

# A Longitudinal Study Identifying the Characteristics and Content Knowledge of Those Seeking Certification to Teach Secondary Biology in the United States

Kimberly Linenberger Cortes,<sup>†\*</sup> Joshua W. Reid,<sup>‡</sup> Rebekah Fallin,<sup>§</sup> Jie Hao,<sup>§</sup> Lisa Shah,<sup>||</sup> Herman E. Ray,<sup>§</sup> and Gregory T. Rushton<sup>‡</sup>

<sup>†</sup>Department of Chemistry and Biochemistry and <sup>§</sup>Analytics and Data Science Institute, Kennesaw State University, Kennesaw, GA 30144; <sup>‡</sup>Tennessee STEM Education Center, Middle Tennessee State University, Murfreesboro, TN 37132; <sup>||</sup>Department of Chemistry, Stony Brook University, Stony Brook, NY 11794

## ABSTRACT

Teacher content knowledge has been identified as a key prerequisite to effective instruction, and current educational policies require measurement of teacher content knowledge to assess candidacy for licensure. The primary instruments used in the United States are the Praxis Subject Assessment exams, which are designed to measure the subject-specific content knowledge needed to be a teacher. The Praxis Biology Subject Assessment exam, used by 42 U.S. states in the past decade, is the most common national measure used to determine biology content knowledge for teacher certification. Demographic and performance data from examinees ( $N = 43,798$ ) who took the Praxis Biology Subject Assessment from 2006 to 2015 were compared to present a much-needed picture of who is seeking certification to teach biology, how different groups of aspiring biology teachers have performed, and how demographic makeup of prospective biology teachers compares with reports in previous studies describing the composition of the biology teacher workforce. Results indicate the majority of students self-reported as White (76%), female (66%), having undergraduate grade point averages (GPAs) at or above a 3.0 (76%) and majoring in biology (45%). Additionally, the demographic data were included in a linear regression model to determine the factors that explained the most variance in performance of the examinee. The model revealed substantial differences in average performance and pass rates between examinees of different genders, races, undergraduate majors, undergraduate GPAs, and census regions. This suggests that if the examinee is a White science, technology, engineering, and mathematics major, man with a 3.5 or higher undergraduate GPA, resides in the western United States, or plans to teach in a suburban school, the examinee will on average outperform their counterparts on the exam. From our analyses, we suggest several measures for the improvement of the biology teaching workforce and establish potential issues in the teacher pipeline that may impact the quality and diversity of U.S. biology teachers.

## INTRODUCTION

One of the largest contributions to the success of students in the secondary classroom is their teacher. Individual teachers and their qualifications (e.g., certification, degree in field teaching) have been shown to have significant impact on students' success in science, technology, engineering, and mathematics (STEM; Fetler, 1999; Darling-Hammond, 2000; Clotfelter *et al.*, 2007). These findings have provided evidence for policy changes that emphasize increased subject-specific course work in disciplinary fundamentals and in-field certifications to ensure a teacher is adequately prepared to teach a specific subject (Shulman, 1987; Veal *et al.*, 1999; Sperandeo-Mineo *et al.*, 2006).

Stephanie Gardner, *Monitoring Editor*

Submitted Aug 31, 2021; Revised Apr 28, 2022;  
Accepted Jul 26, 2022

CBE Life Sci Educ December 1, 2022 21:ar63  
DOI:10.1187/cbe.21-08-0220

\*Address correspondence to: Kimberly Linenberger Cortes (klinienbe@kennesaw.edu).

© 2022 K. L. Cortes *et al.* CBE—Life Sciences Education © 2022 The American Society for Cell Biology. This article is distributed by The American Society for Cell Biology under license from the author(s). It is available to the public under an Attribution–Noncommercial–Share Alike 4.0 Unported Creative Commons License (<http://creativecommons.org/licenses/by-nc-sa/4.0>).

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

Several large-scale science education reform efforts, such as the Next Generation Science Standards (NGSS Lead States, 2013), have emerged over the last two decades to change classroom practice to be more research-based.

Recently, national surveys have been used to identify the personal and professional characteristics of the biology teacher workforce. For instance, biology is the most popular science course relative to enrollment in the United States (Banilower *et al.*, 2013) and has a teacher workforce greater than that of chemistry, physics, and earth science combined (Hill, 2011). Biology teachers are predominantly women and traditionally have been far more prepared to teach biology compared with other science teachers in their respective subjects, with 54% of biology teachers having a degree in their discipline compared with only 22% of other science teachers (Lyons, 2013). Additionally, nearly all (98%) biology teachers report previously taking an introductory biology course during their undergraduate degree program, with an additional 70% having taken a course in more advanced biology such as genetics, anatomy/physiology, and cell biology as of 2012 (Lyons, 2013; Polizzi *et al.*, 2015). This suggests that those teaching biology have strong backgrounds in the discipline, which should therefore be positively correlated with student achievement (Clotfelter *et al.*, 2010). However, achievement gaps have still been identified in these courses (Eddy *et al.*, 2014) and factors associated with these gaps should be investigated further.

### Teacher Knowledge

Osborne *et al.*, (2003) stated that “For the research evidence shows clearly that it is the teacher variables that are the most significant factor determining attitude, not curriculum variables” (p. 1070). One teacher variable, teacher knowledge, has been extensively studied and shown to be composed of several domains of knowledge to effectively facilitate the learning of their students (Evens *et al.*, 2018; Gess-Newsome, 2015). Shulman (1987) subcategorized teacher knowledge into: 1) content knowledge (CK), 2) pedagogical content knowledge (PCK), and 3) curricular knowledge.

The Next Generation Science Standards (NGSS; NGSS Lead States, 2013) have spurred shifts in teaching practices in science classrooms from the memorization of facts to developing a “rich network of connected ideas that serve as a conceptual tool for explaining phenomena, solving problems and making decisions” (Krajcik *et al.*, 2014, p. 159). The NGSS provides a framework for teaching science that includes disciplinary core ideas, science and engineering practices, and crosscutting ideas (NGSS Lead States, 2013). The inclusion of science and engineering practices in the NGSS suggests teachers must be knowledgeable about not only *what* but also *how* knowledge is generated. Additionally, crosscutting ideas can lend themselves to an interdisciplinary curriculum across multiple disciplines. Barbara Nagle (2013) wrote about the challenges and affordances of teaching interdisciplinary biology—teacher preparation was a key challenge. Secondary teachers are often trained in a disciplinary area and might not have the knowledge or preparedness to teach interdisciplinary ideas in biology. As such, in regard to CK, Shulman suggests that secondary biology teachers should be as knowledgeable as a biology undergraduate major, given that they have to teach students not

only *what* the knowledge of the discipline is but also *how* biological knowledge is generated and *why* understanding biology is important.

