

“Anchoring” the World University Rankings

Nicholas A. Bowman and Michael Bastedo

Nicholas A. Bowman is assistant professor, at Bowling Green State University, Ohio. E-mail: nbowman@bgsu.edu. Michael Bastedo is associate professor, at University of Michigan, Ann Arbor, MI. E-mail: bastedo@umich.edu. This article is adapted (with permission) from *Higher Education*, Anchoring Effects in World University Rankings: Exploring Biases in Reputation Scores; more can be read about their rankings research at <http://www.umich.edu/~bastedo>.

Indeed institutional leaders are quite concerned about their position in rankings and league tables. They believe that many constituents within and outside of higher education rely on international or national rankings to reveal which universities are “good” or “great.” If this is true, then these rankings may become virtually synonymous with institutional reputation over time.

It is suspected that rankings might influence reputation through a phenomenon known as the anchoring effect (or anchoring-and-adjustment heuristic). That is, people often rely on some starting point when making judgments for which the answer is ambiguous, such as estimating the number of jelly beans in a jar. For example, people who are initially asked whether the number of jelly beans is higher or lower than a high number (e.g., 2,000) will

generally provide higher estimates for the exact value than those who are initially asked whether it is higher or lower than a low number (e.g., 100). Research has demonstrated that these anchoring effects can occur even among experts in a variety of important, real-world settings, such as assessing real estate values and determining the length of prison sentences.

Until less than a decade ago, the reputations of diverse universities throughout the world were also characterized by considerable uncertainty. While many people would have agreed that Oxford and Harvard are in the top 10, where should the University of Melbourne or Tsinghua University be ranked? And which of these two schools is “better”? Therefore, the introduction of *Times Higher Education Supplement (THES) World University Rankings* in 2004 provided a natural experiment to consider the effect of rankings on the decision-making processes of reputational arbiters around the world. After the first set of *THES* rankings, people who wanted to identify “top” schools—or to see where their own institution stacked up—had a formal list that they could peruse. When asked to identify top schools in future surveys, some reputational raters probably referred to the list directly; and others who had seen the list previously may have reported more favorable ratings for the institutions at the top of the *THES* rankings.

EXAMINING ANCHORING EFFECTS ON UNIVERSITY REPUTATION

To explore empirically whether the early *THES* world university rankings provided an anchoring point for subsequent assessments of institutional reputation, this article examines the first three years of the *THES* rankings and

the corresponding reputation ratings, which were provided by a variety of academics worldwide (e.g., faculty, administrators, lecturers).

Three hypotheses were supported by the data. First, consistent with anchoring theory, the initial *THES* rankings were associated with increases in reputational assessments. Thus, the academics who were asked to provide reputation ratings in the second and third years of the survey appear to have been influenced—whether consciously or unconsciously—by the first year of rankings. Furthermore, the relationship between rankings and reputation was much stronger in the second year of the survey than in the first year, which means that rankings and reputation became more closely aligned in a very short period of time.

We had no reason to anticipate that the opposite pattern would occur (namely, that reputation would lead to improvements in rankings), and we did not find such a link. This “non-finding” supports the interpretation that rankings cause improvements in reputation, not vice versa.

Because the first year of rankings is particularly informative for potential raters (by providing a novel formalized hierarchy among institutions), we expected that the effects of rankings on reputation would be strongest for the first year of rankings than for the second year. Indeed, the first year of rankings was associated with increased reputation in the following two years, but the second year of rankings had no significant effect.

CONCLUSION AND IMPLICATIONS

One of the many uses of rankings is to provide useful information to consumers as they make decisions about college choices. For many years, there has been

concern about the use of rankings based on equity and validity. The reputation assessments of the rankings have come under particular criticism, much of it well deserved.

Our research provides fuel for many of these concerns. It seems highly likely that the anchoring effects we found are influential across ranking schemes worldwide, as this is based on a well-established psychological effect. When also considering the fact that the differences between rankings and reputation are becoming vanishingly small over time and that rankings are stable over time, it is difficult to maintain the fantasy that reputational scores are independent from the rankings themselves.

Engineering effective reputational surveys is a difficult proposition. As mentioned earlier, one of the major issues is that respondents are asked to rate colleges about which they have little first-hand knowledge, and there are longtime lags between changes in quality and subsequent reputation. Therefore, one solution would be to ask respondents only to rate universities about which they have deep knowledge. Unfortunately, this will likely generate a conflict of interest: Universities generally have the deepest knowledge of their closest competitors, and these institutions compete for higher rankings.

Indeed, manipulation of reputational surveys may be a major problem with some college-rankings systems. An organization that conducts college rankings must have specific criteria for identifying survey responses that show evidence of manipulation, and it must ensure that informants have deep knowledge of the institutions they rank. Moreover, sampling a large sample of faculty from many diverse institutions and from a wide range of academic fields would improve the validity of the survey results. And because reputational

surveys are likely here to stay, it is important to use empirical research to make them as informative and unbiased as possible.

Simply eliminating reputational surveys is not the answer. Our “objective” measures of institutions—like those used in the *Shanghai Academic Ranking of World Universities*—are not strong enough to provide a real sense of academic quality. In addition, the more that objective measures are weighted, the more likely it is that institutions will simply purchase the Nobel prize winners and highly cited scientists that they need to rise in the rankings, with fairly negligible improvements in academic program quality more broadly.

Finally, there is simply a demand for knowledge about institutional reputation, and someone will be there to provide it. It is better to encourage those who seek to do it well than to leave it to those who will do it poorly.

We think the new World University Rankings are a step in the right direction. By asking people to rate programs in their own discipline, it is more likely that the ratings will be valid and reliable. Academics keep tabs on one another quite well in their own area of interest, and rankings should take advantage of that. With sociologists rating sociologists and physicists rating physicists, you have the best chance of ascertaining a professional consensus about program quality.