



New tools for measuring academic performance

Numbers and data drive decisions in sports, business, and other fields. Yet in academia, publication lists are still the primary basis for hiring, promotion, and funding decisions. Some administrators and researchers are adopting new methods for assessing academic achievement, however. Large universities may subscribe to comprehensive, costly vendor-based performance panels. Other institutions and individuals design their own algorithms and dashboards. **By Chris Tachibana**

One question guides top scientific institutions in making hiring and promotion decisions, and funders in awarding support, says **Ushma Neill**: “Will the candidate move the needle forward in their field?” Neill, vice president of scientific education and training at Memorial Sloan Kettering Cancer Center (MSKCC), says the bottom line is that organizations want faculty who advance science. Traditionally, she says, methods for evaluating a candidate’s potential have been “largely intuitive.” But that’s changing.

Scientific impact is still mainly measured with publications. Universities don’t set numbers of papers required for hiring or promotion, though, because strict counts don’t reflect author contribution or whether a publication is a review or original findings. In addition, fields vary in how work is presented. Some emphasize peer-reviewed journals or conference proceedings, others books or monographs. Informal discussions suggest annual goals ranging from a few publications for postdocs to five or more for faculty. However, quality matters more than quantity, says Neill. A comprehensive, well-cited study stands out more than several minor articles pulled from a single project.

Academic impact is also measured through advisor and peer recommendations and awards for research, mentoring, and

teaching, says **Morten Pejrup**, associate dean for research in the Faculty of Science at the University of Copenhagen, Denmark. But these “old school methods,” he says, can be supplemented with new tools.

Going beyond traditional measures

When evaluating a job candidate, Pejrup checks Scopus, a subscription database of peer-reviewed journal articles, books, and conference proceedings. “You see citations, journal impact factors, and order of authors,” he says. “But you also see if the person always publishes with the same people or has contributed to different groups.” Publication databases such as Scopus or Web of Science may include bibliometrics such as *h*-index and *m*-index, which combine measures of publications and citations to indicate influence.

Trends in measuring academic performance are driven in part by funders, says **Trine Buhl Monty**, senior executive consultant for the University of Copenhagen’s Faculty of Science. For example, the United Kingdom recently completed the Research Excellence Framework 2014, an initiative to determine the research quality and economic and social impact of its universities. Among its other uses, the evaluation will influence funding. Horizon 2020, the European Union’s biggest funding program for research and innovation, emphasizes impact related to intersectoral partnerships and open science, to ensure influence beyond the lifetime of a project, Buhl Monty says. In response, some faculty now list industry collaborations, open-access publications, and open-source software as “research products.”

Cathy Sarli, a librarian at Washington University School of Medicine in St. Louis, specializes in research impact assessment. She studies the diversification of measures beyond publications, citations, and journal impact factors (calculated as the average number of times articles published over a two-year period are cited in a given journal). Impact factors, she says, were developed in the mid-20th century to compare peer-reviewed journals, not individual researchers. “Organizations are going beyond traditional raw numbers,” she says. “They’re utilizing other metrics and developing their own resources.” **cont.>**

Upcoming features

Postdocs: Tech Transfer—March 17 ■ **Postdocs: Unusual Collaborations—August 25** ■ **Faculty: Managing Multigenerational Teams—September 15**

faculty



“[Metrics] enhance recommendation letters, teaching experience, funding success, and institutional service.” – Ushma Neill, vice president of scientific education and training, Memorial Sloan Kettering Cancer Center

One example is the Relative Citation Ratio (RCR), recently developed by the U.S. National Institutes of Health (NIH) to measure an article’s influence. This year, the NIH released the free *iCite* tool for calculating this metric. The RCR normalizes citation rates to other papers in the same field. This allows comparisons among scientists with different specialties, which are often necessary when allocating internal resources, endowed chairs and leadership positions, and funding. These decisions can require comparing “apples and oranges, or sometimes apples and zucchini,” says Neill.

