GARY WATERS/IKON IMAGES/GETTY

COLUMN

A test that fails

A standard test for admission to graduate school misses potential winners, say **Casey Miller** and **Keivan Stassun**.

Universities in the United States rely too heavily on the graduate record examinations (GRE) — a standardized test introduced in 1949 that is an admissions requirement for most US graduate schools. This practice is poor at selecting the most capable students and severely restricts the flow of women and minorities into the sciences.

We are not the only ones to reach this conclusion. William Sedlacek, professor emeritus of education at the University of Maryland, College Park, who has written extensively on the issue, notes that studies find only a weak correlation between the test and ultimate success in science, technology, engineering and maths (STEM) fields. De-emphasizing the GRE and augmenting admissions procedures with measures of other attributes — such as drive, diligence and the willingness to take scientific risks — would not only make graduate admissions more predictive of the ability to do well but would also increase diversity in STEM.

TEST DISPARITIES

The GRE, like most standardized tests, reflects certain demographic characteristics of test-takers — such as family socioeconomic status — that are unrelated to their intellectual capacity or academic preparation. The exam's 'quantitative score' — the portion measuring maths acumen, which is most commonly scrutinized in admissions to STEM PhD programmes — correlates closely with gender and ethnicity (see 'The great divide'). The effect is powerful. According to data from Educational Testing Service (ETS), based in Princeton, New Jersey, the company that administers the GRE, women score 80 points lower on average in the physical sciences than do men, and African Americans score 200 points below white people. In simple terms, the GRE is a better indicator of sex and skin colour than of ability and ultimate success.

These correlations and their magnitude are not well known to graduate-admissions committees, which have a changing rota of faculty members. Compounding the problem, some admissions committees use minimum GRE scores to rapidly filter applications; for example, any candidate scoring below 700 on the 800-point quantitative test section may be discarded. Using GRE scores to filter applicants in this way is a violation of ETS's own guidelines.

This problem is rampant. If the correlation between GRE scores and gender and ethnicity is not accounted for, imposing such cut-offs adversely affects women and minority applicants. For example, in the physical sciences, ▶

only 26% of women, compared with 73% of men, score above 700 on the GRE Quantitative measure. For minorities, this falls to 5.2%, compared with 82% for white and Asian people.

The misuse of GRE scores to select applicants may be a strong driver of the continuing under-representation of women and minorities in graduate school. Indeed, women earn barely 20% of US physical-sciences PhDs, and under-represented minorities — who account for 33% of the US university-age population — earn just 6%. These percentages are striking in their similarity to the percentage of students who score above 700 on the GRE quantitative measure.

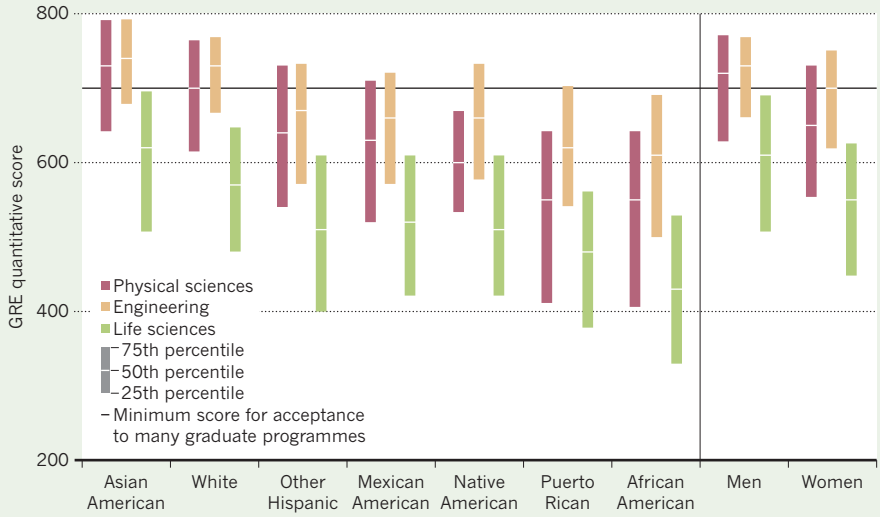
Why is the GRE misused? Admissions committees are busy, and numerical rankings are easy to sort. We believe that faculty members also often presume that higher scores imply that the test-taker has a greater ability to become a PhD-level scientist. Yet research by ETS indicates that the predictive validity of the GRE tests is limited to first-year graduate-course grades, and even that correlation is meagre in maths-intensive STEM fields.

