

THE GIANT PLAN TO TRACK DIVERSITY IN RESEARCH JOURNALS

Efforts to chart and reduce bias in scholarly publishing will ask authors, reviewers and editors to disclose their race or ethnicity.

By Holly Else and Jeffrey M. Perkel

In the next year, researchers should expect to face a sensitive set of questions whenever they send their papers to journals, and when they review or edit manuscripts. More than 50 publishers representing over 15,000 journals globally are preparing to ask scientists about their race or ethnicity – as well as their gender – in an initiative that’s part of a growing effort to analyse researcher diversity around the world. Publishers say that this information, gathered and stored securely, will help to analyse who is represented in journals, and to identify whether there are biases in editing or review that sway which findings get published. Pilot testing suggests that many scientists support the idea, although not all.

The effort comes amid a push for a wider acknowledgement of racism and structural racism in science and publishing – and the need to gather more information about it. In any one country, such as the United States, ample data show that minority groups are under-represented in science, particularly at senior levels. But data on how such imbalances are reflected – or intensified – in research journals are scarce. Publishers haven’t systematically looked, in part because journals are international and there has been no measurement framework for race and ethnicity that made sense to researchers of many cultures.

“If you don’t have the data, it is very difficult to understand where you are at, to make changes, set goals and measure progress,” says Holly Falk-Krzesinski, vice-president of research intelligence at the Dutch publisher

Elsevier, who is working with the joint group and is based in Chicago, Illinois.

In the absence of data, some scientists have started measuring for themselves. Computational researchers are scouring the literature using software that tries to estimate racial and ethnic diversity across millions of published research articles, and to examine biases in who is represented or cited. Separately, over the past two years, some researchers have criticized publishers for not having diversity data already, and especially for being slow to collate information about small groups of elite decision makers: journal editors and editorial boards. At least one scientist has started publicizing those numbers himself.

After more than 18 months of discussion, publishers are now close to agreeing on a standard set of questions – and some have already started gathering information. Researchers who have pushed to chart racial and ethnic diversity at journals say that the work is a welcome first step.

“It is never too late for progress,” says Joel Babdor, an immunologist at the University of California, San Francisco. In 2020, he co-founded the group Black in Immuno, which supports Black researchers in immunology and other sciences. It urges institutions to collect and publish demographic data, as part of action plans to dismantle systemic barriers affecting Black researchers. “Now we want to see these efforts being implemented, normalized and generalized throughout the publishing system. Without this information, it is impossible to evaluate the state of the current



system in terms of equity and diversity,” the group’s founders said in a statement.

Lacking data

The effort to chart researcher diversity came in the wake of protests over the killing of George Floyd, an unarmed Black man, by US police in May 2020. That sparked wider recognition for the Black Lives Matter movement and of the structural racism that is embedded in society, including scientific institutions. The following month, the Royal Society of Chemistry (RSC), a learned society and publisher in London, led 11 publishers in signing a joint commitment to track and reduce bias in scholarly publishing (see go.nature.com/36gqrtp). This would include an effort to collect and analyse anonymized diversity data, as reported by



ILLUSTRATION BY CAMELIA PHAM/FOLIO ART



**IT ALMOST SEEMED
AN INSURMOUNTABLE
CHALLENGE WHEN WE
WERE WORKING ON IT
ON OUR OWN.”**

categories listed are terms chosen by those who conducted a particular survey or study.)

The joint commitment group decided that it would ask scientists about their gender and race or ethnicity when they authored, reviewed or edited manuscripts. The group started by agreeing on a standard schema, or structured list, of questions about gender – although even this wasn't simple, requiring detailed explanatory notes. But what to ask researchers globally about race and ethnicity was a tougher problem, as publishers such as Elsevier had discussed before they joined the group. “It almost seemed an insurmountable challenge when we were working on it on our own,” says Falk-Krzesinski.

Cultural understanding of race and ethnicity differs by country: social categories in India or China, for instance, are different from those in the United States. The historical associations of asking people to disclose these personal descriptors pose another set of problems, and could, if not sensitively handled, intensify concerns about how these data will be used. In countries such as the United States, people might be accustomed to sharing the information with their employers; some companies are required to report this to the federal government by law. But in others, such as Germany, authorities do not collect race or ethnicity data. Here, there is extreme sensitivity around racial classification – rooted in revulsion at the way such information was used in the 1930s and 1940s to organize the Holocaust. Race and ethnicity data must also be carefully processed during collection and storage under Europe's data-protection laws.

