

Review

Review of Racially Equitable Admissions Practices in STEM Doctoral Programs

Sonia F. Roberts ^{1,*}, Elana Pyfrom ^{2,†}, Jacob A. Hoffman ^{2,†}, Christopher Pai ^{2,†}, Erin K. Reagan ^{2,†}
and Alysson E. Light ³

¹ Electrical and Systems Engineering Department, University of Pennsylvania, Philadelphia, PA 19104, USA

² Cell and Molecular Biology Program, University of Pennsylvania, Philadelphia, PA 19104, USA;

pyfrom@pennmedicine.upenn.edu (E.P.); jacob.hoffman@pennmedicine.upenn.edu (J.A.H.);

cpai@pennmedicine.upenn.edu (C.P.); ereagan@pennmedicine.upenn.edu (E.K.R.)

³ Behavioral and Social Sciences Department, University of the Sciences, Philadelphia, PA 19104, USA;

a.light@uscience.edu

* Correspondence: soro@seas.upenn.edu

† Penn Science Policy and Diplomacy Group.

Abstract: This study reviews literature on racially equitable admissions practices relevant to graduate programs in STEM. Graduate Record Exam (GRE) scores correlate more strongly with race, gender, and socioeconomic status than performance metrics for research during or after graduate school. Structural changes to admissions processes that can improve equity of admissions decisions and reduce correlations between admissions decisions and demographic data include using holistic review or composite scores that quantize more components of an application, removing hard limits on course requirements, admitting students as a cohort instead of to individual faculty sponsors, and diversifying admissions committees. Some alternative scoring methods attempt to measure personality traits, but performing these measurements during admissions may present difficulties. Bridge programs—whether they are implemented as collaborations with a minority-serving institution, a personalized educational program for each student admitted to a program, or a stand-alone program before the doctoral degree program—may be able to improve both recruitment and retention of students with underrepresented racial and ethnic identities in their field of study. Finally, financial barriers to applications can disproportionately affect underrepresented applicants due to systemic racism. We end with recommendations for graduate programs to improve equity in admissions processes.

Keywords: composite; GRE; personality; race; PhD; equity; admissions; underrepresented; STEM



Citation: Roberts, S.F.; Pyfrom, E.; Hoffman, J.A.; Pai, C.; Reagan, E.K.; Light, A.E. Review of Racially Equitable Admissions Practices in STEM Doctoral Programs. *Educ. Sci.* **2021**, *11*, 270. <https://doi.org/10.3390/educsci11060270>

Academic Editor: Tabitha Grier-Reed

Received: 5 April 2021

Accepted: 24 May 2021

Published: 31 May 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

1.1. Background

Although Black people represent 12.6% of the United States population [1], in 2019 they earned only 3.5% of doctoral degrees [2]. This underrepresentation is more severe than at the college level, where Black students earned 9% of bachelor's degrees in science and engineering overall in 2018 [1]. Students with Latinx/Hispanic and Indigenous identities are also underrepresented in STEM higher education. Latinx/Hispanic students earned only 4% of doctorates in science and engineering fields in 2019, although they represent 18% of the population [2] and earned 15% of science and engineering bachelor's degrees in 2018 [1]. Indigenous students earned 0.2% of PhDs in 2019 [2], compared with a population of 0.7% [1] and despite earning 0.4% of science and engineering bachelor's degrees in 2018 [1]. In the following, when we refer to “underrepresented” students, we are referring to those students with racial or ethnic identities that are less common in STEM doctoral programs than we would predict from the demographics of the United States, and specifically Black, Latinx/Hispanic, and Indigenous identities.

Students with underrepresented identities may encounter more hardships as undergraduate and graduate students that are directly linked to their race or ethnicity. Black students experience both structural racism [3] and microaggressions from peers and faculty members, sending a constant message that their Black identity is unwelcome in STEM [3–5]. As a result, Black undergraduate students face significant stereotype threat [6,7] in STEM disciplines, where “stereotype threat” refers to stress associated with the risk of confirming a negative stereotype about one’s group. For example, Black students’ experience of group-based performance anxiety—a measure of stereotype threat—predicts their likelihood to drop out of a STEM major [8]. Social isolation and the pressure to represent a demographic group are sources of extra strain for both undergraduate and graduate Black female students [3,9], including exclusion from peer support such as study groups [10]. This result was echoed by a study of Black, Latina, Indigenous, and Asian undergraduate and graduate women in STEM, which found that isolation and microaggressions were common stressors [11]. Some students use negative stereotypes and microaggressions as a motivating factor [12–14], and this “stereotype management” propels them to complete a terminal degree [5]. However, it takes extra time and energy to manage stereotypes using this coping strategy [13,15]—time and energy which other students do not have to spend. When a student is the only representative of their race or ethnicity in a department, many of these problems are made significantly worse because of the increased pressure to represent the whole of their demographic group [3].

1.2. Contributions

To begin to address the issue of underrepresentation of Black, Hispanic/Latinx, and Indigenous students in STEM doctoral programs (“underrepresented” students), we survey the literature for equitable admissions practices with particular attention to practices that STEM graduate programs may be able to adopt. We begin by surveying the literature on the Graduate Record Examination (GRE) and its impact on graduate student admissions, followed by a review of alternative metrics and methods to evaluate applications. We then survey literature on additional institutional factors, including bridge programs, admissions committee makeup, cohort vs. individual admissions, and financial barriers, that should be considered when designing admissions policies. We close with a series of best practices recommendations based on the literature we surveyed.

2. Standardized Tests and Alternatives

We first review the literature on the Graduate Record Exam and some alternative quantifications of application data that can be used for quick comparisons between applicants.

