



Do metrics matter?

Many researchers believe that quantitative metrics determine who gets hired and who gets promoted at their institutions. With an exclusive poll and interviews, *Nature* probes to what extent metrics are really used that way.

No scientist's career can be summarized by a number. He or she spends countless hours troubleshooting experiments, guiding students and postdocs, writing or reviewing grants and papers, teaching, preparing for and organizing meetings, participating in collaborations, advising colleagues, serving on editorial boards and more — none of which is easily quantified.

But when that scientist is seeking a job, promotion or even tenure, which of those duties will be rewarded? Many scientists are concerned that decision-makers put too much weight on the handful of things that can be measured easily — the number of papers they have published, for example, the impact factor of the journals they have published in, how often their papers have been cited, the amount of grant money they have earned, or measures of published output such as the *h*-index.

Last month, 150 readers responded to a *Nature* poll designed to gauge how researchers believe such metrics are being used at their institutions, and whether they approve of the practice. *Nature* also contacted provosts, department heads and other administrators at nearly 30 research

institutions around the world to see what metrics are being used, and how heavily they are relied on. The results suggest that there may be a disconnect between the way researchers and administrators see the value of metrics.

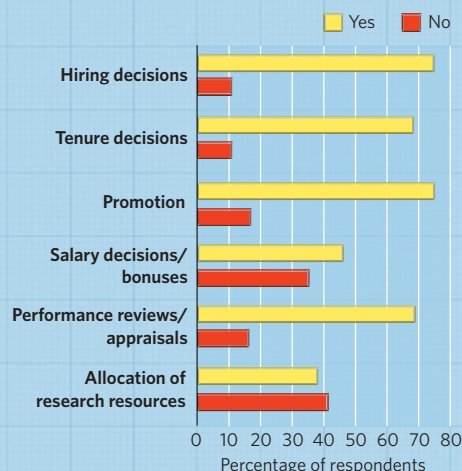
Three-quarters of those polled believe that

metrics are being used in hiring decisions and promotion, and almost 70% believe that they are being used in tenure decisions and performance review (see 'Metrics perceptions'). When asked to rate how much they thought administrators were relying on specific criteria for evaluation, poll respondents indicated that the most important measures were grants and income, number of publications, publication in high impact journals and citations of published research. And a majority (63%) are unhappy about the way in which some of these measures are used (see 'No satisfaction'). "Too much emphasis is paid to these flawed, seemingly objective measures to assess productivity," wrote a biologist from the United States. Respondents doubted that traditional, qualitative review counts for much. From a field of 34 criteria, "Review of your work by peers outside your department or institution" and "Letters of recommendation from people in your field" were tenth and twelfth, respectively — with 20–30% of the respondents stating that their institutions placed no emphasis on these factors at all.

Yet in *Nature*'s interviews, most administrators insisted that metrics don't matter

METRICS PERCEPTIONS

Q: At your institution or department are metrics of scientific performance used to any degree in any of the following?



nearly as much for hiring, promotion and tenure as the poll respondents seem to think. Some administrators said that they ignore citation-based metrics altogether when making such decisions, and instead rely largely on letters of recommendation solicited from outside experts in a candidate's field. "Outside letters basically trump everything," says Robert Simoni, chairman of the biology department at Stanford University in California.

That sentiment was echoed by academic administrators worldwide. "Metrics are not used a great deal," says Alex Halliday, head of the Mathematical, Physical and Life Sciences Division at the University of Oxford, UK. "The most important things are the letters, the interview and the CV, and our opinions of the papers published," he says.

"I don't look at impact factors" of the journals a candidate publishes in, says Kenichi Yoshikawa, dean of the Graduate School of Science at Japan's Kyoto University. "These usually highlight trendy papers, boom fields and recently highlighted topics. We at Kyoto don't want to follow boom."

