Response to Bell et al.

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

Thanks for your interest (see Bell et al., 2014) in my Perspective piece. I am happy to provide more information on the methodology behind and rationale for my analysis.

DEFINING SUCCESS

I asked faculty to identify students who had the most and least positive impact on their labs and University of California, San Francisco (UCSF), based on student performance during their time in graduate school. I thought it was best to compare students when they are all at the same stage in their careers. Furthermore, many of these students have not yet completed their graduate or postgraduate training. I strongly agree that we should be training our students for diverse careers (not just academia), but importantly, my study was based on graduate performance not on postgraduate outcome.

A typical highest-ranked student independently initiated a new area of research for the lab/field and contributed to the success of others at UCSF through collaborations, sharing ideas and reagents, and so on. A typical lowest-ranked student significantly struggled during his or her time in graduate school, possibly due to a lack of appetite and/or aptitude for scientific research. I agree that there are many possible metrics for success, but it would be hard to argue that you would not want to enrich for the first set of students at the expense of the latter.

METHODOLOGY

I chose to interview faculty rather than solicit email or Web-based responses, because the response rates for the latter are typically significantly lower than in-person interviews. I was able to interview more than 95% of the current core faculty at UCSF, which is likely my study, which required independent confirmation of highest- and lowest-ranked students from multiple faculty, would have been underpowered for email or Web-based polls.

In the same way that letters of recommendation were our most meaningful data points for undergraduate applicants to our graduate program, I would argue that faculty evaluations of graduate students are the single most meaningful metric to evaluate our current and past graduate students. The faculty who work directly with the students are in the best position to assess student contributions to published work and all of the other important student contributions that never make it into papers. I think this is vastly preferable to ranking students based on number of published papers, journals in which they published, or other metrics. This strategy had the further advantage of using the collective experience of the faculty to rapidly converge on a highest- and lowest-ranked group. Also including an outside set of experts would have been even better, but this would necessitate independently evaluating each and every student, which was significantly beyond my resources.

PRIVACY CONCERNS

Besides me, no one else was given the list of highest- and lowest-ranked students. I kept the list only long enough to pull the relevant student metrics, which were then deposited into deidentified lists (one randomized list for highest-ranked student grade point average, a separate randomized list for highest-ranked student verbal Graduate Record Examinations, etc.). I excluded myself from the analysis. I received approval for this protocol from the Committee for Human Research (approved protocol 13-12542) and this study also adhered to the standards of the Family and Educational Rights and Privacy Act.

OUTCOMES

During my time as cochair of graduate admissions for the Tetrad program, our proportion of underrepresented minorities significantly increased, and a significant proportion of our class is from socioeconomically disadvantaged backgrounds. This is primarily a testament to the thoughtful efforts of the faculty and staff at UCSF, but emphasizing the importance of previous research experience has certainly not been incompatible with a diverse class. In fact, deemphasizing most of the standard metrics has enabled us to consider many candidates we might have previously overlooked.

I am glad we are having this discussion. On the whole, it is remarkable how nonscientific we have been in choosing the next generation of graduate students. I agree that much more could and should be done. In particular, a major value of these retrospective studies will be to enable programs to address these issues prospectively, using these and other criteria, and I strongly encourage others to expand on this work.

REFERENCE