Computational audits

In the absence of comprehensive data, many studies in the past decade have used computational algorithms to measure gender diversity. Processes that estimate gender from names are far from perfect (particularly for Asian names), but seem statistically valid across large data sets. Some of this work has suggested signs of bias in peer review. An analysis of 700,000 manuscripts that the RSC published between 2014 and 2018, for instance², pointed the organization to biases against women at each stage of its publishing

authors, peer reviewers and editorial decision makers at journals. That group has now grown to 52 publishers. (Springer Nature, which publishes this journal, has joined the group; *Nature's* news team is editorially independent of its publisher.)

But publishers had a problem: they were lacking data. Many had made a start collecting and analysing information on gender, but few had tried to chart the ethnic and racial make-up of their contributors. Some that had done so had relied on their links to scholarly societies to gather regionally limited data.

The American Geophysical Union (AGU) in Washington DC, for instance, which is both a scientific association and a publisher, held information about some US members who had disclosed their race or ethnicity. In

2019, researchers used these data to study manuscripts submitted to AGU journals¹. They cross-checked author information with the AGU member data set, and found that papers with racially or ethnically diverse author teams were accepted and cited at lower rates than were those that had homogenous teams. But the scientists were able to check the race or ethnicity of author teams for only 7% of the manuscripts in their sample.

The UK Royal Society in London, meanwhile, had used annual surveys to collect data for its journals. But by mid-2020, its most recent report (covering 2018) had responses from just 30% of editors and 9% of authors and reviewers, in the categories ‘White British’, ‘White other’ and ‘Black and minority ethnic’. (Here, and throughout this article, the

JOURNAL DIVERSITY DATA

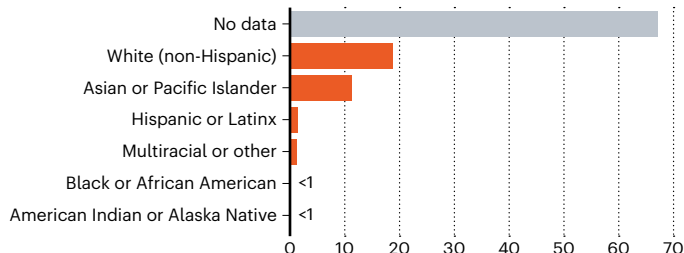
Publishers are still discussing how to collect global data on the race or ethnicity of authors, editors or reviewers at their journals. Some have already published information — but differ in the categories they report*. See go.nature.com/3jrfsfv for more data.

Early data on race and ethnicity

Here are how the American Association for the Advancement of Science (AAAS) and the American Chemical Society (ACS) reported data for authors and reviewers across their journals.

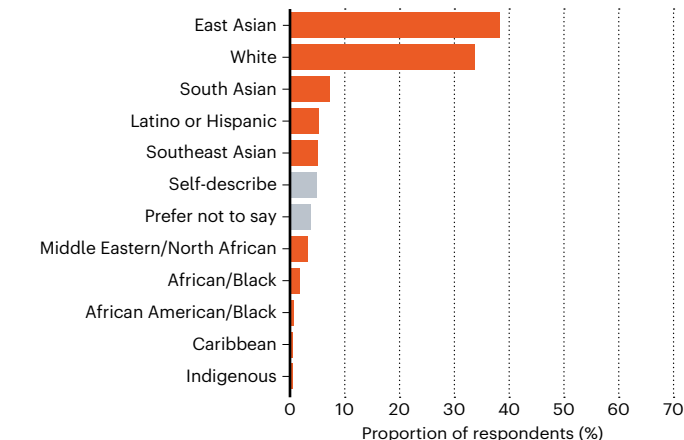
AAAS

Science family authors and reviewers.



ACS

Corresponding authors submitting across ACS journals (for reviewer data, see go.nature.com/3jrfsfv).