Gess-Newsome (2015) extended current models of PCK by proposing six dynamic levels: teacher professional knowledge base, topic-specific professional knowledge, teacher amplifiers and filters, classroom practice, student amplifiers and filters, and student outcomes. The author differentiated the professional knowledge base level from topic specific by stating that topic-specific professional knowledge (TSPK) is “recognized ... by experts and is available for study and use by teachers ... [and] can be recorded in text, diagrams, or tables” (p. 33). This is similar to the mathematical knowledge for teaching model and the specialized content knowledge model presented by Ball and colleagues (Ball *et al.*, 2005, 2008). TSPK represents a deep understanding of the disciplinary content to be taught and the multiple representations that accompany the discipline (i.e., Airey and Linder, 2009). Surprisingly little research has considered the overall biology content knowledge of those seeking certification to teach biology (Carlsen, 1993; Gess-Newsome and Lederman, 1993). Studies have primarily focused on specific topics such as evolution or osmosis, seeking to identify misconceptions that may then be propagated to students (Yip, 1998; Rutledge and Mitchell, 2002; Hakan *et al.*, 2013).

Most states require certification exams as a way to measure a teaching candidate’s subject-specific content knowledge as a requirement for teaching licensure. The Praxis Biology Subject Assessment, administered by the Educational Testing Service (ETS) and used in 42 states over the last decade, is a 150-question multiple-choice exam “designed to measure the knowledge and competencies necessary” for beginning high school biology teachers (ETS, 2014). Items on the exam are intended to measure fundamental knowledge of cellular and molecular biology, genetics, evolution, organismal biology, ecology, and environments. Publicly available exam items are typically written at a level routinely found in introductory college biology courses (ETS, 2014).

### Achievement and Opportunity Gaps

Achievement gaps are reflections of the systemic inequities across educational levels from elementary to graduate education (Milner, 2012). Differences in student performance can be somewhat attributed to disparaging sociopolitical and economic disparities in American education systems (Carter and Welner, 2013). The current educational system privileges some groups with more significant opportunities than others. The prevalence of these gaps have led scholars to shift from achievement-based discourse to opportunity-based discourse and from achievement gaps to opportunity gaps (Milner, 2012, 2017). By considering opportunity gaps, educators and scholars can reflect on systemic changes that need to occur to ensure a more equitable education system.

In science, opportunity gaps related to gender or race have been thoroughly noted in the literature as starting as early as elementary school and persisting through secondary and post-secondary schools (Bacharach *et al.*, 2003; Eddy *et al.*, 2014). In undergraduate biology, gender gaps are seen when comparing student performance on exams and how often students talk, with women being less likely to engage in discourse in biology

classrooms (Eddy *et al.*, 2014), even though women account for more than half of undergraduate biology degrees (Luckenbill-Edds, 2002). Gender and racial gaps are also seen in teachers rather than just students. For instance, Shah *et al.* (2018b) noted both gender and racial performance gaps with regard to examinees taking the Praxis Chemistry Subject Assessment. That is, men outperform women, and White and Asian examinees outperform Black and Hispanic examinees (Shah *et al.*, 2018b). Several strategies for combating these achievement gaps have been identified in the literature, including active-learning environments (Theobald *et al.*, 2020). However, the uptake and implementation of these practices have been slow in undergraduate STEM courses (Stains *et al.*, 2018), perpetuating the opportunity and achievement gaps.

While biology teachers are primarily women and the majority have in-field degrees, we wondered whether the pipeline to producing highly qualified biology teachers was suffering based on the demographics of prospective and current teachers taking the Praxis Biology Subject Assessment. Differences in performance among subgroups of test takers of the Praxis Biology Subject Assessment could be partially explained by achievement gaps and, relatedly, opportunity gaps in K–16 education. Additionally, some variation in test performance could be explained by teacher recruitment strategies. For instance, gender gaps in teachers could be seen if only the lowest-performing women pursue teaching as a career. In this study, we focus on the demographics of the examinees in an effort to highlight achievement gaps found in the Praxis Biology Subject Assessment. The study described herein seeks to build upon these previous works to connect the personal and professional characteristics of those seeking certification to teach biology to the totality of their understanding of basic biology concepts typically taught in introductory biology.

### Research Questions and Rationale

The goal of the presented work was to understand who has intended to teach high school biology in the United States from 2006 to 2016 and how different personal and professional characteristics of candidates are associated with certification exam performance. We report findings from our analysis of Praxis Biology Subject Assessment data for all test takers from 2006 to 2016 to answer the following research questions:

1. What were the personal and professional characteristics of those intending to teach high school biology from 2006 to 2016?
2. How did these personal and professional characteristics correlate with Praxis Biology Subject Assessment performance from 2006 to 2016?

## METHODS

### Data Sources and Standards for Passing

The Praxis Subject Assessments are a series of exams used to measure the subject-specific content knowledge of beginning K–12 teaching candidates. A large percentage of those pursuing a career as a biology teacher take the Praxis Biology Subject Assessment as part of the certification process. In 2016, 70% of U.S. states required the Praxis Biology Subject Assessment as part of certification (see Supplemental Material). Exam questions are prepared by panels of expert educators, teacher prepara-

tion faculty, and subject specialists and subsequently reviewed by ETS for validity, reliability, and issues of bias before full implementation (ETS, 2015). Passing is determined at the state level, after ETS conducts a multistate standard-setting study. Content experts and current/former biology teachers from across the country evaluate the probability of a “just qualified” beginning biology teacher correctly responding to an exam item. After multiple rounds of discussion, these judgments are summed and averaged to produce a final recommended passing score. Individual state’s standard-setting committees are provided with this information, which is considered when determining that state’s yearly passing score (i.e., “cut” score; ETS, 2015). To control for the difficulty of different editions of the exam, ETS converts the raw total score to a scaled score between 100 and 200. Examinees who earn a scaled score at or above the state’s cut score are considered to have passed the exam. Following the precedence of Gitomer *et al.* (2011), we used publicly available data on individual state passing scores to establish a common national passing standard. Using the median cut score across all states using the Praxis Biology Subject Assessment between June 2006 and May 2016 (i.e., academic years 2006–2015; see Supplemental Table S1), a scaled score of 150 was used as the median national cut score for all analyses in this study. To approximate the corresponding percent correct, we determined the median percent correct scores of all examinees who earned an exact scale score of 150 during the analyzed time frame, which resulted in an estimated corresponding percent correct of 56%.

