

# COMMENT

**SPACE** Astrophysics' long relationship with the military **p.173**

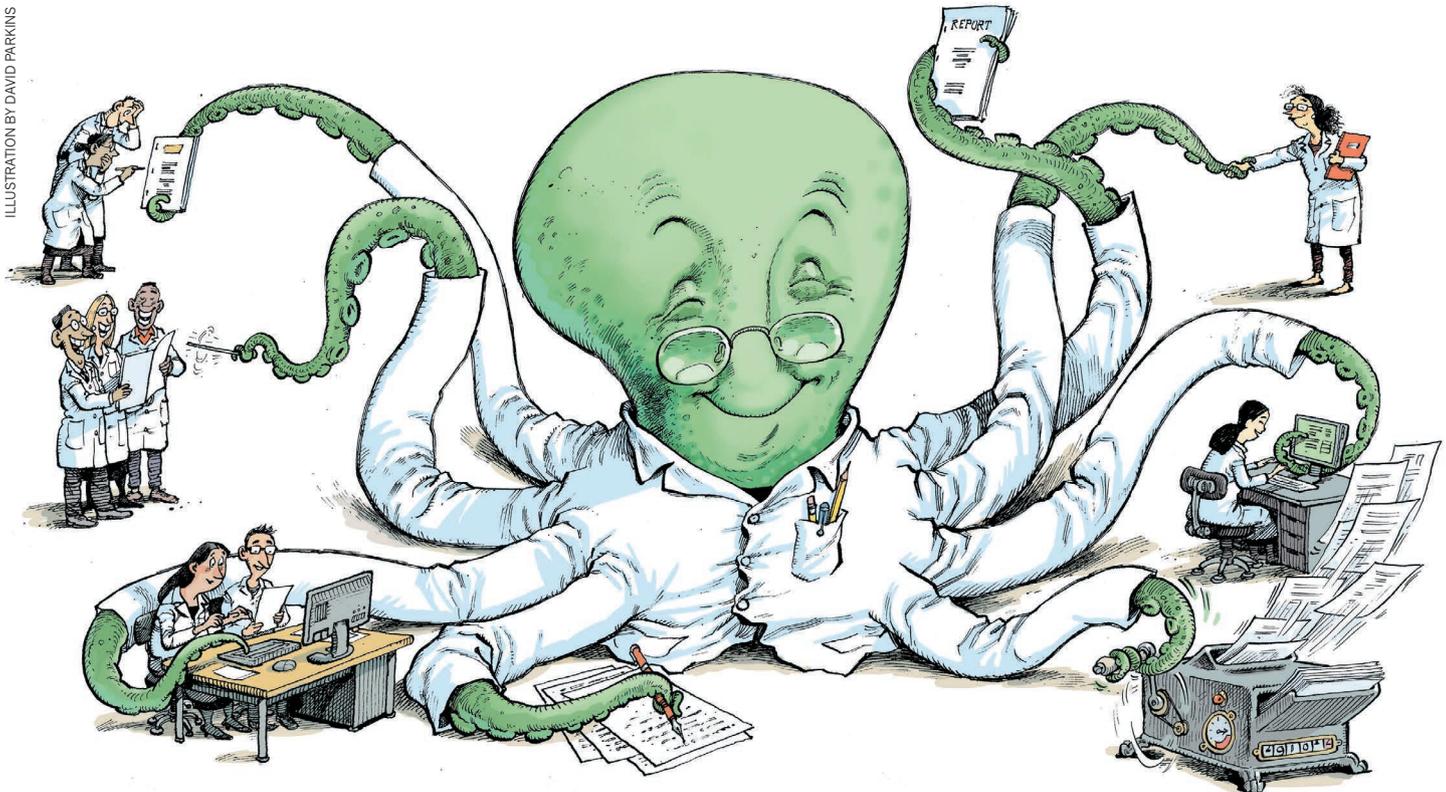


**HISTORY** At last, a definitive biography of Helmholtz **p.175**

**PSYCHOLOGY** How a questionable personality quiz went global **p.176**

**ECOLOGY** Pathogen could wipe out New Zealand's oldest tree species **p.177**

ILLUSTRATION BY DAVID PARKINS



## The scientists who publish a paper every five days

To highlight uncertain norms in authorship, **John P. A. Ioannidis**, **Richard Klavans** and **Kevin W. Boyack** identified the most prolific scientists of recent years.

**A**uthorship is the coin of scholarship — and some researchers are minting a lot. We searched Scopus for authors who had published more than 72 papers (the equivalent of one paper every 5 days) in any one calendar year between 2000 and 2016, a figure that many would consider implausibly prolific<sup>1</sup>. We found more than 9,000 individuals, and made every effort to count only 'full papers' — articles, conference papers, substantive comments and reviews — not editorials, letters to the editor and the like. We hoped

that this could be a useful exercise in understanding what scientific authorship means.