2.1. GRE Scoring Is a Poor Metric for Career Success

The Graduate Record Examination, developed and administered by the private non-profit Educational Testing Service, is a standardized test used by many graduate institutions as a means of evaluating applicants. The GRE is commonly used and trusted during the admissions process [16], likely because early research [17,18] indicated that high GRE scores correlated with success metrics (e.g., likelihood of passing qualifying exams, degree completion, research productivity). However, more recent research contradicts and complicates those findings. A 2017 study of the Vanderbilt University Medical School Biomedical Umbrella Program, a doctoral program at the Vanderbilt medical school, showed that the GRE did not predict graduation with a PhD, qualifier performance, number of first-author presentations, or grants and fellowships received [19]. Another study surveying large graduate institutions with between 1,500 and 14,000 STEM graduate students undermines the positive correlation even further: Researchers found that overall, higher GRE scores did not correlate with PhD program completion, but for male students, a higher GRE score predicted being *more* likely to drop out of the program, while a lower score predicted finishing a terminal degree [20]. Another study attempting to predict PhD program success (defined as student retention and program completion) based on a variety of factors collected during

admissions found that disregarding the GRE actually improved the predictive power of their model [21]. Of course, this sample is not representative of the population since it is limited to students who have already been admitted to the school in question [22], but we were unable to identify any studies that did not suffer from this flaw. Finally, a study of one eighth of the physics PhD student population between 2000 and 2010, sampled from institutions graduating at least 10 physics PhD students per year, found that neither the physics subject test score nor the verbal GRE score correlated with program completion [23]. The quantitative GRE was correlated with completion, but there was only a 7% difference in the completion rate between students who had scored in the 10th and 90th percentiles on the test. This correlation also disappeared for any of the smaller subsamples that the researchers tested: it was present for all physics PhD students and all students who were US citizens, but disappeared when looking at the population of either only female US citizens or male US citizens.

The GRE does, however, strongly correlate with race, gender, and socioeconomic status—much more strongly than with predictions of later academic career success, which have only a weak correlation ($r < 0.3$) [24–26] (See Table 1 for a summary). This correlation has sizable, negative impacts on equity. A 2014 analysis found that a GRE cutoff of 700 eliminated 94.8% of Black applicants from the admissions pool in the physical sciences versus 18% of white and Asian applicants. In a qualitative study of students with underrepresented identities eligible for the American Physical Society’s Bridge program, a third of the graduates, who were interested in applying to physics PhD programs but chose not to, cited GRE scores as a reason. Analysis of admissions decisions for six large physics PhD programs in 2019 revealed that GRE and physics subject GRE scores significantly predicted whether the applicant was shortlisted and admitted [16]. Once shortlisted, applicants with underrepresented racial and ethnic identities were much more likely to be admitted than White applicants. Controlling for GRE and physics GRE scores, underrepresented applicants were also more likely to be admitted than White applicants. Despite this, underrepresented applicants were not favored in the final shortlisted and admitted populations. Applicants’ GRE scores played an outsized role in determining whether an applicant was shortlisted or admitted, and likely disqualified otherwise-highly desirable applicants of underrepresented identities. It is worth considering whether performance on any given metric reflects an applicant’s ability to *sustain* effort throughout the graduate program, or if it simply reflects their ability to *perform* on that metric. Using the GRE as a decision-making metric for graduate school admissions can be a source of inequity while also failing to select the applicants who are most likely to be productive researchers.

Table 1. The top row of the table includes references to the papers surveyed. The papers are arranged in order of publication, with more recent papers to the left and older papers to the right. The rows correspond to specific outcomes that these authors attempted to predict using GRE general scores. The values in each column correspond to positive +, weak positive (+), neutral 0, weak negative (−), or negative—correlations. The difference between a strong and weak correlation is whether $r > 0.3$.

Paper	[23]	[20]	[19]	[24]	[27]	[25]	[26]	[17]	[18]
Qual performance		0						(+)	
Completed PhD	(+)/0 *	0/− **	0					(+)	
Faculty rating			(+)						
Student ranking					0				
GPA (1st yr, cum)			(+)			(+)	(+)	+	+
Research productivity			0	0		0		(+)	
Grants received			0						
Race				+		+	+		

Table 1. Cont.

Paper	[23]	[20]	[19]	[24]	[27]	[25]	[26]	[17]	[18]
Gender				+		+	+		
Socioeconomic						+			

* There was a correlation between physics PhD students' quantitative GRE scores and program completion, but only a 7% difference in program completion for test takers who scored in the 10th and 90th percentiles. There was no correlation between the verbal GRE or physics subject GRE scores and program completion. ** The general population of students had no correlation between GRE score and degree completion; men had a negative correlation, with higher-scoring men more likely to drop out.

2.2. Alternative Methods and Metrics for Applicant Evaluation

The recent research on the GRE's bias and lack of predictive power for graduate school success has motivated the development of many methodologies to de-emphasize or even remove the GRE during the applicant review process. Such methods are collectively referred to as a "holistic review" [28]. There are two main strategies: creating new quantitative measures that incorporate more portions of the application, and including assessments of additional skills or personality traits that may promote research career success.

2.2.1. Composite Scores

Some institutions have attempted to operationalize traditionally less standardized parts of an application, creating new, integrated metrics that can more effectively be used as admissions criteria. In one study, of Puerto Rican biomedical PhD students, students' applications were evaluated using a composite score combining GRE score, undergraduate GPA, research experience, advanced coursework or degrees, conference presentations, and past publications [21]. To validate the composite score, the researchers retroactively evaluated the applications of students and recent graduates. They found a strong correlation between high composite score and success metrics including program completion, fellowship awards, and shorter time to thesis defense. Undergraduate GPA and science GPA were not correlated with success metrics, and removing them from the composite score did not change the correlation of the composite score with the success metrics. Removing the GRE from the composite score, on the other hand, *improved* the strength of the correlations with success metrics.

One particularly strong example of the effect a composite score can have on admissions equity comes from the healthcare field. A study comparing methods of selecting medical students to interview for a psychiatry residency program [29] found that using a composite score resulted in a significantly more diverse set of students than using a traditional scoring method that relies heavily on standardized test scores and society memberships, even when the traditional method was modified to account for race. The composite score was calculated from six main factors—leadership, community service, clinical performance, research, recommendation letters, and professionalism—as well as two metrics of lived experience. The lived experience metrics were "resilience," or having overcome specific obstacles such as illness or discrimination, and "distance traveled," referring to the student's trajectory relative to the expected trajectory of someone with their familial and community background; for example, being a first-generation college student. Race and ethnicity were not explicitly included and were not sufficient on their own to earn recognition of resilience and distance traveled, making it even more striking that the odds of an underrepresented student being selected for interview were more than twice as high under holistic review than traditional review (0.84 vs. 0.35, $p < 0.05$). Holistic review increased the odds of inviting an underrepresented student to interview more than modified traditional review (0.54), which explicitly used race to calculate the score. Using a composite score that combines advanced coursework, research experience, and evidence of personal and professional characteristics such as leadership and community service may produce a student body that is more diverse and more likely to succeed in graduate school.