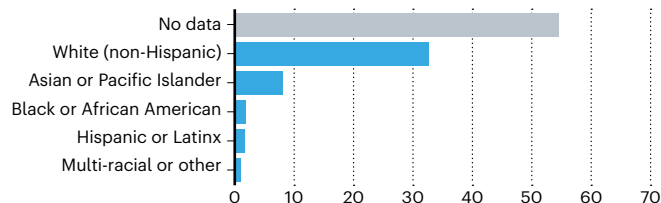


Editors at high-profile journals

Nature reporters asked publishers at nine scientific journals for data on the diversity of their editors or professional staff. Although some gave information on gender, none provided it for race or ethnicity at the journal level. But three (shown below) did so across a family of journals. See go.nature.com/3jrfsfv for detailed responses.

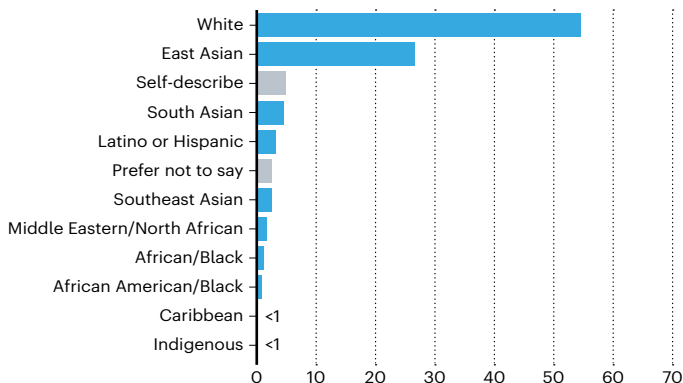
Science

Race or ethnicity data for editors and advisers across the Science family of journals.



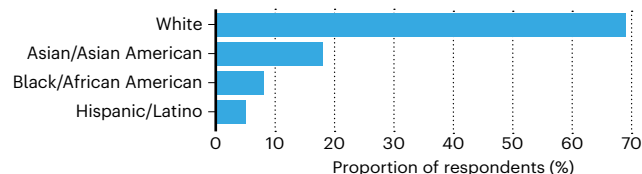
Journal of the American Chemical Society

Race or ethnicity data for editors at American Chemical Society journals*.



JAMA

Race or ethnicity data for editors or editorial board members for JAMA network journals.



*Categories listed are those used by the publishers that collected the data. *Respondents could select more than one category.

SOURCES: AAAS/ACS/JAMA

process; in response, it developed a guide for reducing gender bias. Collecting those data was crucial, says Nicola Nugent, publishing manager at the RSC in Cambridge, UK — without the baseline numbers, it was hard to see where to make changes.

Some researchers have also developed algorithms to estimate ethnicity or geographical origin from names. That idea goes back decades, but has become easier with massive online data sets of names and nationalities or ethnicities, together with growing computer power. Such algorithms can only ever provide rough estimates, but can be run across millions of papers.

US computational biologist Casey Greene at the University of Colorado Anschutz Medical Campus in Aurora argues that publishers could glean insights from these methods, if they apply them to large numbers of names and limit analysis to broad ethnicity classes — especially when examining past papers, for which it might not be possible to ask authors directly.

In 2017, for instance, a team led by computer

scientist Steven Skiena at Stony Brook University in New York used millions of e-mail contact lists and data on social-media activity to train a classifier called NamePrism. It uses people's first and last names to estimate their membership of any of 39 nationality groups — for example, Chinese, Nordic or Portuguese — or six ethnicities, corresponding to categories used by the US Census Bureau³. NamePrism clusters names into similar-seeming groups, and uses curated lists of names with known nationalities to assign nationalities to those groups. It is more accurate for some categories than for others, but has been cited in a few dozen other studies.

Some studies use these kinds of tools to analyse representation. In 2019, Ariel Hippen, a graduate student in Greene's lab, scraped biographical pages from Wikipedia to train a classifier that assigns names to ten geographical regions. A team including Greene, Hippen and data scientist Trang Le at the University of Pennsylvania, Philadelphia, then used the tool to document under-representation of

people from East Asia in honours and invited talks awarded by the International Society for Computational Biology⁴. Last year, Natalie Davidson, a postdoc in the Greene lab, used the same tool to quantify representation in *Nature's* news coverage, finding fewer East Asian names among quoted sources, compared with their representation in papers⁵.