We must be clear: we have no evidence that these authors are doing anything inappropriate. Some scientists who are members of large consortia could meet the criteria for authorship on a very high volume of papers. Our findings suggest that some fields or research teams have operationalized their own definitions of what authorship means.

The vast majority of hyperprolific authors (7,888 author records, 86%) published in physics. In high-energy and particle

physics, projects are done by large international teams that can have upwards of 1,000 members. All participants are listed as authors as a mark of membership of the team, not for writing or revising the papers. We therefore excluded authors in physics.

Of what remained, 909 author records were Chinese or Korean names. Because Scopus disambiguates Chinese and Korean names imperfectly, these may have wrongly combined distinct individuals. For 2016 (when disambiguation had improved for Chinese and Korean names), at least ▶

► 12, and possibly more than 20, authors based in China were hyperprolific, the largest number from any country that year. We believe that this could be connected to Chinese policies that reward publication with cash or to possible corruption<sup>2,3</sup>.

Because of the disambiguation issues, we excluded these names from further analysis, as well as group names and cases in which we found errors (such as journalistic news items misclassified as full articles), duplicate entries, or conference papers misassigned to an organizer.

This left 265 authors (see Supplementary Information). The number of hyperprolific authors (after our exclusions) grew about 20-fold between 2001 and 2014, and then levelled off (see ‘Hyperprolific authors proliferate’). Over the same period, the total number of authors increased by 2.5-fold.

We e-mailed all 265 authors asking for their insights about how they reached this extremely productive class. The 81 replies are provided in the Supplementary Information. Common themes were: hard work; love of research; mentorship of very many young researchers; leadership of a research team, or even of many teams; extensive collaboration; working on multiple research areas or in core services; availability of suitable extensive resources and data; culmination of a large project; personal values such as generosity and sharing; experiences growing up; and sleeping only a few hours per day.

About half of the hyperprolific authors were in medical and life sciences (medicine  $n = 101$ , health sciences  $n = 11$ , brain  $n = 17$ , biology  $n = 6$ , infectious diseases  $n = 3$ ). When we excluded conference papers, almost two-thirds belonged to medical and life sciences (86/131). Among the 265, 154 authors produced more than the equivalent of one paper every 5 days for 2 or more calendar years; 69 did so for 4 or more calendar years. Papers with 10–100 authors are common in these CVs, especially in medical and life sciences, but papers with the hundreds of authors seen in particle physics are uncommon.

Materials scientist Akihisa Inoue, former president of Tohoku University in Japan and a member of multiple prestigious academies, holds the record. He met our definition of being hyperprolific for 12 calendar years between 2000 and 2016. Since 1976, his name appears on 2,566 full papers indexed in Scopus. He has also retracted seven papers found to be self-duplications of previously published work<sup>4</sup>. We searched for news articles in Google detailing retractions for the next 20 most hyperprolific authors and found only one other author (Jeroen Bax) to have one retracted paper.

The 265 hyperprolific authors worked in 37 countries, with the highest number in the United States ( $n = 50$ ), followed by Germany ( $n = 28$ ) and Japan ( $n = 27$ ). The proportion from the United States (19%) is roughly

similar to its share of published science. Germany and Japan are over-represented. There were disproportionately more hyperprolific authors in Malaysia ( $n = 13$ ) and Saudi Arabia ( $n = 7$ ), countries both known to incentivize publication with cash rewards<sup>5</sup>.

Hyperprolific authors also tended to cluster in particular institutions, often as part of a common study. For example, Erasmus University Rotterdam in the Netherlands had nine hyperprolific authors, more than any other institution. Seven of them co-authored mostly papers related to the Rotterdam study, a nearly 30-year-old epidemiological project, or its successor Generation R study, which have followed multiple health parameters in thousands of older adults and yielded thousands of publications. Five hyperprolific investigators from Harvard University in Cambridge, Massachusetts, also often co-authored papers related to cohort studies. Eleven hyperprolific authors across different institutions were on one large cohort study,

**“Whether and how authorship is justified unavoidably varies for each author and each paper, and norms differ by field.”**

## SURVEY

### Criteria fulfilled?

**One-third of the 81 authors identified as hyperprolific in 2016 replied when asked how often they met each of 4 criteria established for authorship of medical studies. Of the 27 responders, 19 admitted they had not met at least 1 criterion more than 25% of the time. Eleven wrote that they had not met two or more criteria upwards of 25% of the time.**

- Substantial contributions to the conception or design of the work; or the acquisition, analysis or interpretation of the data for the work (9 of 27 met this criterion in less than 75% of their papers).
- Drafting the work or revising it critically for important intellectual content (9 of 27 met this criterion in less than 75% of their papers).
- Final approval of the version to be published (3 out of 27 met this criterion in less than 75% of their papers).
- Agreement to be accountable for all aspects of the work (14 out of 27 met this criterion in less than 75% of their papers).

the European Prospective Investigation on Cancer and Nutrition; other large epidemiological studies were also represented. Hyperprolific authors were also concentrated in cardiology and crystallography.