To further decrease bias in a composite score admissions process, admissions officers can generate a standardized set of priorities and criteria to guide subjective evaluations *a priori* during admissions. Using a standardized set of criteria reduces the influence of an application reviewer's emotional state, confirmation bias, or potential priming, and makes the data easier to represent and interpret [30]. Reviewing each part of an application separately (for example, all applicants' research experience, and then all applicants' letters of recommendation) can also reduce cognitive bias. Standardizing the review process is extremely important because bias in the evaluation of an application is not limited to the GRE, but can also appear in the undergraduate GPA, letters of recommendation, and in access to research opportunities [31]. Public and private undergraduate degree-granting institutions give out different percentages of A's and B's, as do institutions in different geographic regions [32]. Letter writers describe people from different demographic backgrounds differently [33,34], and even when the language is the same, readers can interpret it differently depending on the race or gender of the letter's subject. For example, even if the language in the letter is overall very positive, readers may interpret "X is brilliant" to mean "X is brilliant (for X's group)" if X has an underrepresented identity in that field [35,36]. Finally, evaluating research experience and evidence of research project completion ought to be modulated by consideration of the undergraduate institution and other work experience [22,31]. The applicant may not have had access to opportunities for research at their previous institution. Thirty percent of recent physics graduates with underrepresented identities, who were interested in applying to PhD programs but ultimately chose not to, cited a lack of research experience as a reason [37]. Although an applicant may not have research experience, they may have been able to demonstrate successful independent work in other ways through independent studies, open-ended projects in courses, or leadership in student organizations. Some applicants may also have not been able to volunteer for unpaid research positions if they need to financially support themselves or their families.

2.2.2. Personality Scoring in Admissions

Programs at other universities attempt the difficult task of measuring personality traits and features of emotional intelligence that may indicate preparedness for graduate school. One example comes from two PhD programs in physics. The University of South Florida and Fisk-Vanderbilt, a partnership program between Fisk University, a Historically Black College and University (HBCU), and Vanderbilt University [19] use a careful, multi-stage interview process to examine "college and research experiences, key relationships, leadership experience, service to community and life goals" in an attempt to measure commitment to research, maturity, perseverance, adaptability, conscientiousness [25], and "grit," a combination of passion and perseverance which has been studied in depth by Angela Duckworth and colleagues (e.g., [38]) [39]. The USF and Fisk-Vanderbilt PhD completion rate is 80%, compared with 50% on average in the country for physics, with a more diverse student body than average as well: 84% of students in the first year of the program were underrepresented [25]. There is independent evidence to support the approach of predicting graduate school success from personal and emotional traits: A study of physicists and chemists found that a desire to learn predicted career success (determined by primary or first-author publications and grant funding) better than a desire to follow up on high grades in an undergraduate program or a "good student" identity [40].

However, the use of personality scores in admissions processes should be approached with caution. The "big-five" personality scoring system is the most commonly used system in psychology, with five traits: extraversion, agreeableness, openness, conscientiousness, and neuroticism [41]. Though there is some evidence that self-reported affinity with certain adjectives associated with big-five traits is stable over time for working-age adults [42], behavior associated with these personality traits shows marked variability [43]. Furthermore, personality development occurs during emerging adulthood, and the social roles that we assume can influence the development of big-five personality traits [44]—in other words, the experience of being a graduate student could foster personality development

and increase traits associated with success in graduate school and beyond. It is, therefore, not reasonable to assume that a low score on a certain trait during an admissions assessment reflects a stable low expression of that trait—especially not for the population of mostly young adults applying to graduate school. Grit may be a better predictor of success in graduate school than big-five factors. However, it is difficult to differentiate grit from conscientiousness in simple-to-administer tests [45]. Participants can also display reduced grit and conscientiousness when they feel they have a scarcity of time or other resources [46]. This could lead to large amounts of potential measurement variability due to the time constraints on any student during the admissions process, and may particularly affect students with early-life low socioeconomic status, since early experiences of scarcity predict adulthood responses to scarcity [47]. The scarcity of time alone for applicants, most of whom are working or studying full-time while applying, would likely be enough to bring out these differences. Ad-hoc measurements of traits such as grit will also be highly subject to admissions officers' own life experiences and ability to relate to the hardships of the students they are reviewing [48,49]. Finally, traits such as conscientiousness or grit have high "face validity" when used during higher education admissions: It is usually very easy for the person being tested to guess both what they are being tested for, and to guess how to respond in a way that fits the desired profile [50]. There is also evidence that the ability to fake desired responses is based on what the subjects know, meaning that applicants with insider knowledge about graduate school may fare better on these kinds of assessments [51]. Due to the difficulty of accurate measurement, it is our opinion that attempts to use personality traits in admissions processes ought to be approached with significant care and ample consultation from experts in personality psychology.

2.2.3. Hybrid Holistic Review

A hybrid approach that uses a holistic review method on some but not all the students during the first round of application assessments can confer the benefits of holistic review while reducing both the time spent reviewing applications and the bureaucratic challenge of consistently applying holistic review throughout a decentralized admissions process [28]. A study assessing the hybrid approach to admissions to the doctoral program at the University of Texas Medical School shows promising results [52]. Applicants to their doctoral program were "administratively triaged based on Quantitative and Verbal Reasoning GRE scores". Applicants scoring above the 50th percentile automatically continued to the discussion round of applicant selection. Applicants scoring below the 50th percentile were further reviewed by three faculty members, who were instructed to evaluate the applicant based on their personal statement, letters of recommendation, and research experience. If at least two of the three faculty agreed that the application merited further discussion, it advanced to the discussion round with the applications of those scoring above the 50th percentile. Importantly, reviewers in the discussion round were not informed of the process through which the application entered the discussion pool. Use of this method resulted in admissions decisions that were independent of race, gender or citizenship status without requiring that every applicant undergo the holistic review process. This hybrid approach appears to provide the full benefit of increasing admitted student diversity to the levels in the applicant pool, while mitigating the time spent performing holistic reviews by faculty and administrators.