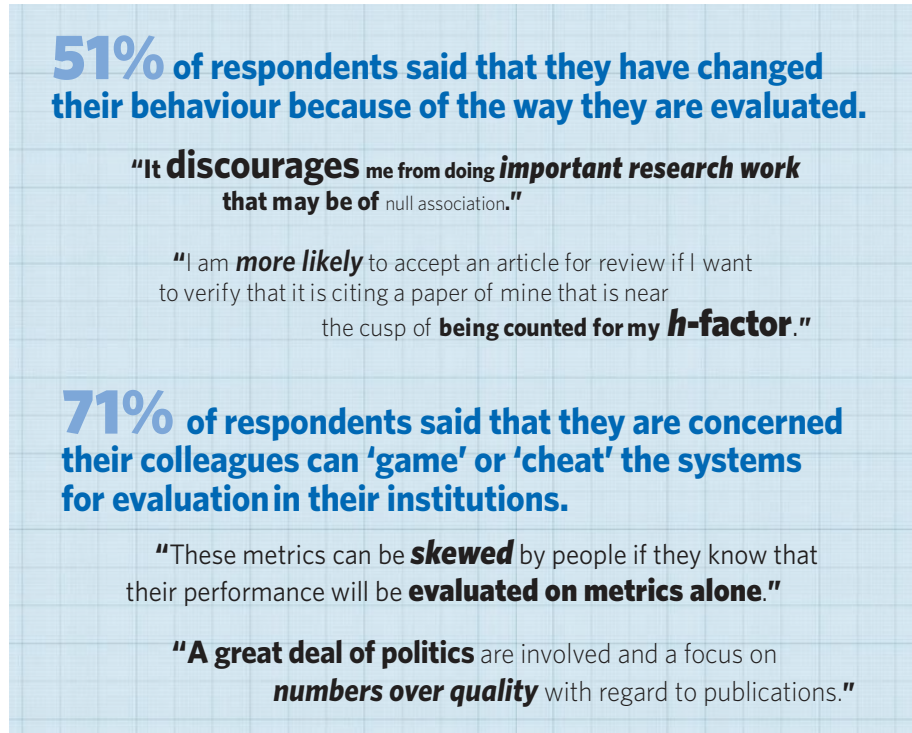
Metrics are not wholly excluded, of course. Those 'qualitative' letters of recommendation sometimes bring in quantitative metrics by the back door. "We do not look at publication records or tell the reviewers to," says Yigong Shi, dean of the School of Life Sciences at Tsinghua University in Beijing. "But in reality, they do have an impact, because the reviewers will look at them."

Mixed messages

Administrators may also send mixed signals: metrics don't matter, except that they do. "Each year we collect the average performances of people across various different things: student evaluations of lectures, teaching loads, research income, paper output, *h*-indices," says Tom Welton, head of the chemistry department at Imperial College London. Welton insists that this information is reported back to researchers as a guideline, "not a hurdle that has to be leapt over to get a promotion". Nevertheless, the fact that such measures are being made could give the impression that they are being relied on heavily.

At the Massachusetts Institute of Technology in Cambridge, Claude Canizares, vice-president for research and associate provost, says that "we pay very little attention, almost zero, to citation indices and counting numbers of publications". But, he says, "if someone has multiple publications in a higher-impact journal, it's like getting another set of letters — the peers that reviewed that paper gave it high marks".

A separate reason for the disparity is that



the use of metrics can vary markedly between countries (see 'Around the world with metrics', page 862) — or even between disciplines.

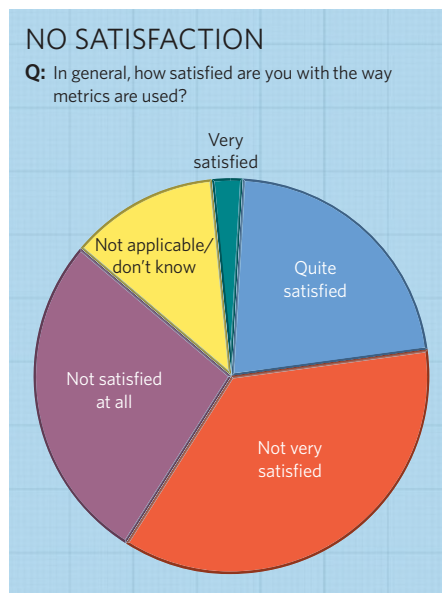
Poll respondents and administrators agree that metrics have potential pitfalls. For example, 71% of respondents said that they were concerned that individuals at their institutions could manipulate the metrics, for example by publishing several papers on the same basic work. Most deans and provosts seemed less concerned about that possibility, arguing that such practices were unlikely to slip past

reviewers. But they were wary of the more insidious effects of using metrics.

"If you decide that publishing a large number of papers is important, then you've decided that's what quality is," says Gregory Taylor, dean of the Science Faculty at the University of Alberta in Edmonton, Canada. "That's always a very dangerous route to go down, because then you get people working to achieve by the formulae, which isn't a very good way to encourage people to use their imagination." Indeed, half the poll respondents said that they shaped their research behaviours on the basis of the metrics being used at their university. Although many of the altered behaviours given were fairly innocuous — for example, "work harder" — some had the potential to compromise scientific ideals. "It discourages me from doing important research work that may be of null association," said one respondent, a US postdoctoral fellow.

Breaking the old-boys' networks

Despite general dissatisfaction with the way in which metrics are being used, some poll respondents welcome them. Many said that they appreciated the transparency and objectivity that quantitative metrics could provide. "I prefer this to qualitative metrics," wrote one, a department head in chemistry and engineering from Europe. Others who were dissatisfied with the use of metrics at their institution said they felt that the metrics weren't being used enough or weren't being used consistently.



Around the world with metrics

North America

Academic institutions in the United States comprise a complex system of state-supported and private universities, all largely independent of one another. Although they use a similar system for evaluating researchers for hiring, promotion and tenure (in most cases, letters from experts in the field hold the most sway) *Nature's* interviews with nearly a dozen deans, provosts and department heads showed considerable variation.