### Study Sample

Human subjects exemption was granted by the Institutional Review Boards at each principal investigator’s institution before the start of the data analysis. The study sample includes all examinees of the Praxis Biology Subject Assessment during the analyzed time frame (AY 2006–2015)<sup>1</sup> who reported an age between 18 and 75 ( $N = 43,798$ ). Restricting the age range eliminated outliers who may have misreported their date of birth. Demographic information was obtained from the examinees’ self-reported responses to survey questions asking about their personal and professional backgrounds (e.g., gender, race/ethnicity, undergraduate major, undergraduate GPA) at the beginning of the exam. Details about demographic characteristics analyzed in this work are listed in Supplemental Table S2. Regarding gender, we wish to note that the exam asks individuals to self-report their gender using male/female options. We recognize these are terms associated with biological sex rather than gender. As such, we have decided to use this language in referencing our results, to keep consistent with reporting of the Praxis Biology Subject Assessment. When not discussing results directly, we use more gender-appropriate terminology, such as man and woman. The Praxis Biology Subject Assessment does not provide a nonbinary or other option for respondents. To protect the privacy of the individuals, geographic information is presented at the level used by the U.S. Census,

<sup>1</sup>ETS convened an National Advisory Committee and conducted a job survey, and revised specifications went into effect in 2014; however, the changes were small enough that the new version of the test was able to be equated to the old version, avoiding rescaling and allowing equated scaled scores from the old and new versions to be comparable.

determined by the state postal codes of the examinee's home address. Because some examinees repeated the exam, only the highest score was included in our analyses to ensure a single score for each individual (Nettles et al., 2011). When performing demographic-specific analyses on the sample (e.g., impact of gender on exam performance), only those who reported on these specific prompts were included in the analysis. As a result, each analysis conducted for the results displayed in this work includes slightly different populations. Given that the entire test-taking population from June 2006 to May 2016 is represented in our analysis, the use of statistical significance parameters (e.g.,  $p$  values) were not present in our analyses. This is because  $p$  values are generally used to quantify differences between samples within a population; however, our study used data from the entire population.

### Model Selection

As a result of an apparently normal distribution of percentage of correct responses, a linear regression model was used to uncover characteristics of examinees most strongly associated with performance for examinees in each academic year. Given the 14 demographic attributes of interest including race/ethnicity, gender, and undergraduate major, as well as the corresponding two-way and three-way interactions between them, a total of 469 variable interactions were possible. A stepwise linear regression technique was used to determine the best linear model from the set of candidate independent variables using the Schwarz Bayesian criterion (SBC; Schwarz, 1978) for variable selection and 10-fold cross-validation (CV) as the stopping criterion. The SBC was selected as it tends to suggest simpler models with lower dimensionality than other selection criteria (Akaike, 1974), whereas the 10-fold CV is used to reduce bias caused by the variable selection technique. The SBC is defined to be:

$$\text{SBC} = n * \ln(\text{SSE}/n) + p * \ln(n)$$

where SSE is the sum of squared errors,  $n$  is the sample size, and  $p$  is the number of parameters included in the model. The stepwise procedure is able to yield a single optimal model based on the specified criterion.

The aforementioned process was conducted to identify which personal and professional characteristics were most strongly correlated with Praxis Biology Subject Assessment performance in each year of 2006–2016. Characteristics or interactions (combinations of characteristics) that were identified in 50% or more of years were then used to build an aggregate model that identifies the characteristics most strongly correlated with performance on the exam across all

years 2006–2016 explaining 29% of the variation in scores (total  $\eta^2$ ; Table 1). These variables were chosen because additional variables provided little additional explanation of the variance. The total  $\eta^2$  for each model offers an indication of the partial contribution of each factor to the overall variance. A more detailed analysis of these variables/interactions was performed and is described in the following sections (see Supplemental Table S5).

### Differential Item Functioning Analysis

To ensure differences seen in performance reflected the differences seen in the population and not necessarily built into the assessment, differential item functioning (DIF) analysis was performed following the guidelines set forth by ETS (Zieky, 2003; Zwick, 2012). To summarize this procedure, for each test form administered between 2006 and 2016, examinees were matched based on the percent correct calculated after removing questions that show a difference in the odds of getting the response correct based on race or gender. As a result, quartiles of similarly performing examinees were produced. A logistic regression model was used to estimate the relative likelihood of the focal group (e.g., females, Blacks, Hispanics, other) answering a question correctly as compared with the reference group (males, Whites) within each performance quartile. The MH D-DIF statistic (Zieky, 2003; Zwick, 2012) was then calculated for each item on each exam form using the equation  $-2.35 \ln \hat{\alpha}_{.MH}$ , where  $\hat{\alpha}_{.MH}$  represents the odds ratio of performance for the focal and reference group. For example, an MH D-DIF value of  $-1.1$  for a specific item in regard to race would provide an  $\hat{\alpha}_{.MH}$  value of 1.597, indicating that White examinees (“reference group”) would have a 60% higher chance of answering the item correctly than a Black examinee (“focal group”) when performance on the exam as a whole is matched via quartile performance. To determine the degree of bias in the items on the respective forms, the MH D-DIF statistic was used to classify the items into one of three levels: A, B, or C, where 1) A-level items had absolute values of MH D-DIF statistics that were less than 1.0, indicating small differences in performance between groups; 2) C-level items had absolute values of MH D-DIF statistics that were statistically greater than 1.5, indicating large differences in performance between groups; and 3) all remaining items with MH D-DIF statistics between levels A and C were classified as B-level items. The results of these analyses are reported in Supplemental Tables S3 and S4.

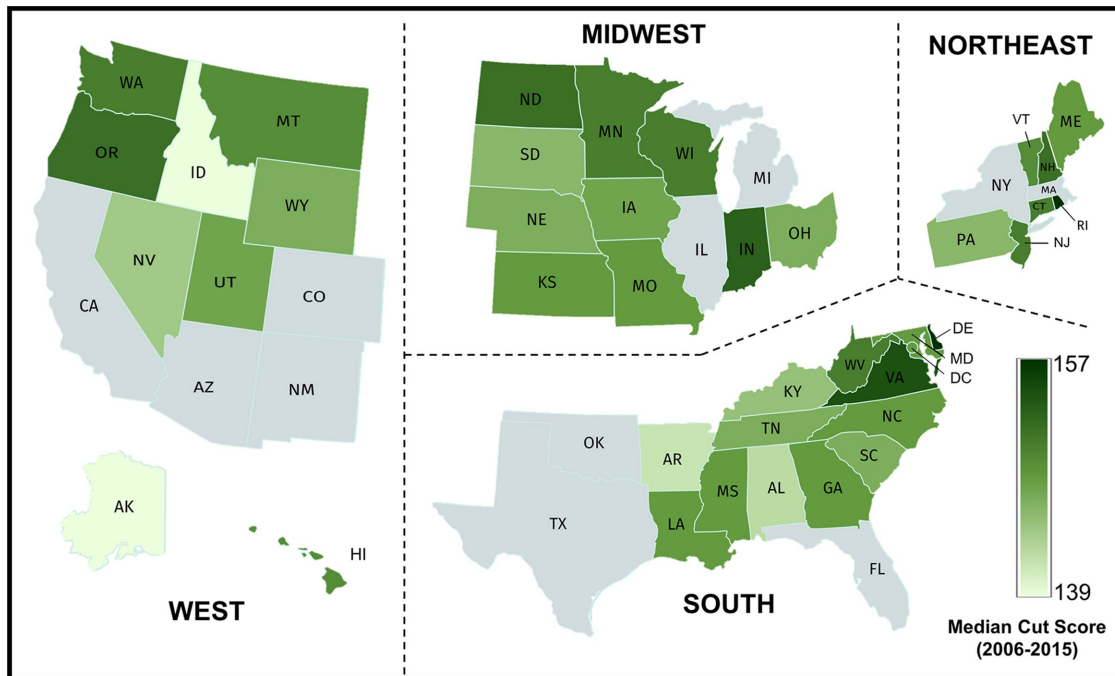
## RESULTS

1. What were the personal and professional characteristics of those intending to teach high school biology from 2006 to 2016?