3. Structural Changes that Can Support Increased Diversity during Admissions

Replacing biased metrics with some form of holistic review is not the only way to decrease bias during admissions and increase diversity in the student body. Structural changes can also have an impact, including bridge programs, the makeup of admissions committees, using cohort-based rather than sponsor-based admissions, and structuring the admissions process to consider applicants with lower incomes.

3.1. Bridge Programs

A resource-intensive structural change with substantial promise is the development of bridge programs that provide additional training for admitted doctoral students, either before the doctoral program starts or concurrently. Brown University's Initiative to Maximize Student Development program in the Division of Biology and Medicine [53] built several such "bridges." First, Brown partnered with several undergraduate institutions that served underrepresented students in the biological sciences to ensure that the students graduating from those programs in good standing would be well prepared for Brown's graduate programs. They also developed module-based personalized education programs for students, which had the added benefit of engaging faculty in supporting these diversity initiatives. The training modules were available to all students, though students in the bridge program were prioritized in enrollment, and were deemed useful enough by faculty to be recommended to many students outside of the bridge program. Finally, students were not denied admissions on the basis that they needed additional undergraduate coursework, because the personalized education program could include undergraduate courses. As a result of these partnerships, underrepresented student enrollment dramatically increased, from 3% to 20%. These students also went on to receive more fellowships, publish at higher rates, earn more non-minority travel grants, and present more at conferences. Institutional investments such as this can result in major benefits for all students, not just marginalized students.

3.2. Admissions Committee Diversity and Cohort-Based Admissions

The makeup of admissions committees and whether students are admitted as a cohort or sponsored by an individual lab can also affect inequities in the admissions process. Although institutions usually consider diversity and individual excellence to be complementary and believe that excellent institutions must admit diverse students [54], most faculty admissions officers at predominantly White institutions do not consider diversity during the first (or even first few) rounds of application review [49]. Instead, they focus on easily quantifiable metrics of "excellence" such as the GRE, which, as we argued previously, do not correlate with graduate school success. Faculty of color, on the other hand, typically resist relying on standardized metrics such as the GRE during application review [48], and prefer to work with students who express interest in or have done previous work on diversity, equity, and inclusion efforts—which also correlates with a more diverse population of students [49]. One way to address the mismatch between the university's stated goals and the admission committees' behavior would be to diversify admissions committees throughout all parts of the admissions process. However, given that the diverse faculty themselves are underrepresented in the faculty population, faculty of color face a considerable burden when they are expected to "fix" diversity failures, often without adequate support or compensation.

Especially as departments transition their hiring and recruitment practices, diverse admissions officers could include current graduate students, postdoctoral fellows in the department, and alumni of the program. Although institutions cannot grant individuals access to an applicant's educational records without the applicant's consent due to the Family Educational Rights and Privacy Act of 1974, it would be trivial to ask applicants to grant such consent by including a descriptive explanation of why their consent is requested. This change would have the added benefit of distributing labor, which might facilitate the adoption of more onerous but advantageous holistic admissions procedures. Departments should also recognize that by incorporating trainees and those early in their career into this process, they are asking people to perform work that is and should be entirely the responsibility of departmental leadership; as such, trainees should be compensated accordingly for both their time and expertise in the multifaceted nature of marginalization [48].

Another structural change that could have a large impact is to use cohort-based admissions instead of sponsor-based admissions; that is, programs which use a model

where faculty choose individual students to sponsor could switch to a model where the graduate students are admitted as a cohort and gain faculty sponsorship through exposure to professors in classroom settings, department open houses at the start of the school year, or a rotation program. When the admissions officers can review the cohort as a group instead of considering each applicant individually during admissions, it is more apparent to the admissions team whether certain demographics are overrepresented in the cohort [55], and makes creating a more diverse cohort more attractive [56,57]. Given that universities and individuals on the admissions committees claim to value diversity in the student body [54] but only underrepresented admissions officers behave in accordance with those values [48,49], since cohort-based admissions encourage all admissions officers to consider the diversity of the forming student body earlier in the admissions process [55], changing to cohort-based admissions could bring the behavior of admissions committees more in line with their stated values.

3.3. Reducing or Removing Financial Barriers to Applicants

Financial barriers serve as a strong selector that can reduce diversity in the applicant pool [9]. Application fees in particular can be a strong financial disincentive, preventing lower-income students from applying to graduate programs. Due to systemic racism, financial burdens will disproportionately affect students who are already underrepresented. Although many schools offer fee waivers to lower-income students, these waivers are frequently offered late in or at the end of an application process, thus discouraging students who need the fees to apply from submitting an application [58]. Furthermore, fee waivers often involve more paperwork than an applicant may anticipate, particularly when facing an application deadline [58]. A relevant example of the discouraging effect that financial assistance paperwork can have on narrowing an application pool is the Free Application for Federal Student Aid (FAFSA) form for undergraduate financial assistance; for many applicants, filling out the FAFSA is seen as an insurmountable task, preventing them from applying to undergraduate programs [59]. Considering the strong correlation between economic status and race due to systemic racism, the financial burden of application fees combined with the amount of paperwork required to overcome them could be substantially limiting diversity in the applicant pool.

Institutions could solve this problem by transitioning to low-paperwork fee waivers. A 2013 study of high-achieving, low-income high school students found that access to low-paperwork fee waivers significantly increased the number of schools that an individual student would apply to [58]. Another study looking at the effect of universities offering FAFSA assistance to students found that this assistance increased the rate of FAFSA submission [59]. Taken together, these studies indicate that reducing the financial barriers to application can prevent the loss of low-income students from an applicant pool. These types of interventions have the added benefit of requiring minimal technical effort, as many schools already offer application fee waivers [58]. Moving these fee waivers to the beginning of an application and ensuring that they require low levels of information gathering to complete will make it easier for lower-income students to apply to graduate programs, removing another barrier to entry for a very modest cost.

There may be other opportunities to reduce the financial burden for lower-income applicants during admissions processes. For example, if the applicant is burdened by the application fee, they will likely also be burdened by the cost of traveling for an interview and waiting an uncertain and sometimes lengthy time for reimbursement. Programs covering this cost upfront only alters *when* university resources are required, not *how much* of that resource is required.