Why should graduate-admissions committees care about fixing the problem? First, diversity, in the form of individuals with different perspectives, backgrounds and experiences, is a key component of innovation and problem solving, a concept that business and industry have come to recognize. Less diversity in STEM graduate programmes means slower progress in tackling today's scientific and technical challenges. Second, the overall PhD completion rate in US STEM graduate programmes is a disappointing 50%. Although graduate programmes certainly produce successful students who continue on to productive science careers, we think that many faculty members would agree that such a low PhD completion rate is a poor return on the investment in recruiting and training students. Indeed, STEM graduate programmes are failing not only from the diversity standpoint, but also from a success standpoint.

THE GREAT DIVIDE

The data represent the scores typically achieved in the quantitative reasoning test of the graduate record examinations (GRE) by US students from different ethnic groups applying for graduate school. In the physical sciences, a minimum score of 700 is required by many PhD programmes.

SOURCE: ETS



ALTERNATIVE SELECTION

So what should universities do? Instead of filtering by GRE scores, graduate programmes can select applicants on the basis of skills and character attributes that are more predictive of doing well in scientific research and of ultimate employability in the STEM workforce. Appraisers should look not only at indicators of previous achievements, but also at evidence of ability to overcome the tribulations of becoming a PhD-level scientist.

A few innovative PhD programmes, including the bridge programmes at the University of South Florida in Tampa and Fisk–Vanderbilt in Nashville, Tennessee (in which we are involved) are doing this. They have achieved completion rates above 80%, well above the national average, and are greatly boosting participation by women and minorities (see *Nature* **504,** 471–473; 2013). The admissions process includes an interview that examines college and research experiences, key relationships, leadership experience, service to community and life goals. The result is a good indication of the individual's commitment to scientific research and a good assessment of traits such as maturity, perseverance, adaptability and conscientiousness atop a solid academic foundation. The combination of academic aptitude and these other competencies points to the likelihood of high achievement in graduate school and in a STEM career.

How have the students admitted to these courses performed? In the Fisk–Vanderbilt programme, 81% of the 67 students who have entered the programme — including 56 under-represented minorities and 35 women — have earned, or are making good progress towards, their PhDs. And all students who have completed PhDs are employed in the STEM workforce as postdocs, university faculty members or staff scientists in national labs or industry. From the standpoint of optimal outcomes — earning a PhD and obtaining employment in the STEM workforce — the GRE has proved irrelevant. Indeed, 85% of these young scientists would have been eliminated from consideration for PhD programmes by a GRE quantitative cut-off score of 700.

The only downside is that interviews take about 30 minutes each. But the number of interviews need not be large, and the tremendous insight garnered justifies the time. ETS is even marketing a tool for referees to evaluate applicants' personal attributes. The company developed it in part as a response to calls from applicants and graduate programmes for alternative measures of student potential for long-term achievement that is not captured by GRE.

We often hear admissions committee members say, 'We would admit women and minorities if they were qualified'. This mindset reflects long-standing admissions practices that systematically, if inadvertently, filter out women and minorities. At the same time, these practices are no better than a coin flip at identifying candidates with the potential — and the mettle — to earn a PhD.

Let us be frank: we believe that many STEM faculty members on admissions committees and upper-level administrators hold a deep-seated and unfounded belief that these test scores are good measures of ability, of potential for doing well in graduate school and of long-term potential as a scientist, and that students who score poorly on standardized exams are not likely to become PhD-level scientists. These assumptions are false.

"In simple terms, the GRE is a better indicator of sex and skin colour than of ability and ultimate success."

This is not a call to admit unqualified students in the name of social good. This is a call to acknowledge that the typical weight given to GRE scores in admissions is disproportionate. If we diminish reliance on GRE and instead augment current admissions practices with proven markers of achievement, such as grit and diligence, we will make our PhD programmes more inclusive and will more efficiently identify applicants with potential for long-term success as researchers. Isn't that what graduate school is about? ■

Casey Miller *is an associate professor in the physics department at the University of South Florida in Tampa.* **Keivan Stassun** *is a professor in physics and astronomy at Vanderbilt University and Fisk University in Nashville, Tennessee.*