**TABLE 1. Stepwise linear regression models including top examinee characteristics most strongly associated with performance on the Praxis Biology Subject Assessment from 2006 to 2016**

Variables	Cumulative $\eta^2$
G_UNDERMAJO*G_ETHNIC	0.19
G_UNDERMAJO*G_ETHNIC GENDER*UGPA	0.24
G_UNDERMAJO*G_ETHNIC GENDER*UGPA EDUCATIONLEVEL	0.26
G_UNDERMAJO*G_ETHNIC GENDER*UGPA EDUCATIONLEVEL REGION	0.28
G_UNDERMAJO*G_ETHNIC GENDER*UGPA EDUCATIONLEVEL REGION GEOGRAPHICAREATOTEAC	0.29





**FIGURE 1.** Praxis Biology Subject Assessment cut scores from 2006 to 2016. Individual data reported by each state were used to determine the minimum scaled score needed to be awarded certification in the state. The median of these values for each state between 2006 and 2016 is depicted. States were assigned a score if they accepted Praxis Biology Subject Assessment testing for certification in any year within this time frame. Those that did not are shaded gray. A scaled score of 150 corresponds to a raw percentage of approximately 56%. Source: Derived from data provided by ETS.

A specific breakdown of examinee personal and professional characteristics for those seeking certification to teach high school biology can be found in Supplemental Table S2. Two-thirds of examinees self-identified as female, which is significantly higher than the ratios of females taking the Praxis Chemistry and Physics Subject Assessments (58% and 37%, respectively; Shah *et al.*, 2018a,b). The overwhelming majority of students self-reported to be White (76%) and have undergraduate GPAs at or above a 3.0 (76%). Approximately, 45% of examinees were biology majors, with another 10% of examinees coming from other STEM disciplines or STEM education. Almost half (49%) of the examinees reported earning at most a bachelor's degree, with 46% of examinees reporting currently attending college. For those with a graduate major, the disciplinary trend holds, with 46% reporting a graduate major in biology or another STEM or STEM education discipline. There is relatively equal representation from the four age groups sampled (18–23, 24–28, 29–38, and 39–75). The majority of examinees (52%) were currently enrolled in or planning to enroll in a teacher preparation program with relatively equal participation in undergraduate (22%), master's degree (21%), and alternative route (22%) preparation programs. The majority (43%) of examinees took the exam in the southern United States, and 24% of examinees reported plans to teach in a suburban school district.

2. How did these personal and professional characteristics correlate with Praxis Biology Subject Assessment performance from 2006 to 2016?

### State-Level Median "Cut" Scores

Figure 1 depicts the median scaled score required for an examinee to be considered for a biology teaching certificate in each state. The cut scores from all states accepting the Praxis Biology Subject Assessment over the last decade range from a low of 139 to a high of 157 on a scale of 100 to 200 (which correspond to approximate percent correct values of 47% and 62%, respectively). The national median cut score during the analyzed time frame was 150 (estimated at 55.94%) and will serve as the standard for passing in all analyses presented.

The results from the stepwise regression indicated that that model was statistically significant,  $F(64, 21,165) = 135.52$ ,  $p < 0.001$ , and explained 29% of the variance in the scaled scores ( $\eta^2 = 0.29$ ). Two statistically significant interaction effects were identified in the model. First, a student's undergraduate major and ethnicity produced an interaction that explained 19% ( $\eta^2 = 0.19$ ) of the variance in the model. This interaction indicates that, while undergraduate major predicts an individual's score on the exam, this is dependent on the individual's ethnicity. Furthermore, a statistically significant interaction was identified between gender and GPA, which explained 5% of the variation ( $\eta^2 = 0.05$ ).

### Demographics

The personal and professional characteristics of Praxis Biology Subject Assessment test takers most closely correlated with exam performance are shown in Table 2, along with additional characteristics that may be of interest. A total of 43,798 examinees took the exam between academic years 2006 and 2016, with 82.52% of examinees meeting or exceeding the previously

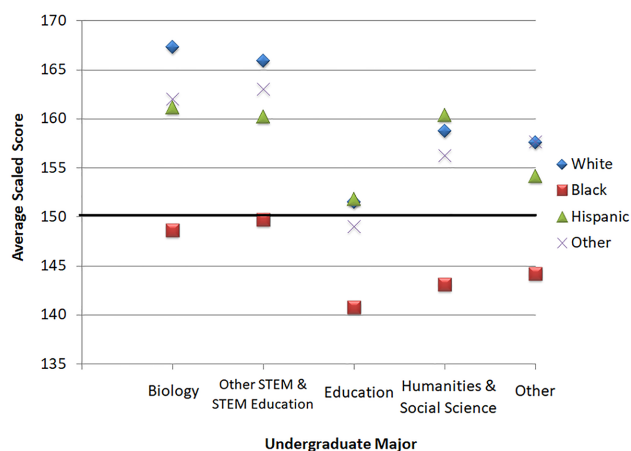
**TABLE 2. Personal and professional characteristics of Praxis Biology Subject Assessment examinees from 2006 to 2016**

Gender	N	% Pass
Female	28,875	80.88
Male	14,696	85.67
<b>Ethnicity</b>		
White	33,268	86.4
Black	4058	50.91
Hispanic	992	80.54
Other	3300	78.82
<b>Undergraduate major</b>		
Biology	19,703	86.57
Other STEM & STEM education	4577	85.86
Education	2432	57.03
Humanities and social sciences	3052	71.63
Other	3543	71.15
<b>Undergraduate GPA</b>		
3.5–4.0	15,447	87.95
3.0–3.49	17,652	83.34
2.5–2.99	7916	71.48
Below 2.5	995	62.01
<b>Educational level</b>		
Current undergraduate	8997	88.92
Earned bachelor's	21,536	79.05
Earned master's	8121	82.44
Earned PhD	1435	92.33
<b>Reported (planned) geographic areas where examinees plan to teach</b>		
Urban	9723	80.13
Rural	8702	77.63
Suburban	10,670	88.11
No plan to teach	2025	90.17
<b>Region</b>		
Northeast	11,933	88.21
Midwest	8615	90.38
South	18,782	73.32
West	4286	91.23
Overall (Total sample and the total sample pass rate)	43,798	82.52

discussed passing score of 150. While females passed the Praxis Biology Subject Assessment at a greater rate (81%) than females taking the Praxis Chemistry and Physics Subject Assessments (68% vs. 60%, respectively), they still on average passed at a lower rate than males. There is little difference in the percentage of passing between the various STEM degrees. While 24% of examinees come from the southern United States, these examinees have the lowest pass rate (73%) compared with the other census regions, and will be considered in more detail later.