4. General Recommendations for Admissions Processes

Our assessment of the literature indicates that there are at least five practices that ought to be adopted as part of any equitable admissions process:

1. The GRE should not be collected nor considered during applications. It fails to consistently predict a graduate student's success, and using the GRE to reduce the number of applicants reduces the number of racial and gender minorities considered for admission. In its place, we recommend a hybrid approach with a GPA threshold and holistic review; to minimize the amount of time holistic review of all students requires, holistic review should be performed for all applicants except those exceeding a predetermined GPA threshold. A preferable approach would be to perform holistic review or calculate a composite score for all students, but we acknowledge that universities which are currently using soft or hard GRE thresholding may not be able to move directly to performing holistic review or composite scores for all students.
2. Admissions committees must be diversified. Recognizing that there may be insufficient faculty available to provide this diversity, we suggest involving a diverse group of graduate students, recent alumni, or postdoctoral trainees to temporarily increase the diversity of admissions committees while faculty diversity is improved. These students, alumni, and postdoctoral trainees should be financially compensated for this work.
3. At minimum, graduate institutions should engage with HBCUs and minority-serving institutions and organizations to explore how the graduate programs can assist in the development and recruitment of their students.
4. Any formal requirements for applicants to have taken specific undergraduate courses should be discarded. Instead, incoming students can work with their advisers to determine how they can best learn the material they need for their research or qualifiers. Examples could be by taking a class, by auditing the portion of a class most relevant to their graduate work, or by forming a study group with other students interested in learning or reviewing that material.
5. Application fee waivers should be publicized at the start of the admissions process, be simple to complete, and be accessible to all underrepresented students, not only to those who have attended a specific conference or formal post-baccalaureate program.

The above recommendations are only a starting place, and graduate institutions can certainly do more. Next, we suggest practices that support excellence in admissions equity rather than the bare minimum.

1. Rather than only performing holistic review of students whose GPA is below the predetermined threshold, the admissions team could perform holistic review of all students. The review process should be standardized and include generation of a composite score based on an applicant's personal statement, letters of recommendation, research experience, leadership experience, and community service.
2. Graduate programs should develop training modules for early-stage PhD students to learn necessary research practices instead of denying admission for the sake of admitting students in a cohort who all have previous exposure to the same set of lab techniques.
3. Universities with graduate programs should work with undergraduate institutions, particularly minority-serving institutions, to tailor their programs to a graduate institution's desires for an applicant. This could significantly improve both the diversity and excellence of applicants.
4. Given that costs during the application process can be a barrier to graduate school, the other financial costs of attending graduate school could be alleviated by providing underrepresented students with additional internal fellowships and paying all graduate students' first month stipend plus a moving bonus before they are expected to move in. Before and during the program, financial equity could also take the form of an internal financial incentive given to anyone who applies for external fellowships.

Finally, we present two broader, more structural suggestions for institutions that are committed to building an equitable graduate program.

1. Departments should matriculate graduate students as a cohort rather than based on individual faculty sponsorship, and allow students to find faculty mentors after arrival. It is by no means a simple task to convert to this process throughout an institution, especially for fields which cannot easily break graduate student work into a trial-sized piece that a student can complete during a rotation. However, the ability to evaluate matriculating student diversity is vital in identifying ongoing racist patterns in admissions.
2. Universities should develop preparatory programs for future applicants to graduate school, such that interested individuals can build the requisite skills and knowledge directly under the university's supervision. Although these programs will no doubt take time, resources, and careful consideration to make them not only available, but accessible, they are some of the best ways to recruit and support underrepresented students and improve equity throughout and after degree completion.

5. Discussion

We review literature relevant to racial equity in admissions to graduate STEM programs. First, we review the results of analyses linking parts of admitted students' data with their performance in graduate school, finding that standardized tests are not accurate predictors of success in graduate school and subsequent research careers. We describe some existing interventions that other universities have performed, such as using alternative scoring methods, bridge programs, and personality scores, and we discuss some potential pitfalls of implementing these interventions. We also suggest strategies to mitigate those pitfalls, such as consulting with personality psychologists instead of having individual departments or universities develop ad-hoc personality measurements. We address the importance of reducing financial barriers, which disproportionately affect underrepresented students due to systemic racism. Finally, we provide recommendations to graduate programs, ordered according to the level of structural overhead required to make the change.

It is important to note that the scope of this review covers only equitable admissions practices, but there are many other practices that may improve recruitment and retention of doctoral students with underrepresented racial and ethnic identities. Increasing the number of underrepresented students may also relieve some aspects of the strain of isolation or the pressure to represent an entire demographic [3]. However, Black students encounter extra barriers due to bias in graduate school and spend more energy on stereotype management than White students [5]. Studies of Black academics who persist in STEM show that developing a robust STEM identity strongly predicts future involvement in science [60–63]. Increasing the number of mentors and role models for Black students may aid in identity development and thus retention. Creating “counterspaces,” or supportive spaces for Black, Latinx/Hispanic, Indigenous, and Asian students that counter the dominant racist STEM culture, may also be protective [9,11,14]. These can take several forms, including groups with heterogeneous identities and relationships between people at different power levels in an institution. Increasing the number of matriculating underrepresented graduate students will only go so far towards increasing the number of these students graduating with a PhD if the campus climate is not conducive to their wellbeing.