Other studies analyse citation patterns. For instance, one analysis⁶ of US-based authors found that papers with authors of different ethnicities gained 5–10% more citations, on average, than did papers with authors of the same ethnicity, a finding that has been interpreted as a benefit of diverse research groups. And a 2020 preprint⁷ from a team led by physicist Danielle Bassett at the University of Pennsylvania found that authors of colour in five neuroscience journals are undercited relative to their representation; the team's analysis suggests that this is because white authors preferentially cite other white authors.

Instead of training a classifier, a different idea is to estimate ethnicity directly from

census information – although this approach is limited to names from the country that did the census. In January, a team used⁸ US Census Bureau data to assign US names a probability distribution of being associated with any of four categories: Asian, Black, Latinx or White. The researchers then studied papers by 1.6 million US-based authors, and found that work from what they describe as minoritized groups is over-represented in topics that tend to receive fewer citations, and that their research is less cited within topics.

Still, Cassidy Sugimoto, an information scientist at the Georgia Institute of Technology in Atlanta who worked on that study, says computational methods are largely incapable of addressing the most pressing questions about racial diversity and inclusion in science. This is because ethnicity is only loosely associated with family name (most obviously in the case of surname changes after marriage), and has many more dimensions than gender. “Race and ethnicity classification is infinitely more complicated than gender disambiguation,” she says.

Given those complex dimensions, the best option for collecting data is simply to invite scientists to self-identify, says Jory Lerback, a geochemist at the University of California, Los Angeles, who worked with the AGU on its studies of academic diversity.

Hippen, Davidson and Greene agree. In a correspondence article⁹ this year, they advise those using automated tools to be transparent, to share results with affected communities and to ask people how they identify, if possible.

Called out for inaction

As publishers discussed how to follow up their June 2020 commitment, they faced outside pressure. An increasing number of scientists began calling out the publishing industry for its inaction on providing diversity data.

In October 2020, *The New York Times* reported how several US scientists, including Babdor, were unhappy that publishers, despite their commitment, had no idea of how many Black researchers were among their authors.

That same month, Raymond Givens, a cardiologist at Columbia University Irving Medical Center in New York City, had begun privately tallying editors’ ethnicities himself. He counted the number of what he classed as Black, brown, white and Hispanic people on the editorial boards of two leading medical journals, *The New England Journal of Medicine* (*NEJM*) and *JAMA*, after reading a now-retracted article¹⁰ on affirmative-action programmes, published in a different society journal. Givens categorized the editors by looking at their photographs online, together with other contextual clues, such as surname and membership of associations that might indicate identity, and determined that just one of *NEJM*’s 51 editors was Black and one was Hispanic. At *JAMA*, he found that 2 of 49 editors were Black and 2

were Hispanic. Givens e-mailed the journals his data; he had no response from *JAMA* and got an acknowledgement from *NEJM*, but editors there didn’t get back to him.

Within months, *JAMA* had become embroiled in controversy after a deputy editor, Edward Livingston, hosted a podcast in which he questioned whether structural racism could exist in medicine if it was illegal. More than 10,000 people have now signed a petition calling for *JAMA* to take measures to review and restructure its editorial staff and processes, as well as to commit to a series of town-hall conversations with health-care staff and patients who



A LOT OF JOURNALS HAVE ALL OF A SUDDEN BEEN SHOCKED BY BEING CONFRONTED IN THIS WAY.”

are Black, Indigenous and people of colour (BIPOC). Livingston, and Howard Bauchner, the then-editor-in-chief of *JAMA*, have also stepped down from their posts.

Givens’ efforts became public in April 2021, when news website *STAT* reported his findings. “A lot of journals have all of a sudden been shocked by being confronted in this way,” says Givens. But it’s important to ask why it has taken them so long to start thinking about how to collect this kind of information, he says. He acknowledges that making his own categorizations is an “imperfect” method, but says someone had to undertake the project to confront journals with the problem.

Both *JAMA* and *NEJM* say they have added BIPOC editors to their boards, although *NEJM* did not provide a breakdown of editorial staff ethnicities when asked. *JAMA*, meanwhile, has published aggregate data only on editors and editorial board members across its 13 *JAMA* Network journals (see ‘Journal diversity data’).