### Undergraduate Major and Ethnicity

The interaction between a prospective biology teacher's undergraduate major and ethnicity explains 19% of the variance in performance on the Praxis Biology Subject Assessment (Table 1). There is little difference in average performance (at most 6 points on the total scaled scale) between White test takers and those who identify as Hispanic or as another ethnicity, regard-

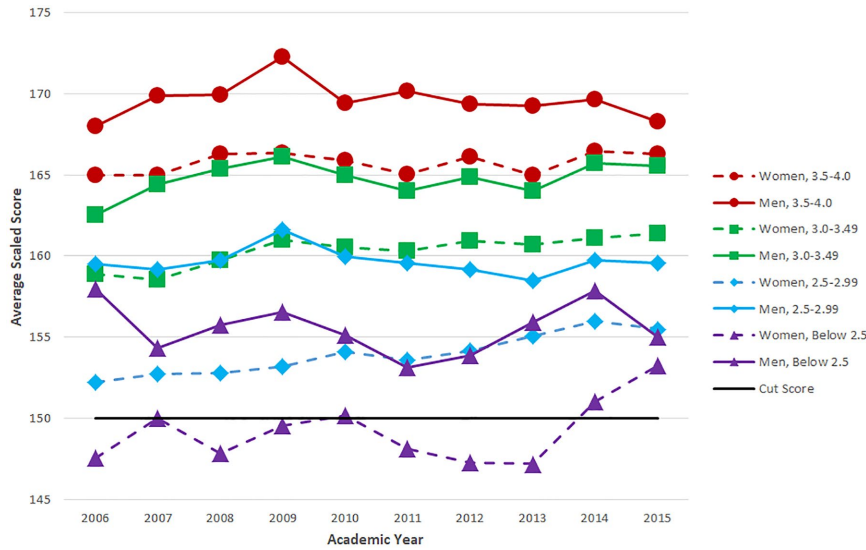


**FIGURE 2. Praxis Biology Subject Assessment performance by examinees' reported undergraduate major and ethnicity. The black line represents the national median passing score of 150. Source: Derived from data provided by ETS.**

less of major, as indicated in Figure 2. The disparity in performance can be seen when looking at the performance of examinees who identify as Black. The average performance for these examinees was below 150 regardless of major over the time period under consideration. This suggests that, on average, examinees who identified as Black were less likely to pass the exam and hence be able to become a biology teacher. Figure 2 also shows consistent performance between biology majors and all other STEM and STEM education majors. Humanities and social science majors and those in other non-STEM majors make up 15% of examinees and tend to score approximately 7 scaled points lower than their STEM colleagues. Regardless of race, education majors taking the Praxis Biology Subject Assessment underperform compared with all other undergraduate majors.

### Gender and GPA

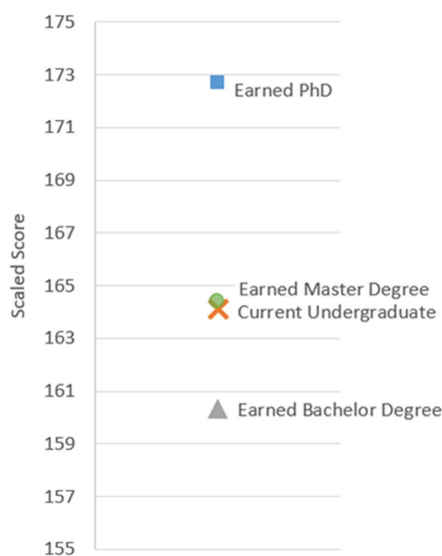
An examinee's undergraduate GPA and gender is correlated with performance on the Praxis Biology Subject Assessment (Figure 3) and accounts for 5% of the variation in performance (Table 1). Similar to what was found for prospective chemistry teachers' performance on the Praxis Chemistry Subject Assessment, within each gender, examinees with high GPAs consistently outperform those with lower GPAs (Shah et al., 2018b). The range of scores remained relatively consistent across the GPA groups for males, as the averages shifted downward as the GPA decreased. This is in comparison to the female population, where the ranges tended to shift down along with the averages. In looking at the data over time in Figure 4, the average score for males at each GPA is higher compared with the average score for females with the same GPA by an average of 4–6 scaled points across 2006–2015. Until 2014, females with GPAs below 2.5 were the only examinees in this set of characteristics who on average consistently scored just below the standard cut score. This is also the only group of females who have seemed to significantly narrow the achievement gap between themselves and their male counterparts over the time span analyzed.



**FIGURE 3.** Praxis Biology Subject Assessment performance by examinees’ reported gender and undergraduate GPA. Scaled score is plotted against academic year (i.e., academic year 2006 = June 2006–May 2007). The black line represents the national median passing score of 150. Source: Derived from data provided by ETS.

**Educational Level**

In general, the highest educational level attained correlates with performance on the Praxis Biology Subject Assessment (Figure 4). Current undergraduates represent the exception, as they perform similarly to those with earned master’s degrees on average. This exception may be a result of current undergraduates having recently learned the material in contrast to other examinees who already earned a bachelor’s degree. As a result of their lower average score, examinees with bachelor’s degrees exhibited a 10% lower passing rate compared with their counterparts currently enrolled in a bachelor’s degree program (Table 1).



**FIGURE 4.** Praxis Biology Subject Assessment performance by examinees’ reported educational level. Source: Derived from data provided by ETS.

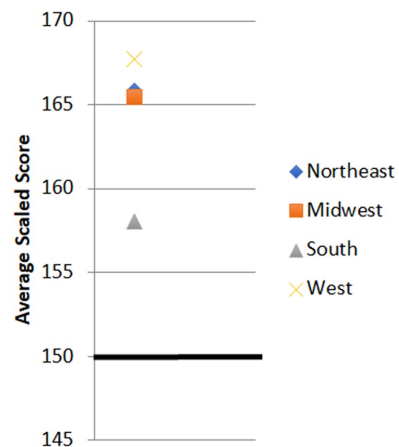
**Region**

Although the region of the United States that the examinee resides in only explains 2% of the variance in performance on the Praxis Biology Subject Assessment (Table 1), the results of exploring this characteristic warrant further explanation. There is a 10-point gap between the average scores of examinees living in the western United States and those living in the South (Figure 5), resulting in an 18% difference in the rate of passing the exam for those living in these two regions of the United States (Table 2). This difference may be explained by the fact that the vast majority (78%) of Black test takers come from the South compared with only 29% of Hispanic and 40% of White and other ethnicity examinees living in the South. However, it is important to contextualize this finding, given that Black test takers may be more likely to receive lower-quality K–16 instruction compared with White test takers (due to socioeconomic differences, access to educational resources, etc.).