In 2015, Neill and two MSKCC colleagues published their method for considering researchers in different areas and stages in their careers, using data from Scopus and other sources. She explains, “We wanted to create a level playing field when comparing, for example, the productivity and impact of a structural biologist and a surgeon.” Neill points to another trendsetter in tracking academic productivity: New York University (NYU) Langone Medical Center.

Developing dashboards and datasets

The story of the NYU Langone dashboard system for tracking academic and clinical performance begins with the appointment of Robert Grossman as dean and CEO in 2007. Grossman wanted standardized reports to guide strategic decisions. NYU Langone’s information technology group built a system that presents data from across the organization. Graphical displays show clinical measures, research performance, educational statistics, and finances. **Laura Ahlborn**, vice president for research enterprise and strategy, says that for measuring basic research, the dashboard has metrics in four areas: grants, publications, innovations (such as patents), and efficiency (for example, productivity per full-time equivalents). Administrators use the dashboard to track the performance of departments, but it’s also helpful for individual faculty members.

“We use dashboard metrics to drive strategic decision-making, to identify where we have strengths and weaknesses, and to guide where to invest our time and resources,” says **Dafna**

Bar-Sagi, NYU Langone’s senior vice president and vice dean for science, and its chief scientific officer. “The dashboard shows performance over time, so it’s also useful for monitoring the career trajectory of faculty and advising them accordingly.”

Ahlborn and Bar-Sagi emphasize that the dashboard is used in combination with human judgment. “Academic endeavors can’t be judged strictly by numbers,” says Bar-Sagi. “So the dashboard is one of several elements of an ongoing dialog between faculty and leadership.”

Both say that interest in tools for measuring academic impact is high and that they often field requests for information about the dashboard. Not surprisingly, it’s a lot of work. Setting up the dashboard required an initial investment in infrastructure and decisions about what data to collect and how systems would talk to each other. The dashboard is continually updated, including moving to add the RCR as a metric. Developing and maintaining this resource works, Ahlborn says, because NYU Langone has a culture of quantitative, precise, data-driven decision-making.

Tracking trainees

Starting in 2006, the Danish government decided to invest in increasing the intake of Ph.D. students at its universities. Schools have used the data on the increase, much of which is self-reported, to track how their graduates are filling jobs in education, research, government, and other sectors. Because of public concern about using taxpayer funds to educate non-Danes, the universities also want to know where international students go and what jobs they get.

“We must be transparent and show policy-makers and taxpayers what we are producing,” Pejrup says. “The Faculty of Science is responsible for 40 percent of the university’s budget.” Pejrup’s unit has now generated two substantial reports with longitudinal data on Master’s and doctoral graduates: the jobs they took, the countries they worked in, and their salaries. It could inform graduate curricula, training strategies, and policy, for example, by showing if the country is training the appropriate number of Ph.D.s for the national and international job market.

Buhl Monty says that the reports are also useful for guiding Ph.D. students by showing career options outside of academia. For students committed to a faculty position, the data indicate the importance of being able to move to another country for a job. “It shows students what’s possible,” she says. “It adjusts their expectations.”

Pejrup is also implementing a measure of scientific contribution that he would like to see become widespread. “More large consortia are producing publications,” he says, “such as in physics or genomics collaborations. Unless someone is first, last, or corresponding author, it’s hard to know what they contributed.” In fields such as bioinformatics, publications can have long author lists. Knowing who did what is critical for evaluating individual collaborators.