Underrepresented students may also be less likely to apply to PhD programs in the first place, particularly at predominantly White institutions, because these programs are not currently the best way to achieve their goals [37]. When deciding to pursue a STEM graduate degree, Black doctoral students are similar to students of other races in that they cite passion for the discipline as a primary motivator [15]. However, they are also likely to cite positive mentoring experiences, a desire to mentor others, and the social justice paradigm of giving back to one's community [10,15,64]. These goals are mirrored in undergraduate Black and Latinx/Hispanic populations [64]. The cut-throat and highly competitive hyper-focus on research at predominantly White universities may make it difficult for students to work towards these additional goals in a traditional academic

position [15]. Additionally, recruitment efforts for STEM disciplines often focus on financial gain as a motivation for a STEM career, whereas Black students find equity ethics more compelling as a motivation for joining or remaining in STEM [64]. Finally, the academic culture at the faculty level may influence students' interest in an academic career. Black faculty experience frequent microaggressions [65], sometimes even in the presence of their students [66], with efforts to confront the aggressors proving futile [65]. Students who have already experienced racism from professors as undergraduates may have reasonable concerns about the amount of highly subjective assessment involved in important milestones for academics, such as passing qualifiers as a PhD student or tenure review as a professor. Over time, underrepresented students may become less interested in pursuing a PhD rather than a professional terminal degree such as a masters or a medical degree if they perceive that becoming an academic would require either abandoning some of their primary motivations towards a STEM career or accepting a higher risk of job insecurity than White peers experience due to biased subjective assessments. It is possible that an academic climate that celebrates multiple routes to success, such as valuing community leadership, equity and administrative work, and course development more highly for PhD students and tenure-track faculty, might improve retention of underrepresented scholars. However, it is impossible to retain underrepresented students in STEM disciplines if these students are not admitted to the program, and to improve the equity of outcomes in higher education, work is needed on both fronts.

Author Contributions: Conceptualization, E.K.R.; methodology, S.F.R., E.P., and J.A.H.; investigation, S.F.R., E.P., J.A.H., and A.E.L.; writing—original draft preparation, C.P., S.F.R., E.P., and J.A.H.; writing—review and editing, S.F.R., A.E.L., and E.P.; supervision, E.K.R. and A.E.L.; project administration, E.K.R. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: No data were collected for this review.

Acknowledgments: This work was initiated and largely conducted by members of the Penn Science Policy and Diplomacy Group, an extracurricular student organization at the University of Pennsylvania. We would like to thank Shannon Wolfman, PhD, of the University of Pennsylvania, Ishita Kamboj and Alex Hsain of North Carolina State University, Annabelle Lolinco of Iowa State University, and Ansha Clement, EdD of Organizational Development Student for assistance with revisions. We would also like to thank all other members of the Penn Science Policy and Diplomacy Group's Racial Justice Working Group for assistance and support throughout the researching, writing, and revising process.

Conflicts of Interest: The authors declare no conflict of interest. As this work was unfunded, no funders had any role in the design of the review, or interpretation of its findings, or in the writing of the manuscript, or in the decision to publish the results.

References

1. National Science Foundation, National Center for Science and Engineering Statistics. 2021. Available online: <https://nces.nsf.gov/pubs/nsf21321/data-tables> (accessed on 21 May 2021).
2. National Science Foundation, National Center for Science and Engineering Statistics. 2019. Available online: <https://nces.nsf.gov/pubs/nsf21320> (accessed on 21 May 2021).
3. McGee, E.O.; Bentley, L. The troubled success of Black women in STEM. *Cogn. Instr.* **2017**, *35*, 265–289. [CrossRef]
4. Martin, D.B. Race, racial projects, and mathematics education. *J. Res. Math. Educ.* **2013**, *44*, 316–333. [CrossRef]
5. McGee, E.O.; Martin, D.B. "You would not believe what I have to go through to prove my intellectual value!" Stereotype management among academically successful Black mathematics and engineering students. *Am. Educ. Res. J.* **2011**, *48*, 1347–1389. [CrossRef]
6. McGee, E. "Black genius, Asian fail": The detriment of stereotype lift and stereotype threat in high-achieving Asian and Black STEM students. *AERA Open* **2018**, *4*, 2332858418816658. [CrossRef]