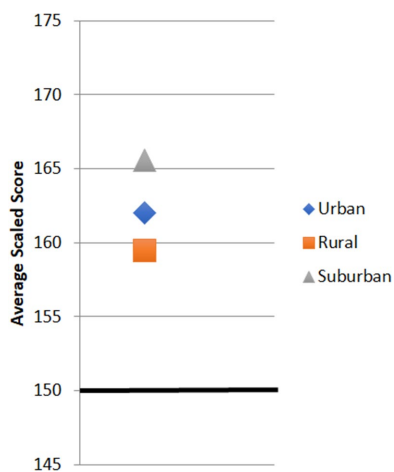
There is little difference in performance or passing rate for those living in the Northeast versus the Midwest. Note though, that there are major population centers in each region that are not included in this analysis (as shown in Figure 1) due to these areas not using the Praxis Biology Subject Assessment as means for certification.

**Geographic Area to Teach In**

The final characteristic included in our model is the geographic area examinees intended to teach in during the year following their exam administration. Despite this question only beginning to appear on the demographic questionnaire in 2008, the results are telling in regard to performance. Of the 93.5% of examinees



**FIGURE 5.** Praxis Biology Subject Assessment performance by the region of the United States where the examinee resides. The black line represents the national median passing score of 150. Source: Derived from data provided by ETS.



**FIGURE 6.** Praxis Biology Subject Assessment performance by reported (planned) geographic areas where examinees plan to teach. The black line represents the national median passing score of 150. Source: Derived from data provided by ETS.

planning to teach in the year following their examination, the 37% planning on teaching in a suburban school district consistently outperformed those planning on teaching in an urban or rural setting by 4 and 7 points, respectively (Figure 6). Although the average scores for all subgroups are above the cut score, rural communities seem to be attracting the lower-performing candidates consistently from 2006 to 2015. This trend is true for both White and Black examinees planning to teach in rural areas, with average scores being 5–10 scaled points lower than those planning to teach in the other communities. Hispanic examinees had little differences in averages across the communities of interest.

## DISCUSSION

Our analysis of the personal and professional characteristics of Praxis Biology Subject Assessment examinees over the past decade reveals much information regarding the demographic makeup of prospective high school biology teachers and their performance on this specific exam. Over the last decade, examinees with performance as low as 139, or approximately 49% correct, have been certified to teach biology at the secondary level. Demographic analysis shows that the percentage of women seeking certification to teach biology is consistent with the proportion of women teaching biology in the United States (Lyons, 2013; Polizzi et al., 2015) and those seeking a degree in biology in the United States (Eddy et al., 2014). When we look at the ethnic makeup of those seeking certification, minority groups continue to be underrepresented, with only 10% of the sample identifying as Black and 2% of the sample identifying as Hispanic. While this severely underrepresents the 17% of Hispanics in the U.S. population, the percentage of Black examinees nearly mirrors the U.S. population of 13% (U.S. Census Bureau, 2020). However, the average for Black examinees is consistently lower compared with other ethnicities, regardless of undergraduate major. The combination of passing rate and those sitting for the exam, taken together, may provide possible explanations for the lack of diversity in those teaching biology in the United States (Polizzi et al., 2015).

While biology majors do have the highest passing rate on the exam (87%), other STEM majors have performed nearly as well. A disparity in performance is also seen by gender when controlling for prior academic performance; on average, males tend to outperform their female counterparts by 4–6 scaled points. Not surprisingly, examinees with more advanced degrees performed better on the exam. While this study was cross-sectional rather than longitudinal in nature, our evidence does suggest a lack of retention of knowledge, with those having only an earned bachelor's degree showing the lowest performance. Finally, those living in the South and planning to teach in rural districts tend to be the least prepared in regard to biological content understanding, further exacerbating the inequities found within these regions. For instance, rural areas are often regions with low socioeconomic status and large numbers of underrepresented minority groups (i.e., Black, Hispanic). Placing the lowest-performing teachers in these schools widens the opportunity gap for students.

These results have several implications for future research, policy decisions, and teacher preparation. While publicly available exam questions seem to cover topics consistent with a college level introductory general biology course, even the highest-achieving students (GPA 3.5–4.0) and biology majors average about 160–167 (Figures 2 and 3). In addition, while the freshness of the information for current undergraduates may be the explanation for the similar performance as those with an earned master's degree, the change in performance shown by those with an earned bachelor's degree is troublesome. This drop in performance should provide additional motivation for biology faculty at the university level to enact research-based education reform initiatives, such as those outlined in *Vision and Change* (American Association for the Advancement of Science [AAAS], 2011) and in subsequent research initiatives (AAAS, 2015; Banger and Brownell, 2014).

Given an ever-increasing demand for teachers (Sutcher et al., 2016), the fact that biology is the most frequently taught science subject in the United States (Lyons, 2013), and a lack of supply of STEM teachers, (Sutcher et al., 2016), it is not surprising that the bar for passing the Praxis Biology Subject Assessment has been set by some states as low as 139. In addition, attrition of teachers is the leading factor in the current teacher shortage, and one of the leading causes of this attrition is lack of preparation, including understanding the content well enough to teach it (Ingersoll et al., 2014; Sutcher et al., 2016). Interestingly, in this study, we found test takers from the western United States outperformed those from the southern United States by 10-points on the scale. Taken as a whole, this finding suggests the lack of content knowledge preparation, as evidenced by performance on the Praxis Biology Subject Assessment, may add to the attrition of teachers in these populations. This could be explained by the recruitment of teachers into teacher education programs. If only low-performing students (e.g., women) pursue teaching-related careers, then these teachers might also be more vulnerable to attrition due to lack of understanding of the content. Therefore, there is a great need to help provide content support for teachers, especially in high turnover areas (e.g., the South, rural communities). This need for content knowledge support may also be present for biology majors who earned their degree in a specific field of biology and may have difficulties with biology subject matter outside their specific areas of expertise.