Pejrup would like to see descriptions of scientific contribution become more standard in academic reporting. He started with a requirement that Ph.D. theses that insert publications as chapters include a description of the student’s contribution. “Many journals now require this description,” Pejrup says. “Having it be public information would help people evaluate the authors’ work.” **cont.>**

PHOTO: COURTESY OF MEMORIAL SLOAN KETTERING CANCER CENTER

faculty

Featured participants

<p>American University in Cairo www.aucegypt.edu</p> <p>Faculty of Science, University of Copenhagen www.science.ku.dk/english</p> <p>Memorial Sloan Kettering Cancer Center www.mskcc.org</p> <p>NYU Langone Medical Center nyulangone.org</p> <p>School of Aquatic and Fishery Sciences, University of Washington fish.uw.edu</p>	<p>Washington University School of Medicine in St. Louis medicine.wustl.edu</p> <p>Additional resources</p> <p>Altmetric www.altmetric.com</p> <p>iCite (Relative Citation Ratio) icite.od.nih.gov</p> <p>Impactstory impactstory.org</p> <p>Web of Science ipscience.thomsonreuters.com/product/web-of-science</p>
--	---



Laura Ahlborn



Ahmed Moustafa

Individual scientists are also developing their own performance metrics. At the University of Washington School of Aquatic and Fisheries Sciences, Associate Professor **Trevor Branch** created a database of the most cited papers in his field. It includes a metric for estimating individual contributions to studies with multiple authors. The calculation is not automated but has been tested in a 2016 proof-of-concept paper. Branch says the database is mostly a hobby, but he finds it strategically useful. “I’m a modeler and a data analyst,” he says. “The database helps me look at the big picture—where can I have the most impact, what fields are of interest to other people, what should I work on, and where should I send my papers?”

Including infographics and social media

Adding metrics to your own research portfolio is simple, says **Ahmed Moustafa**, associate professor of bioinformatics and genomics at the American University in Cairo. His 2013 tenure dossier included data from Impactstory, a web-based tool for measuring shares and mentions of a research product. The publication list in his online curriculum vitae displays the dissemination of his publications as colorful donut graphs from the digital science company Altmetric. These measures reflect public outreach and engagement, an area that funders and institutions now include in academic performance evaluations.

Moustafa believes the information strengthened his dossier. “It was well received,” he says. “My committee used it when presenting my tenure case to the department and the school. It gave information about the performance of my research in the scientific community and an assessment of the global visibility and significance of my work.”

Setting up Impactstory and Altmetric tools doesn’t take much time and effort, Moustafa says, and updates are automatic. He has also used the information to introduce himself to potential collaborators. In his field of life science, he says it’s not unusual for grant proposals to include Altmetrics or Impactstory profiles among other measures of productivity.

Branch has included social media activity in the engagement plans of his grant applications, but reports a mixed reception

from reviewers. Outreach and engagement can be done in many ways, he says. Some people might gravitate to using social media to interact with the community, but others might prefer public events, blogging, or another method. “Learn what style of scientific outreach works for you,” he says. But having an engagement outlet in place is an advantage when the opportunity arises to promote your work.

When Branch’s student, Cole Monnahan, published a 2014 study about blue whales, they knew the general public would be interested, so Branch promoted it on his social media accounts. “It really took off,” he says, resulting in a BBC interview that was picked up by other news sources. The results reached a lot of people, but the approach worked because the topic was large marine mammals. “The same student has a statistical methods paper coming out that will probably have a bigger impact on the field,” says Branch, “but it just isn’t going to be as newsworthy.”

Using metrics correctly: Always put them in context

As new measures of academic performance appear, experts stress that quantitative indicators are only part of what counts as scientific impact. “These metrics are never used in a vacuum,” says Neill. “They enhance recommendation letters, teaching experience, funding success, and institutional service.” A solid publications profile can lead to a candidate interview, Pejrup says, but the interviewee’s ability to describe research goals and impact in conversation is what really counts.

Sarli agrees that numbers and graphs are powerful but should always be presented with context. When using bibliometrics or other performance indexes, she advises providing a description of the measure and its limitations. Quantitative indicators, she says, should always be part of a narrative about your scientific influence. “No single metric is sufficient for measuring performance, quality, or impact,” she says. “Go beyond numbers to tell your story.”

Chris Tachibana is a science writer based in Seattle, USA, and Copenhagen, Denmark.

DOI: 10.1126/science.opms.r1700173

PHOTOS: (FROM LEFT) NYU LANGONE MEDICAL CENTER, ALAA SOLIMANI