7. Steele, C.M.; Aronson, J. Stereotype threat and the intellectual test performance of African Americans. *J. Personal. Soc. Psychol.* **1995**, *69*, 797. [[CrossRef](#)]
8. Beasley, M.A.; Fischer, M.J. Why they leave: The impact of stereotype threat on the attrition of women and minorities from science, math and engineering majors. *Soc. Psychol. Educ.* **2012**, *15*, 427–448. [[CrossRef](#)]
9. Ong, M.; Wright, C.; Espinosa, L.; Orfield, G. Inside the double bind: A synthesis of empirical research on undergraduate and graduate women of color in science, technology, engineering, and mathematics. *Harv. Educ. Rev.* **2011**, *81*, 172–209. [[CrossRef](#)]
10. Rosa, K.; Mensah, F.M. Educational pathways of Black women physicists: Stories of experiencing and overcoming obstacles in life. *Phys. Rev. Phys. Educ. Res.* **2016**, *12*, 020113. [[CrossRef](#)]
11. Ong, M.; Smith, J.M.; Ko, L.T. Counterspaces for women of color in STEM higher education: Marginal and central spaces for persistence and success. *J. Res. Sci. Teach.* **2018**, *55*, 206–245. [[CrossRef](#)]
12. Hodari, A.K.; Ong, M.; Ko, L.T.; Smith, J.M. Enacting agency: The strategies of women of color in computing. *Comput. Sci. Eng.* **2016**, *18*, 58–68. [[CrossRef](#)]
13. Hodari, A.K.; Ong, M.; Ko, L.T.; Smith, J. Enabling courage: Agentic strategies of women of color in computing. In Proceedings of the 2015 Research in Equity and Sustained Participation in Engineering, Computing, and Technology (RESPECT), Charlotte, NC, USA, 13–14 August 2015; pp. 1–7.
14. Ko, L.T.; Kachchaf, R.R.; Hodari, A.K.; Ong, M. Agency of women of color in physics and astronomy: Strategies for persistence and success. *J. Women Minor. Sci. Eng.* **2014**, *20*, 171–195. [[CrossRef](#)]
15. McGee, E.O.; White, D.T.; Jenkins, A.T.; Houston, S.; Bentley, L.C.; Smith, W.J.; Robinson, W.H. Black engineering students' motivation for PhD attainment: Passion plus purpose. *J. Multicult. Educ.* **2016**, *10*, 167–193. [[CrossRef](#)]
16. Posselt, J.R.; Hernandez, T.E.; Cochran, G.L.; Miller, C.W. Metrics first, diversity later? Making the short list and getting admitted to physics PhD programs. *J. Women Minor. Sci. Eng.* **2019**, *25*, 283–306. [[CrossRef](#)]
17. Kuncel, N.R.; Hezlett, S.A. Standardized tests predict graduate students' success. *Science* **2007**, *315*, 1080–1081. [[CrossRef](#)]
18. Kuncel, N.R.; Hezlett, S.A.; Ones, D.S. A comprehensive meta-analysis of the predictive validity of the graduate record examinations: implications for graduate student selection and performance. *Psychol. Bull.* **2001**, *127*, 162. [[CrossRef](#)] [[PubMed](#)]
19. Moneta-Koehler, L.; Brown, A.M.; Petrie, K.A.; Evans, B.J.; Chalkley, R. The limitations of the GRE in predicting success in biomedical graduate school. *PLoS ONE* **2017**, *12*, e0166742. [[CrossRef](#)] [[PubMed](#)]
20. Petersen, S.L.; Erenrich, E.S.; Levine, D.L.; Vigoreaux, J.; Gile, K. Multi-institutional study of GRE scores as predictors of STEM PhD degree completion: GRE gets a low mark. *PLoS ONE* **2018**, *13*, e0206570. [[CrossRef](#)] [[PubMed](#)]
21. Pacheco, W.I.; Noel Jr, R.J.; Porter, J.T.; Appleyard, C.B. Beyond the GRE: Using a composite score to predict the success of Puerto Rican students in a biomedical PhD program. *CBE Life Sci. Educ.* **2015**, *14*, ar13. [[CrossRef](#)]
22. Miller, A.; Crede, M.; Sotola, L.K. Should research experience be used for selection into graduate school: A discussion and meta-analytic synthesis of the available evidence. *Int. J. Sel. Assess.* **2021**, *29*, 19–28. [[CrossRef](#)]
23. Miller, C.W.; Zwickl, B.M.; Posselt, J.R.; Silvestrini, R.T.; Hodapp, T. Typical physics Ph. D. admissions criteria limit access to underrepresented groups but fail to predict doctoral completion. *Sci. Adv.* **2019**, *5*, eaat7550. [[CrossRef](#)] [[PubMed](#)]
24. Hall, J.D.; O'Connell, A.B.; Cook, J.G. Predictors of student productivity in biomedical graduate school applications. *PLoS ONE* **2017**, *12*, e0169121. [[CrossRef](#)]
25. Miller, C.; Stassun, K. A test that fails. *Nature* **2014**, *510*, 303–304. [[CrossRef](#)]
26. Miller, C.W. Admissions criteria and diversity in graduate school. *arXiv* **2013**, arXiv:1302.3929.
27. Weiner, O.D. How should we be selecting our graduate students? *Mol. Biol. Cell* **2014**, *25*, 429–430. [[CrossRef](#)]
28. Kent, J.D.; McCarthy, M.T. *Holistic Review in Graduate Admissions*; Council of Graduate Schools: Washington, DC, USA, 2016.
29. Barceló, N.E.; Shadravan, S.; Wells, C.R.; Goodsmith, N.; Tarrant, B.; Shaddox, T.; Yang, Y.; Bath, E.; DeBonis, K. Reimagining Merit and Representation: Promoting Equity and Reducing Bias in GME Through Holistic Review. *Acad. Psychiatry* **2020**, *45*, 34–42. [[CrossRef](#)] [[PubMed](#)]
30. Sukumar, P.T.; Metoyer, R. A visualization approach to addressing reviewer bias in holistic college admissions. In *Cognitive Biases in Visualizations*; Springer: Berlin/Heidelberg, Germany, 2018; pp. 161–175.
31. Woo, S.E.; LeBreton, J.; Keith, M.; Tay, L. Bias, Fairness, and Validity in Graduate Admissions: A Psychometric Perspective. *Preprint* **2020**. [[CrossRef](#)]
32. Rojstaczer, S.; Healy, C. Where A is ordinary: The evolution of American college and university grading, 1940–2009. *Teach. Coll. Rec.* **2012**, *114*, 1–23.
33. Grimm, L.J.; Redmond, R.A.; Campbell, J.C.; Rosette, A.S. Gender and racial bias in radiology residency letters of recommendation. *J. Am. Coll. Radiol.* **2020**, *17*, 64–71. [[CrossRef](#)]
34. Madera, J.M.; Hebl, M.R.; Dial, H.; Martin, R.; Valian, V. Raising doubt in letters of recommendation for academia: Gender differences and their impact. *J. Bus. Psychol.* **2019**, *34*, 287–303. [[CrossRef](#)]
35. Biernat, M.; Villicana, A.J.; Sesko, A.K.; Zhao, X. Effects of dyadic communication on race-based impressions and memory. *Group Process. Intergroup Relat.* **2018**, *21*, 302–318. [[CrossRef](#)]
36. Collins, E.C.; Biernat, M.; Eidelman, S. Stereotypes in the communication and translation of person impressions. *J. Exp. Soc. Psychol.* **2009**, *45*, 368–374. [[CrossRef](#)]
37. Cochran, G.L.; Hodapp, T.; Brown, E.E. Identifying barriers to ethnic/racial minority students' participation in graduate physics. In Proceedings of the Physics Education Research Conference, Washington, DC, USA, 1–2 August 2018.