Finally, the gender and ethnicity achievement gaps evidenced in these data provide additional areas for improvement. The 4–6 point difference in performance in favor of male examinees is similar to previous reports of a gender gap found in introductory biology (2.8%; Eddy *et al.*, 2014) and a 400-level biology course (3.5%; Creech and Sweeder, 2012) when academic ability is controlled. This correlation suggests that the small gender gap that is evident at the start of a biology degree continues unchanged throughout the pursuit and acquisition of the degree. Research-based instructional strategies have been evidenced to help narrow achievement gaps related to gender and race (Theobald *et al.*, 2020). However, undergraduate courses and instructors have been slow to adopt these practices (Stains *et al.*, 2018), perpetuating the gaps. However, it is important to note that we did not collect or analyze data related to the instruction examinees received. Further research should consider controlling for the type of instruction each examinee receives in relation to individual performance on the Praxis exam. Additionally, research is warranted to determine whether this gap seen in women seeking certification to teach biology is evident in regard to self-efficacy and thus transferred either knowingly or unknowingly to their students, or if this is just another case of stereotype threat. As this consistent gap is seen across Praxis Subject Assessments in biology, chemistry, and physics, it may warrant further efforts by ETS similar to their efforts to decrease the gap between White and Black examinees (Barton and Coley, 2010; Nettles *et al.*, 2011). Further efforts need to be made in regard to the difference in performance between Black and White examinees taking the Praxis Biology Subject Assessment. Additionally, we recognize that a plethora of other factors may contribute to these achievement gaps seen in the Praxis Biology Subject Assessment, such as economic factors and school-based factors (Hung *et al.*, 2020). Specifically, additional efforts need to be made in biology course work with regard to providing support to these prospective teachers to retain and integrate biology content in order to pass the exam and gain certification.

In looking at the DIF analysis by gender, at most, 10 questions (9%) on any given form were categorized as level C, indicating a question that was included based on the importance of the content, even though significant differences in performance exist between the groups. The values for the difference in Black and White examinee performance were as high as 18 questions (15%) on a form categorized as level C. Further research needs to be conducted on the specifics of the content covered in these items to determine what aspects of this content seem to be causing the difference. This analysis would provide insight into a continuous area of weakness in biology education that may help overcome these gaps in performance.

### Limitations

The findings from this study should be viewed while understanding the limitations inherent within. While the Praxis Biology Subject Assessment presents the most comprehensive representation of content knowledge of teachers seeking certification in biology, there are several large population centers not represented in the data, as they use their own certification exams (e.g., California, New York, Texas). Additionally, the majority of the southeastern United States are without representation in this population. As a result of examinees being able to use their exam scores in any state

accepting the Praxis Biology Subject Assessment for certification, and each state having a relatively different score, the median cut score used in this analysis is a means of interpretation of passing created for this study as a national representation of ability to pass. Furthermore, our study was cross-sectional rather than longitudinal, limiting claims about knowledge retention; we found evidence to suggest a lack of retention of knowledge in those with bachelor's degrees, but we did not collect data on the amount of time between graduation and taking the test. This gap, if large enough, could provide some explanation as to why this potential lack of retention of knowledge was observed. Additionally, due to only 37% of examinees providing a complete demographic profile (as seen in Supplemental Table S2), the demographic characteristics listed are not representative of the entire population but of the majority of the population. Finally, we acknowledge the fact that there is far more that goes into an effective teacher than simply content knowledge and that the breadth of knowledge captured using the Praxis Biology Subject Assessment may not capture the depth of a prospective biology teacher's knowledge.

### ACKNOWLEDGMENTS

The authors would like to thank Jonathan Steinberg of ETS. This work was funded through the National Science Foundation Award to Kennesaw State University (DUE-1557285) and Stony Brook University (DUE-1557292). The opinions set forth in this publication are those of the author(s) and not ETS or the NSF.

### REFERENCES

- Airey, J., & Linder, C. (2009). A disciplinary discourse perspective on university science learning: Achieving fluency in a critical constellation of modes. *Journal of Research in Science Teaching*, 46, 27–49.
- Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19, 716–723.
- American Association for the Advancement of Science (AAAS). (2011). *Vision and change in undergraduate biology education: A call to action*. Washington, DC.
- AAAS. (2015). *Vision and change in undergraduate biology education: Chronicling change, inspiring the future*. Washington, DC.
- Bacharach, V. R., Baumeister, A. A., & Furr, R. M. (2003). Racial and gender science achievement gaps in secondary education. *Journal of Genetic Psychology*, 164(1), 115–126.
- Ball, D. L., Hill, H. C., & Bass, H. (2005). Knowing mathematics for teaching: Who knows mathematics well enough to teach third grade, and how can we decide? *American Educator*, 29, 14–17, 20–22, 43–46.
- Ball, D. L., Thames, M. H., & Phelps, G. (2008). Content knowledge for teaching: What makes it special? *Journal of Teacher Education*, 59, 389–407. doi: 10.1177/0022487108324554
- Bangera, G., & Brownell, S. E. (2014). Course-based undergraduate research experiences can make scientific research more inclusive. *CBE—Life Sciences Education*, 13(4), 602–606.
- Banilower, E. R., Smith, P. S., Weiss, I. R., Malzahn, K. A., Campbell, K. M., & Weis, A. M. (2013). *Report of the 2012 National Survey of Science and Mathematics Education*. Chapel Hill, NC: Horizon Research.
- Barton, P. E., & Coley, R. J. (2010). *The Black-White achievement gap: When progress stopped*. Retrieved May 16, 2018, from [www.ets.org/Media/Research/pdf/PICBWGAP.pdf](http://www.ets.org/Media/Research/pdf/PICBWGAP.pdf)
- Carlsen, W. S. (1993). Teacher knowledge and discourse control: Quantitative evidence from novice biology teachers' classrooms. *Journal of Research in Science Teaching*, 30, 471–481.
- Carter, P. L., & Welner, K. G. (2013). *Closing the opportunity gap: What America must do to give every child an even chance*. Oxford, UK: Oxford University Press.

