NATURE | COLUMN: MUSE

Phil Ball

Richard Haughton

Philip Ball.

The mathematics of science's broken reward system

Theoretical models of how science works provide valuable insights, says Philip Ball.

16 November 2016

Science has been peculiarly resistant to self-examination. During the ‘science wars' of the 1990s, for instance, scientists disdained sociological studies of their culture. Yet there is now a growing trend for scientists to use the quantitative methods of data analysis and theoretical modelling to try to work out how, and how well, science works — often with depressing conclusions. Why are these kinds of studies being produced, and what is their value?

Take a study published on 10 November[1] by psychologists Andrew Higginson of the University of Exeter and Marcus Munafò of the University of Bristol, UK. It considers how scientists can maximize their ‘fitness’, or career success, in a simplified ecosystem that allows them to invest varying amounts of time and effort into exploratory studies. The study finds that in an ecosystem that rewards a constant stream of high-profile claims, researchers will rationally opt for corner-cutting strategies, such as small sample sizes. These save on the effort required for each study, but they raise the danger that new findings will not prove robust or repeatable.

A slightly different perspective — but a similar conclusion — comes from work published on 21 September[2], by information scientist Paul Smaldino at the University of California, Merced, and evolutionary ecologist Richard McElreath, at the Max Planck Institute for Evolutionary Anthropology in Leipzig, Germany. They take an evolutionary view, imagining that laboratories are in competition for rewards, and that the most successful of them produce more 'progeny': new research groups that use the same techniques and strategies. There is generally a trade-off between productivity and rigour: producing more statistically secure, replicated findings takes more time and effort, but generating too many false positives will eventually take its toll on reputations. Under selection for productivity, however, less-rigorous methods spread and false discovery rates increase.

Metric incentives

It is no coincidence that this kind of game-theory approach to modelling science has arisen as researchers worry about the rise of metric-driven management of science’s funding and reward systems. Bibliometric data sets mean that scientific success can be quantified with achievement metrics such as the h-index — a quantitative measure of the number of highly cited papers that an academic produces — or the journal impact factor, which proposes to identify the most visible journals.

Many researchers have warned that the availability of these metrics, which are supposed to make the management of science and funding more systematic and objective, may be changing the nature of science. They are starting to dominate how science is structured and steered and place great pressure on researchers, especially in the earlier stages of their career, to publish often and prominently. As a result, understanding science as a social phenomenon has become a matter of some urgency.

And by quantifying goals and rewards, metrics transform science into the kind of game-theoretical play of utility maximization and cost-benefit analyses with which economists and ecologists have long become familiar. It is only, I think, because science is now so driven by metrics-based incentives that such models and analyses can meaningfully be applied at all.

The models suggest that when researchers apply strategies that boost individual and institutional performance metrics, by publishing papers as often as they can in high-profile journals, scientific research as a whole becomes less efficient and reliable. (Rigorous journals and referees can partly ameliorate the situation).

“Whenever quantitative metrics are used as proxies to evaluate and reward scientists,” write Smaldino and McElreath, “those metrics become open to exploitation if it is easier to do so than to directly improve the quality of research.” That’s basically a statement of Goodhart’s law, familiar to economists: when a measure of success becomes a target, it loses its worth.

Munafò thinks that placing so much emphasis on high-profile publications is a mistake, and that it would relieve the pressure on junior scientists if they were given more opportunity to develop their ideas and to report ‘null’ findings. That fits with a conclusion from a study published last year[3] by bioinformaticist Andrey Rzhetsky of the University of Chicago, Illinois, and his colleagues that modelled the rate of scientific discoveries. Using a data set from biomedical chemistry, they suggest that over the past three decades, scientists have been gradually choosing more conservative, low-risk research strategies. They argue that papers reporting replications and null results are important for increasing the efficiency of science, but because they don't tend to get highly cited, there is little incentive to produce them.

Why models are useful

A question hanging over these studies is whether they really tell us anything we could not already intuit. Don't we already know that science's reward schemes encourage 'safe' research and encourage corner-cutting?

Perhaps. But relying on what 'everyone knows', without inspecting the evidence, has never been a good way to do science. Grumbling about over-assessment and bad incentives is one thing, but if they can be rigorously shown to create problems, then they are harder to ignore.

While the current wave of science models may not yet have told us anything utterly unexpected, studies of game-theoretic competition and complexity in other systems, both social and natural, often turn out to have some counter-intuitive implications; apparently well-motivated incentives can lead to unintended, detrimental outcomes.

The real benefit of mathematical modelling is that it gives you some idea of how the problems arise, and therefore which dials should be turned to ease them. How, for example, do you foster more risk-taking research strategies? By earmarking funding and creating Grand Challenges? By developing teams and institutions that spread the risk? It's not as obvious as it might seem.

Scientists have nothing to fear, and much to gain, from having the lens of science turned back on its own conduct. It's the old injunction: physician, heal thyself.

Nature | doi:10.1038/nature.2016.20987

References

1. Higginson, A. D. & Munafò, M. R. *PLoS Biol.* **14**, e2000995 (2016).
2. Smaldino, P. E. & McElreath, R. *R Soc. Open Sci.* **3**, 160384 (2016).
3. Rzhetsky, A., Foster, J. G., Foster, I. T. & Evans, J. A. *Proc. Natl Acad. Sci. USA* **112**, 14569 (2015).