2010b). Current status of research in teaching and learneing evolution: II. Pedagogical issues. *Science & Education*, *19*(6–8), 539–571. https://doi.org/10.1007/s11191-009-9216-4

Society for the Study of Evolution (SSE). (2017). *Diversity statement*. https://www.evolutionsociety.org/content/diversity-statement.html

Tashiro, J., & Talanquer, V. (2021). Exploring inequities in a traditional and a reformed general chemistry course. *Journal of Chemical Education*, *98*(12), 3680–3692. https://doi.org/10.1021/acs.jchemed.1c00821

Theobald, E. J., Hill, M. J., Tran, E., Agrawal, S., Arroyo, E. N., Behling, S., Chambwe, N., Cintrón, D. L., Cooper, J. D., Dunster, G., Grummer, J. A., Hennessey, K., Hsiao, J., Iranon, N., Jones L., 2nd, Jordt, H., Keller, M., Lacey, M. E., Littlefield, C. E., ... Freeman, S. (2020). Active learning narrows achievement gaps for underrepresented students in undergraduate science, technology, engineering, and math. *Proceedings of the National Academy of Sciences of the United States of America*, *117*(12), 6476–6483.

Thiry, H., Weston, T., Harper, R. P., Holland, D. G., Koch, A. K., Drake, B. M., Hunter, A., & Seymour, E. (2019). In E. Seymour & A. Hunter (Eds.), *Talking about leaving revisited*. Springer.

Tseng, M., El-Sabaawi, R. W., Kantar, M. B., Pantel, J. H., Srivastava, D. S., & Ware, J. L. (2020). Strategies and support for Black, Indigenous, and people of colour in ecology and evolutionary biology. *Nature Ecology & Evolution*, *4*(10), 1288–1290. https://doi.org/10.1038/s41559-020-1252-0

Van Dusen, B., & Nissen, J. (2019). Associations between learning assistants, passing introductory physics, and equity: A QuantCrit investigation. *arXiv*, arXiv:1912.01533.

Van Dusen, B., Nissen, J., Talbot, R. M., Huvard, H., & Shultz, M. (2022). 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. https://doi.org/10.1021/acs.jchemed.1c00352

White, K., Haas, J. S., & Williams, D. R. (2012). Elucidating the role of place in health care disparities: The example of racial/ethnic residential segregation. *Health Services Research*, *47*(3 Pt 2), 1278–1299. https://doi.org/10.1111/j.1475-6773.2012.01410.x

Williams, D. R. (1999). Race, socioeconomic status, and health: The added effects of racism and discrimination. *Annals of the New York Academy of Sciences*, *896*(1), 173–188. https://doi.org/10.1111/j.1749-6632.1999.tb08114.x

Young, N., Rypkema, H., Bell, E., & Singer, S. (2022). *Leveraging institutional data to advance equity in STEM courses*. AAAS and NSF. https://www.aaas-iuse.org/leveraging-institutional-data/