For example, many universities pay close attention to a candidate's ability to raise grant money. And for some the source counts: at Arizona State University in Tempe, provost Betty Capaldi says that funding from the US National Science Foundation is regarded as more prestigious than funding from the US National Institutes of Health, because the former is considered more competitive. By contrast, Marc Kastner, dean of the School of Science at Massachusetts Institute of Technology in Cambridge, says that the amount of funding almost never comes into consideration for a researcher's promotion. "If someone can have a strong impact on an international community without any funding, more power to them," he says.

Many university officials in Canada say that they evaluate scientists on the basis of 40% research, 40% teaching and 20% 'service' to the institution and the scientific community. But most that *Nature* spoke to admit to a loose interpretation of this breakdown — particularly when it comes to teaching. Although weak teaching can be overlooked for someone with an influential research programme, several North American administrators said that someone who is an excellent teacher but a poor researcher will stand little chance of earning tenure.

Europe

In the United Kingdom, colleges and universities are beholden to the country's Research Assessment Exercise (RAE), in which government funding is based on the quality of research at a department level. Thus institutions are systematic about collecting metrics, including statistics on papers published and student-evaluation measures. "It's not surprising that procedures that universities use internally tend to map on to and emulate the judgements that we know are being made by the RAE," says Rod Coombs, vice-president for innovation and economic development at the University of Manchester, UK. Quantitative metrics may become even more important under a proposed new system, the Research Excellence Framework, but details are still being debated.

Italian universities have a very different system: neither a department nor a university can choose its own staff. An ad hoc committee is set up for every post in the country, with just one representative from the university involved. For entry-level positions, selection committees must be guided by internationally recognized metrics, such as total number of citations, average number of citations per publication and the impact factor of journals published in. The law does not prescribe how much weight these metrics should have in decision-making. Some Italian research institutions have more independence. Pier Paolo di Fiore, former director of the FIRCO Institute of Molecular Oncology (IFOM), for example, says that criteria such as impact factor serve only as orientation.

In Germany, the 80 basic research institutes run by the Max Planck Society have a well-defined evaluation system based partly on metrics and partly on letters of evaluation. Elsewhere in Germany, there is little consensus on how to measure academic performance objectively.

Asia

The norm throughout much of China, South Korea and Japan is for short-term academic appointments, commonly five years. But because renewals are mostly automatic, employment is effectively for life. The flip side is that there is no set schedule for promotion; a scientist has the chance to move to the next stage in his or her career only when someone else leaves or dies.

Metrics are used heavily, although some Chinese universities have begun to align their evaluation methods with North America, and to de-emphasize metrics. Yi Rao, dean of the School of Life Sciences at Peking University, and Yigong Shi, dean of Tsinghua University's School of Life Sciences, both say that their institutions now base much of their assessment on recommendations from international experts in a candidate's field. International experience is also looked on very favourably, says Makoto Oka, dean of the School of Science at the Tokyo Institute of Technology.

Other Asian institutions have moved towards more metrics. Ben Chao, dean of the Earth sciences college at Taiwan's National Central University in Jongli City, says that the system has become more rigid and quantitative in the past five years. Candidates for hiring or promotion are judged 50% on research, 30% on teaching and 20% on community service. Each category has scores, says Chao, and minimum scores are required for advancement from assistant to associate professor. In the evaluation, the impact factor of journals is considered, as are metrics about the individual to some degree. "It's OK," he says, "but we always find cases where numbers don't reflect how hard they work and how much they contribute."

"The metrics can be nullified at the college or provost level," complained a US professor of neuroscience. If nothing else, says Welton, the use of quantitative measures can reassure young researchers that the institution is not perpetuating an old-boys' network, in which personal connections are valued over actual achievement. Administrators who say that they do consider metrics in the decision-making process stress that they recognize the limitations of such measures in defining the career of an individual. Researchers in different fields and different specialities publish and cite at different rates. An intimate understanding of the fields — and more importantly the individuals being assessed — is crucial, they say. This ultimately makes the use of metrics more subjective by necessity.

Surprisingly, if poll respondents desire change, it's not necessarily away from quantitative metrics. When *Nature* gave respondents a list and asked them to choose the five criteria that they thought should be used to evaluate researchers, the most frequently chosen was "Publication in high-impact journals", followed by "Grants earned", "Training and mentoring students" and "Number of citations on published research". In other words, what respondents think they are being measured on roughly matches what they want to be measured on.

The challenge for administrators, it seems, is not to reduce their reliance on metrics, but to apply them with more clarity, consistency and transparency. "The citation index is one of those things that is interesting to look at, but if you use it to make hiring decisions or use it as a

sole or main criterion, you're simply abrogating a responsibility to some arbitrary assessment," says Jack Dixon, vice-president and chief scientific officer of the Howard Hughes Medical Institute in Chevy Chase, Maryland. While he says that the institute eschews such metrics, he recognizes that they will continue to be used. "All decisions are based on various criteria. The thing you hope for is that the decisions are fair, and are based upon criteria that the reviewers know and understand."

Alison Abbott, David Cyranoski, Nicola Jones, Brendan Maher, Quirin Schiermeier and Richard Van Noorden all contributed to this article.

See Editorial, page 845, and metrics special at www.nature.com/metrics. Full results of the survey are available at go.nature.com/em7auj.