38. Duckworth, A.L.; Quirk, A.; Gallop, R.; Hoyle, R.H.; Kelly, D.R.; Matthews, M.D. Cognitive and noncognitive predictors of success. *Proc. Natl. Acad. Sci. USA* **2019**, *116*, 23499–23504. [[CrossRef](#)] [[PubMed](#)]
39. Powell, K. Higher education: On the lookout for true grit. *Nature* **2013**, *504*, 471–473. [[CrossRef](#)]
40. Hazari, Z.; Potvin, G.; Tai, R.H.; Almarode, J. For the love of learning science: Connecting learning orientation and career productivity in physics and chemistry. *Phys. Rev. Spec. Top. Phys. Educ. Res.* **2010**, *6*, 010107. [[CrossRef](#)]
41. Digman, J.M. Personality structure: Emergence of the five-factor model. *Annu. Rev. Psychol.* **1990**, *41*, 417–440. [[CrossRef](#)]
42. Damian, R.I.; Spengler, M.; Sutu, A.; Roberts, B.W. Sixteen going on sixty-six: A longitudinal study of personality stability and change across 50 years. *J. Personal. Soc. Psychol.* **2019**, *117*, 674. [[CrossRef](#)]
43. Jayawickreme, E.; Zachry, C.E.; Fleeson, W. Whole trait theory: An integrative approach to examining personality structure and process. *Personal. Individ. Differ.* **2019**, *136*, 2–11. [[CrossRef](#)]
44. Atherton, O.E.; Grijalva, E.; Roberts, B.W.; Robins, R.W. Stability and change in personality traits and major life goals from college to midlife. *Personal. Soc. Psychol. Bull.* **2021**, *47*, 841–858. [[CrossRef](#)]
45. Ponnock, A.; Muenks, K.; Morell, M.; Yang, J.S.; Gladstone, J.R.; Wigfield, A. Grit and conscientiousness: Another jangle fallacy. *J. Res. Personal.* **2020**, *89*, 104021. [[CrossRef](#)]
46. Shafir, E. Decisions in poverty contexts. *Curr. Opin. Psychol.* **2017**, *18*, 131–136. [[CrossRef](#)]
47. Mittal, C.; Laran, J.; Griskevicius, V. How early-life resource scarcity influences self-confidence and task completion judgments. *J. Assoc. Consum. Res.* **2020**, *5*, 404–414. [[CrossRef](#)]
48. Squire, D.D. Shifting Narratives in Doctoral Admissions: Faculty of Color Understandings of Diversity, Equity, and Justice in a Neoliberal Context. Ph.D. Thesis, Loyola University Chicago, Chicago, IL, USA, 2015.
49. Posselt, J.R. Toward inclusive excellence in graduate education: Constructing merit and diversity in PhD admissions. *Am. J. Educ.* **2014**, *120*, 481–514. [[CrossRef](#)]
50. Niessen, A.S.M.; Meijer, R.R.; Tendeiro, J.N. Measuring non-cognitive predictors in high-stakes contexts: The effect of self-presentation on self-report instruments used in admission to higher education. *Personal. Individ. Differ.* **2017**, *106*, 183–189. [[CrossRef](#)]
51. Geiger, M.; Olderbak, S.; Sauter, R.; Wilhelm, O. The “g” in faking: Doublethink the validity of personality self-report measures for applicant selection. *Front. Psychol.* **2018**, *9*, 2153. [[CrossRef](#)] [[PubMed](#)]
52. Wilson, M.A.; Odem, M.A.; Walters, T.; DePass, A.L.; Bean, A.J. A Model for Holistic Review in Graduate Admissions That Decouples the GRE from Race, Ethnicity, and Gender. *CBE Life Sci. Educ.* **2019**, *18*, ar7. [[CrossRef](#)] [[PubMed](#)]
53. Thompson, N.L.; Campbell, A.G. Addressing the challenge of diversity in the graduate ranks: Good practices yield good outcomes. *CBE Life Sci. Educ.* **2013**, *12*, 19–29. [[CrossRef](#)] [[PubMed](#)]
54. Lamont, M.; Da Silva, G.M. Complementary rather than contradictory: diversity and excellence in peer review and admissions in American higher education. *Twenty-First Century Soc.* **2009**, *4*, 1–15. [[CrossRef](#)]
55. Michel, R.S.; Belur, V.; Naemi, B.; Kell, H.J. Graduate admissions practices: A targeted review of the literature. *ETS Res. Rep. Ser.* **2019**, *2019*, 1–18. [[CrossRef](#)]
56. Feng, Z.; Liu, Y.; Wang, Z.; Savani, K. Let’s choose one of each: Using the partition dependence effect to increase diversity in organizations. *Organ. Behav. Hum. Decis. Process.* **2020**, *158*, 11–26. [[CrossRef](#)]
57. Chang, E.H.; Kirgios, E.L.; Rai, A.; Milkman, K.L. The isolated choice effect and its implications for gender diversity in organizations. *Manag. Sci.* **2020**, *66*, 2752–2761. [[CrossRef](#)]
58. Hoxby, C.; Turner, S. Expanding college opportunities for high-achieving, low income students. *Stanf. Inst. Econ. Policy Res. Discuss. Pap.* **2013**, *12*, 014.
59. Bettinger, E.P.; Long, B.T.; Oreopoulos, P.; Sanbonmatsu, L. *The Role of Simplification and Information in College Decisions: Results from the H&R Block EAFSA Experiment*; Working Paper #15361; National Bureau of Economic Research: New York, NY, USA, 2009.
60. Ortiz, N.A.; Morton, T.R.; Miles, M.L.; Roby, R.S. What about us? Exploring the challenges and sources of support influencing black students’ STEM identity development in postsecondary education. *J. Negro Educ.* **2019**, *88*, 311–326. [[CrossRef](#)]
61. Young, J.L.; Young, J.R.; Ford, D.Y. Standing in the gaps: Examining the effects of early gifted education on Black girl achievement in STEM. *J. Adv. Acad.* **2017**, *28*, 290–312. [[CrossRef](#)]
62. Hazari, Z.; Sadler, P.M.; Sonnert, G. The science identity of college students: Exploring the intersection of gender, race, and ethnicity. *J. Coll. Sci. Teach.* **2013**, *42*, 82–91.
63. James, M.; Bertschinger, E.; Beckford, B.; Dobbins, T.; Fries-Britt, S.; Gates, S.J.; Isler, J.; Ong, M.; Richardson, A.; Williams, Q.; et al. *The Time is Now: Systemic Changes to Increase African Americans with Bachelor’s Degrees in Physics and Astronomy*; Technical Report MSU-CSE-06-2, AIP National Task Force to Elevate African American Representation in Undergraduate Physics and Astronomy (TEAM-UP); American Institute of Physics: College Park, MD, USA, 2019.
64. McGee, E.; Bentley, L. The equity ethic: Black and Latinx college students reengineering their STEM careers toward justice. *Am. J. Educ.* **2017**, *124*, 1–36. [[CrossRef](#)]
65. Louis, D.A.; Rawls, G.J.; Jackson-Smith, D.; Chambers, G.A.; Phillips, L.L.; Louis, S.L. Listening to our voices: Experiences of Black faculty at predominantly White research universities with microaggression. *J. Black Stud.* **2016**, *47*, 454–474. [[CrossRef](#)]
66. Johnson-Bailey, J. Academic incivility and bullying as a gendered and racialized phenomena. *Adult Learn.* **2015**, *26*, 42–47. [[CrossRef](#)]