These biological and medical disciplines with many hyperprolific authors exhibit different patterns from those found in particle and high-energy physics. Papers with hundreds to thousands of authors are the norm across a community of many thousands of scientists working in projects based at CERN, Europe’s particle-physics laboratory near Geneva, Switzerland. In crystallography, papers tend to have few co-authors. In epidemiology and cardiology, long lists of authors appear only in relationship to specific research teams that seem to have a tradition of extensive authorship lists.

This raises the question of what authorship entails. The US National Institutes of Health, for example, has guidelines on the activities that qualify: actively supervising, designing and doing experiments, and data acquisition and analysis outside “very basic” work plus drafting the manuscript. Collecting funds or distant mentorship do not qualify. Most of the 6,000 authors in a recent survey across many geographical regions and disciplines felt that drafting a paper, interpreting results and analysing data should qualify for authorship, but attitudes varied by region and field<sup>6</sup>.

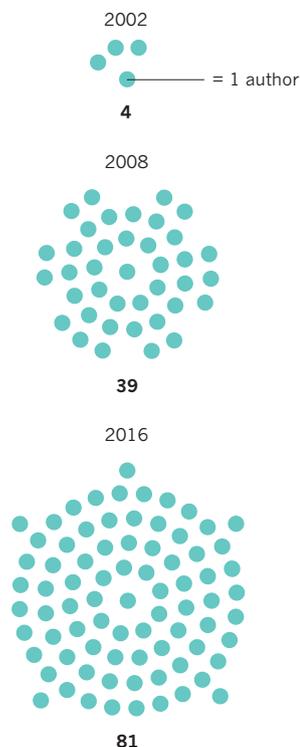
## AUTHORSHIP CRITERIA

Perhaps the most widely established requirements for authorship are the Vancouver criteria established by the International Committee of Medical Journal Editors in 1988. These specify that authors must do all of four things to qualify: play a part in designing or conducting experiments or processing results; help to write or revise the manuscript; approve the published version; and take responsibility for the article’s contents.

The International Committee of Medical Journal Editors does not count supervision, mentoring or obtaining funding as sufficient for authorship. We did observe that some authors seemed to become hyperprolific on becoming full professors, department chairs or both. It is common and perhaps expected for scientists who assume leadership roles in large centres to accelerate their productivity. For example, clinical cardiologists publish more papers after they assume director roles (despite heavy clinical and administrative duties). Occasionally, the acceleration is stunning: at the peak of their productivity, some cardiologists publish 10 to 80 times more papers in one year compared with their average annual productivity when they were 35–42 years old. There was also often a sharp decrease after passing the chair to a successor. Another study noted similar patterns two decades ago<sup>7</sup>.

## HYPERPROLIFIC AUTHORS PROLIFERATE

Numbers of authors with more than 72 papers a year increased dramatically over time.



20 authors, another might give credit only to 3 people or none. For example, genome-wide studies typically include many dozens of authors. As a dramatic counter-example, one recent publication of a genome-wide study had only one author<sup>9</sup>, and apparently that researcher did the same amount of work for which perhaps dozens would get authorship credit in similar papers spearheaded by different teams. Some evidence suggests that the increase in the average number of authors per paper does not reflect so much the genuine needs of team science as the pressure to ‘publish or perish’<sup>10</sup>.

Widely used citation and impact metrics should be adjusted accordingly. For instance, if adding more authors diminished the credit each author received, unwarranted multi-authorship might go down. We found that the 30 hyperprolific authors who seemed to benefit the most from co-authorship numbered 6 cardiologists and 24 epidemiologists (including those working on population genetics studies). (For these scientists, the ratio of their Hirsch H index to their co-authorship-adjusted Schreiber  $H_m$  index was higher; see Supplementary Information.)

