HIGHER EDUCATION

US school ranking names no winners

Graduate programmes are assessed and measured, but stale data could reduce impact of long-awaited report.

BY EMMA MARRIS

Which US chemistry department is the biggest? As of autumn 2005, the University of California, Berkeley, had a whopping 406 graduate students. That must be some departmental picnic. Which ecology programme takes the longest? The median time to complete a PhD degree in the ecology and evolutionary biology department at Tulane University in Louisiana is 8.5 years. Which genetics programme has the highest average number of citations per faculty publication? The Massachusetts Institute of Technology in Cambridge dominates, with a knockout 10.08. Which physics programme is the best? A new report that supplies all of the other answers doesn't make the call.

Released on 28 September, the long-awaited National Academies study on US PhD programmes, *A Data-Based Assessment of Research-Doctorate Programs in the United States* (see go.nature.com/tqvokc), is notable for not ranking programmes in 1-2-3 order. But it aims to offer comparisons that are detailed enough both to help students determine where to apply and to help job-seekers judge offers. The findings could also guide spending by administrators at a state or school level — whether by lavishing funds on standout programmes or by spending money to improve less-successful ones.

The report was delayed by funding problems, and the National Research Council had to charge institutions up to US$10,000 apiece to be included. The underlying data are now five years old, which could limit the report's impact. But it is accompanied by a huge trove of raw data, which can be manipulated to answer specific questions. And the rankings are less subjective than previous versions of the report, the last of which appeared in 1995. "We thought doing it right was more important than doing it fast," says the report's committee chair Jeremiah Ostriker, an astronomer at Princeton University in New Jersey.

The new rankings derive from quantitative measures, such as publications or citations per faculty member, weighted in two different ways. In one scheme, members of a field were asked to evaluate the importance of various measures. In the other, the specialists had to rank programmes, and statistical analysis determined the weights that various measures would have to be given to reproduce those rankings. "It is not really based on reputation, it is based on the things that seem to predict reputation," says Ostriker.

The two methods produced subtle differences (see 'Grading the schools'). For example, although few faculty members stressed the importance of programme size, they tended to give higher rankings to the programmes that awarded many PhDs. Both ranking schemes, however, gave surprisingly little importance to other measures. "How well the students are taken care of and how well they do after they graduate is obviously important, but it isn't what the faculty put the most emphasis on," says Ostriker. "They care more about the research output of the faculty."

Each programme's position is expressed as a range rather than an average to communicate the uncertainties and fluctuations in the data. The overall result is a lot of data — 20 variables for more than 5,000 programmes at 212 universities — but no clear 'winners'. "The committee believes that the concept of a precise ranking of doctoral programs is mistaken," the report reads. "The reader who seeks a single, authoritative declaration of the 'best programs' in given fields will not find it in this report."

> ***"We thought doing it right was more important than doing it fast."***

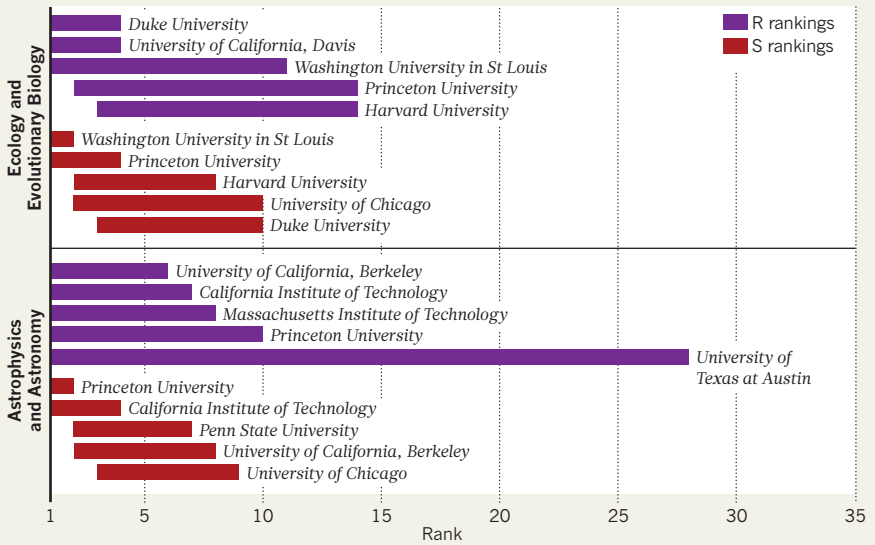
Harvey Waterman, associate dean for academic affairs at the graduate school of Rutgers University in New Brunswick, New Jersey, who helped to advise on the surveys used by the project, predicts a fair amount of nitpicking about old data and new methodology. For example, 'interdisciplinarity' is measured by how many of a programme's faculty members are listed as 'associate'. Programmes that were interdisciplinary by nature scored zero because their faculty are full members (not associates) regardless of speciality.

Debra Stewart, president of the Council of Graduate Schools in Washington DC, calls the report's two ranking systems and the ranges of outcomes "perplexing in a very healthy way". For Stewart, the varied rankings prove that different criteria make sense for different programmes, depending on their priorities. A school that prides itself on diversity might focus on the various measures of faculty and student diversity; a school that has no plans for expanding a small programme might compare itself only with other small programmes. The fact that the data are a bit stale, she says, "only becomes important if there is no effort to update this on a regular basis".

Ostriker says that this is on the cards. "We hope that in a couple of years we can get data on new faculty and then repeat it. That is the only thing that changes very quickly." ■

GRADING THE SCHOOLS

R rankings (based on a statistical model derived from from faculty members' opinions) and S rankings (based on the criteria that faculty members named as important) produced different lists of the top-five programmes in two different fields. To reflect uncertainties, the report expressed each programme's position as a range.



SOURCE: NATIONAL ACADEMIES

© 2010 Macmillan Publishers Limited. All rights reserved