- Clotfelter, C. T., Ladd, H. F., & Vigdor, J. L. (2007). Teacher credentials and student achievement: Longitudinal analysis with student fixed effects. *Economics of Education Review*, 26(6), 673–682.
- Clotfelter, C. T., Ladd, H. F., & Vigdor, J. L. (2010). Teacher credentials and student achievement in high school. *Journal of Human Resources*, 45, 655–681.
- Creech, L. R., & Sweeder, R. D. (2012). Analysis of student performance in large-enrollment life science courses. *CBE—Life Sciences Education*, 11, 386–391.
- Darling-Hammond, L. (2000). Teacher quality and student achievement. *Education Policy Analysis Archives*, 8, 1.
- Eddy, S. L., Brownell, S. E., & Wenderoth, M. P. (2014). Gender gaps in achievement and participation in multiple introductory biology classrooms. *CBE—Life Sciences Education*, 13(3), 478–492.
- Educational Testing Service (ETS). (2014). *The Praxis study companion biology: Content knowledge*. Retrieved August 25, 2017, from [www.ets.org/praxis/pdf/5235pdf](http://www.ets.org/praxis/pdf/5235pdf)
- ETS. (2015). *Technical manual for the Praxis Series and related assessments*. Princeton, NJ.
- Evens, M., Elen, J., Larmuseau, C., & Depaepae, F. (2018). Promoting the development of teacher professional knowledge: Integrating content and pedagogy in teacher education. *Teaching and Teacher Education*, 75, 244–258.
- Fetler, M. (1999). High school characteristics and mathematics test results. *Education Policy Analysis Archives*, 7(9). <https://doi.org/10.14507/epaa.v7n9.1999>
- Gess-Newsome, J. (2015). A model of teacher professional knowledge and skill including PCK: Results of the thinking from the PCK summit. In Berry, A., Friedrichsen, P., & Loughran J. (Eds.), *Re-examining pedagogical content knowledge in science education*. New York: Routledge.
- Gess-Newsome, J., & Lederman, N. G. (1993). Preservice biology teachers' knowledge structures as a function of professional teacher education: A year-long assessment. *Science Education*, 77, 25–45.
- Gitomer, D. H., Brown, T. L., & Bonett, J. (2011). Useful signal or unnecessary obstacle? The role of basic skills tests in teacher preparation. *Journal of Teacher Education*, 62, 431–445.
- Hakan, K. U. R. T., Ekici, G., Aktas, M., & Aksu, O. (2013). Determining biology student teachers' cognitive structure on the concept of "diffusion" through the free word-association test and the drawing-writing technique. *International Education Studies*, 6(9), 187.
- Hill, J. G. (2011). *Education and certification qualifications of departmentalized public high school-level teachers of core subjects: Evidence from the 2007–08 Schools and Staffing Survey* (NCES 2011–317). Washington, DC: National Center for Education Statistics. Retrieved July 10, 2018, from <http://nces.ed.gov/pubsearch>
- Hung, M., Smith, W. A., Voss, M. W., Franklin, J. D., Gu, Y., & Bounsanga, J. (2020). Exploring student achievement gaps in school districts across the United States. *Education and Urban Society*, 52(2), 175–193.
- Ingersoll, R. M., Merrill, L., & May, H. (2014). *What are the effects of teacher education and preparation on beginning teacher retention?* Philadelphia: Consortium for Policy Research in Education.
- Krajcik, J., Codere, S., Dahsah, C., Bayer, R., & Mun, K. (2014). Planning instruction to meet the intent of the Next Generation Science Standards. *Journal of Science Teacher Education*, 25(2), 157–175. doi: 10.1007/s10972-014-9383-2
- Luckenbill-Edds, L. (2002). The educational pipeline for women in biology: No longer leaking? *BioScience*, 52, 513–521.
- Lyons, K. C. (2013). *2012 National Survey of Science and Mathematics Education: Status of high school biology*. Chapel Hill, NC: Horizon Research.
- Milner, H. R., IV. (2012). Beyond a test score: Explaining opportunity gaps in educational practice. *Journal of Black Studies*, 43(6), 693–718.
- Milner, H. R., IV. (2017). Where's the race in culturally relevant pedagogy? *Teachers College Record*, 119(1), 1–32.
- Nagle, B. (2013). Preparing high school students for the interdisciplinary nature of modern biology. *CBE—Life Sciences Education*, 12(2), 144–147.
- Nettles, M. T., Scatton, L. H., Steinberg, J. H., & Tyler, L. L. (2011). *Performance and passing rate differences of African American and White prospective teachers on Praxis™ examinations: A joint project of the National Education Association (NEA) and Educational Testing Service (ETS)*. ETS Research Report Series. Princeton, NJ: ETS.
- NGSS Lead States. (2013). *Next Generation Science Standards: For states, by states*. Retrieved June 9, 2017, from [www.nextgenscience.org](http://www.nextgenscience.org)
- Osborne, J., Simon, S., & Collins, S. (2003). Attitudes towards science: A review of the literature and its implications. *International Journal of Science Education*, 25(9), 1049–1079.
- Polizzi, S. J., Jaggernauth, J., Ray, H. E., Callahan, B., & Rushton, G. T. (2015). Highly qualified or highly unqualified? A longitudinal study of America's public high school biology teachers. *BioScience*, 65, 812–821.
- Rutledge, M. L., & Mitchell, M. A. (2002). High school biology teachers' knowledge structure, acceptance & teaching of evolution. *American Biology Teacher*, 64(1), 21–28.
- Schwarz, G. (1978). Estimating the dimension of a model. *Annals of Statistics*, 6, 461–464.
- Shah, L., Hao, J., Rodriguez, C. A., Fallin, R., Cortes, K. L., Ray, H. E., & Rushton, G. T. (2018a). Who are our perspective physics teachers? An analysis of Praxis Physics Subject Assessment examinees and performance. *Physical Review Physics Education Research*, 14(1), 010126.
- Shah, L., Hao, J., Schneider, J., Fallin, R., Cortes, K. L., Ray, H. E., & Rushton, G. T. (2018b). Repairing leaks in the chemistry teacher pipeline: A longitudinal analysis of Praxis Chemistry Subject Assessment examinees and scores. *Journal of Chemical Education*, 95(5), 700–708.
- Shulman, L. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard Educational Review*, 57(1), 1–23.
- Sperandeo-Mineo, R. M., Fazio, C., & Tarantino, G. (2006). Pedagogical content knowledge development and pre-service physics teacher education: A case study. *Research in Science Education*, 36, 235.
- Stains, M., Harshman, J., Barker, M. K., Chasteen, S. V., Cole, R., DeChenne-Peters, S. E., ... & Young, A. M. (2018). Anatomy of STEM teaching in North American universities. *Science*, 359(6383), 1468–1470.
- Sutcher, L., Darling-Hammond, L., & Carver-Thomas, D. (2016). *A coming crisis in teaching? Teacher supply, demand, and shortages in the U.S.* Retrieved August 2, 2018, from [https://learningpolicyinstitute.org/sites/default/files/product-files/A\\_Coming\\_Crisis\\_in\\_Teaching\\_BRIEF.pdf](https://learningpolicyinstitute.org/sites/default/files/product-files/A_Coming_Crisis_in_Teaching_BRIEF.pdf)
- Theobald, E. J., Hill, M. J., Tran, E., Agrawal, S., Arroyo, E. N., Behling, S., ... & Freeman, S. (2020). Active learning narrows achievement gaps for underrepresented students in undergraduate science, technology, engineering, and math. *Proceedings of the National Academy of Sciences*, 117(12), 6476–6483.
- U.S. Census Bureau (2020). *Quick facts United States*. Retrieved December 15, 2021, from [www.census.gov/quickfacts/fact/table/US/PST045217](http://www.census.gov/quickfacts/fact/table/US/PST045217)
- Veal Tippins, W. R., & Bell, D. J., & John, (1999). The evolution of pedagogical content knowledge in prospective secondary physics teachers. <https://files.eric.ed.gov/fulltext/ED443719.pdf>
- Yip, D. Y. (1998). Identification of misconceptions in novice biology teachers and remedial strategies for improving biology learning. *International Journal of Science Education*, 20(4), 461–477.
- Zieky, M. (2003). *A DIF primer*. Princeton, NJ: Educational Testing Service Center for Education in Assessment.
- Zwick, R. (2012). *A review of ETS differential item functioning assessment procedures: Flagging rules, minimum sample size requirements, and criterion refinement*. ETS Research Report Series. Princeton, NJ: ETS.