Overall, hyperprolific authors might include some of the most energetic and excellent scientists. However, such modes of publishing might also reflect idiosyncratic field norms, to say the least. Loose definitions of authorship, and an unfortunate tendency to reduce assessments to counting papers, muddy how credit is assigned. One still needs to see the total publishing output of each scientist, benchmarked against norms for their field. And of course, there is no substitute for reading the papers and trying to understand what the authors have done. ■

**John P. A. Ioannidis** is a professor of medicine at the Meta-Research Innovation Center at Stanford (METRICS), Stanford University, California, USA. **Richard Klavans** and **Kevin W. Boyack** are researchers at SciTech Strategies in Philadelphia, Pennsylvania, and Albuquerque, New Mexico, USA. e-mail: [jioannid@stanford.edu](mailto:jioannid@stanford.edu)

1. Wager, E., Singhvi, S. & Kleinert, S. *PeerJ* **3**, e1154 (2015).
2. Quan, W., Chen, B. & Shu, F. Preprint at <https://arxiv.org/abs/1707.01162> (2017).
3. Hvistendahl, M. *Science* **342**, 1035–1039 (2013).
4. *Nature* **483**, 246 (2012).
5. Abritis, A., McCook, A. & Retraction Watch. *Science* **357**, 541 (2017).
6. Patience, G. S., Galli, F., Patience, P. A. & Boffito, D. C. Preprint at bioRxiv <https://doi.org/10.1101/323519> (2018).
7. Drenth, J. J. *Am. Med. Assoc.* **280**, 219–221 (1998).
8. Sauermaun, H. & Haeussler, C. *Sci. Adv.* **3**, e1700404 (2017).
9. Kim, S. K. *PLoS ONE* **13**, e0200785 (2018).
10. Papatheodorou, S. I., Trikalinos, T. A. & Ioannidis, J. P. *J. Clin. Epidemiol.* **61**, 546–551 (2008).

Supplementary Information accompanies this Comment online (see [go.nature.com/2phadg2](https://go.nature.com/2phadg2)).

One unexpected result was that some hyperprolific authors placed many publications in a single journal. Prominent in this regard were *Acta Crystallographica Section E: Structure Reports Online* (relaunched in 2014 as Section E: Crystallographic Communications, with brief structural data reports now published in IuCrData) and *Zeitschrift für Kristallographie New Crystal Structures*. Three authors have each published more than 600 articles in the former (Hoong-Kun Fun, Seik Weng Ng and Edward Tiekink); three authors have each published more than 400 papers in the latter (Karl Peters, Eva Maria Peters and Edward Tiekink). Three other authors (Anne Marie Api, Charlene Letizia, Sneha Bhatia) published many papers in single supplement issues of *Food and Chemistry Toxicology* focused on reviews of fragrance materials.

Journals indexed in Scopus are generally considered to be quality journals. The citation impact of hyperprolific authors was usually high, but there was large variability: with a median of 19,805 citations per author (range: 380 to 200,439). The median number of full papers per hyperprolific author in 2000–2016 was 677; across all hyperprolific authors, last author positions accounted for 42.5%, first author positions for 7.1%, and single authorships for 1.4%. Across the years, the median proportion of papers with middle author positions (that is, not a single, first or last author) was 51%, but varied from 2.1% to 98.5% for individual authors.

Our work to identify hyperprolific authors is admittedly crude. It is mainly intended to raise the larger question of what authorship entails. Whether and how authorship is justified unavoidably varies for each author and each paper, and norms differ by field. It is likely that sometimes authorship can be gamed, secured through coercion or provided as a favour. We could not assess these patterns in our data. We did not examine contributorship statements<sup>8</sup>, which are not archived in Scopus. Nevertheless, even contributorship statements can be gamed and might not be accurate.

Further work is needed to explore how to best normalize these data and what is the optimal level of normalization: for example, adjusting for wide discipline, relatively narrow field and/or highly specific research team.

### WHAT AUTHORS SAY

To better understand authorship norms, we e-mailed a survey to the 81 hyperprolific authors of 2016 (see Supplementary Information). We asked whether they fulfilled all four Vancouver criteria. Of the 27 who completed the survey, most said they did not (see ‘Survey’). Almost all the responders were from US and European institutions. The only two responders from elsewhere stated that they failed Vancouver criteria in most of their papers. It is likely that the survey underestimates the proportion not

meeting Vancouver criteria.

Not all authors had approved the final versions of their own papers, but all considered approval of the final version necessary for authorship. Fifty-nine per cent (16 of 27) said that they had contributed more than any other listed author for 25 or more of the papers they authored in 2016.

Responses to the question “What, in your own words, do you think should be required for authorship?” generally reflected a requirement for “significant contributions”, but also dissatisfaction with how authorship was assessed. One scientist said, “I personally don’t count them as ‘my papers’ and don’t have them on my CV as such, as there is a distinction between being a ‘named author’ versus a ‘consortium member’ authorship.” Another observed that authorship was often awarded for seniority, and another that better distinctions were essential. “I think there should be levels of authorship — and not those implied by order!” It will be interesting to monitor how innovations in assigning credit, such as data citation or formal author contribution taxonomies, could alter authorship conventions. Authorship norms can vary within each field and even within each team. For example, some teams in epidemiology and cardiology apparently offer authorship more generously; others stick to stricter (and probably more appropriate) authorship criteria. For a similar task and contribution, one cohort study might credit