How should we be selecting our graduate students?

Orion D. Weiner
Cardiovascular Research Institute, Department of Biochemistry and Biophysics, University of California, San Francisco, San Francisco, CA 94158

ABSTRACT We use many quantitative undergraduate metrics to help select our graduate students, but which of these usefully discriminate successful from underperforming students and which should be ignored? Almost everyone has his or her own pet theory of the most predictive criteria, but I hoped to address this question in a more unbiased manner. I conducted a retrospective analysis of the highest- and lowest-ranked graduate students over the past 20 years in the Tetrad program at the University of California at San Francisco to identify undergraduate metrics that significantly differed between these groups. Only the number of years of research experience and subject graduate record exams (GREs) were strong discriminations between the highest- and lowest-ranked students, whereas many other commonly used admissions metrics (analytical, verbal, and quantitative GREs, grade point average, and ranking of undergraduate institution) showed no correlation with graduate performance. These are not necessarily the same criteria that matter at other graduate programs, but I would urge faculty elsewhere to conduct similar analyses to improve the admissions process and to minimize the use of useless metrics in selecting our students.

INTRODUCTION

As cochair of graduate admissions of the Tetrad program at the University of California at San Francisco (UCSF), I conducted a retrospective analysis of our highest- and lowest-ranked graduate students over the past 20 years to identify undergraduate metrics that significantly differed between these groups, I interviewed 30 core faculty members with a significant history of Tetrad graduate students and asked them to identify the very best versus most underperforming students they have known over the past two decades from their own labs, thesis committees, rotations, and so on. Because different faculty members have different criteria for ranking students, I next whittled this list down to unanimous highest- and lowest-ranked groups by interviewing the students’ thesis committees and other core faculty members and removed students for whom there was any disagreement. From this I obtained a unanimous cohort of 31 highest-ranked and 21 lowest-ranked students, and I analyzed how various undergraduate metrics differed between the groups.

RESULTS

1. The single largest discriminator of the highest-ranked versus lowest-ranked group was the number of years of previous research experience (p < 0.001). More than half (52%) of our lowest-ranked group versus 0% of our highest-ranked group had performed <2 years of previous research by the time they interviewed. It is likely that students who have done <2 years of research do not know what they are getting into. Furthermore, the letters of recommendation (our dominant criteria for choosing students) are more meaningful when written by principal investigators who have interacted with the students for long enough to know them well. Beyond a point, more research did not necessarily correlate with higher performance. The same proportion of highest-ranked versus lowest-ranked students did more than 3 years of research, but too little research is bad. On an educational note, making sure that students who are bound for experimental work in graduate school get significant hands-on exposure to real, not just course-based, lab work should be a high priority component of their undergraduate training.

2. None of the standard graduate record exams (GREs) exhibited a highly significant difference between the highest- and lowest-ranked students. The means of the verbal, analytical, and quantitative GREs, grade point average, and ranking of undergraduate institution showed no correlation with graduate performance.
To the extent that we use universal metrics such as GPA, GRE, and the amount of research experience, it is worth noting that some of these are significantly more discriminating than the others. Figure 1B shows the percentage of lowest-ranked students that are excluded for thresholds that capture 90% of the highest-ranked students for each selection criteria. Only number of years of previous research and subject GREs significantly enrich for highest-ranked students. These are not necessarily the same criteria that will matter at other schools, but I would urge faculty elsewhere to conduct similar analyses to improve the admissions process and to minimize the use of useless metrics in selecting our students.

3. There were no significant differences in grade point averages (GPAs) between our highest- and lowest-ranked students, although it is important to note that we generally only accepted students with GPAs of 3.0 or above.

4. There were no significant differences between the proportion of the highest-ranked (45%) versus lowest-ranked (38%) students that came from top 10 life sciences universities, as ranked by U.S. News & World Report and the National Research Council.

Of course, you cannot simply select a class with these numbers alone. Other factors, such as the strength of the letters (particularly from the primary research advisor) and the quality of the applicant’s previous research and essays are massively important. These were strong for all of the students that we accepted and were (and continue to be) our primary criteria for selecting interviewees.

To the extent that we use universal metrics such as GPA, GRE, and the amount of research experience, it is worth noting that some of these are significantly more discriminating than the others. Figure 1B shows the percentage of lowest-ranked students that are excluded for thresholds that capture 90% of the highest-ranked students for each selection criteria. Only number of years of previous research and subject GREs significantly enrich for highest-ranked students. These are not necessarily the same criteria that will matter at other schools, but I would urge faculty elsewhere to conduct similar analyses to improve the admissions process and to minimize the use of useless metrics in selecting our students.

ACKNOWLEDGMENTS

I thank the faculty, staff, and Tetrad admission committee at UCSF for their help in identifying highest- and lowest-ranked graduate students and Rahul Deo for thoughtful discussion of the results and for helping with the analysis used in Figure 1B. This work was supported by National Institutes of Health Grant R01 GM084040.