Givens still has concerns that those who have joined editorial boards have peripheral influence compared with white men who retain central, powerful positions. He has continued his work, gathering gender and race data by eye on more than 7,000 editors at around 100 cardiology journals – finding that fewer than 2% are Black and almost 6% are Latinx – and looking at networks between the editors.

“When you look at the networks, white men are central: they are the hub from which all the spokes emanate,” he says. “Sometimes you really have to shake the system to force it to change. Until you are going to reshape the

system, we will still be having this conversation a decade from now.”

When it comes specifically to information on editorial board members, Givens says that’s not difficult to collect – if publishers truly put in the effort. He says it took him only a few months to do it. “It’s just counting,” he says. “When people say you have to start with collecting the data, I never have confidence that it will lead to anything. There needs to be intense pressure on them.”

Nature’s news team asked seven high-profile journals besides *JAMA* and *NEJM* (including *Nature*) for information about the diversity of editorial board members and professional staff. None provided it at the journal level, but some shared information about the make-up of staff across their entire company, or wider family of journals. These broader metrics might not reflect diversity at any one journal.

Ethnicity surveys

While the joint group of publishers started work on its race and ethnicity schema, some US publishers – who were not all in the group at the time – raced ahead with data collection.

As far back as 2018, the American Association for the Advancement of Science (AAAS) in Washington DC had begun working on how best to ask manuscript authors and reviewers about their race and ethnicity. It decided to use categories that closely followed US census descriptions, because that is a vetted system familiar to those in the United States, a spokesperson says.

In October 2020, the AAAS published data it had collected over the past year. The respondents covered only 12% of authors and reviewers in the Science family of journals. A report covering the subsequent year, released in January 2022, upped that coverage to 33%, because, the publisher said, it had improved the way it collected information using its electronic submission system for manuscripts and peer review. But data are still limited, and the AAAS is concerned that some researchers might not feel confident disclosing their ethnicity, its spokesperson says. The overall proportion identifying as African American or Black was less than 1%. Of the proportion who did report ethnicity, 57% identified as white (non-Hispanic) and 34% as Asian or Pacific Islander (which the AAAS grouped together in its reporting). The publisher is refining its race and ethnicity questions and last month added its name to the joint commitment. It is now looking at whether to adopt that group’s schema, when the framework is ready.

Another publisher that raced ahead was the American Chemical Society (ACS) in Washington DC, an early signatory of the joint commitment. It also pledged in June 2020 to collect demographic data to make its journals more representative of the communities it serves. From February to September 2021, it started to

ask authors and reviewers across its more than 75 journals for their gender and racial or ethnic identities (with a choice of ten categories), among other questions. Designing the categories required some market research, with a goal of being inclusive and crafting questions that are clear and easy to answer, says Sarah Tegen, a senior vice-president in the ACS journals publishing group. In December 2021, the ACS announced aggregate results from more than 28,000 responses; only around 5% of respondents chose not to disclose race or ethnicity. It noted that, among authors who gained their PhD more than 30 years ago, just under two-thirds identified as white – but among those who gained it less than 10 years ago, only about one-quarter did. Among editors of all ACS journals, 55% were white, 27% East Asian and 1.2% African/Black. Tegen says the data are a useful baseline for understanding the demographics of ACS journals.

For its part, the joint group of publishers was ready in February 2021 to consult a specialist – demographer Ann Morning at New York University – about its draft framework for asking about race and ethnicity. “It was a neat challenge,” says Morning, who advises the US government on its census process. She was intrigued by the difficulty of coming up with a standard schema that could apply across cultures. At that time, she says, publishers had thrown together a list of terms describing race and ethnicity, but they had not thought about how it would all fit together. “It was immediately obvious it was very confused.” She advised separating ethnicity and race into two questions. The first covered geographical ancestry and provided 11 options, including illustrative examples. The second covered race, in six options. (In both cases, respondents can choose not to answer.)

The draft was then sent to researchers for pilot testing, with a short accompanying survey. Of more than 1,000 anonymous respondents, greater than 90% reported their race and ethnicity, and more than two-thirds said they felt well represented in the schema. About half said they would be comfortable providing this information when submitting a paper.

The results suggest that some respondents were not willing to give information. But Falk-Krzesinski, who led the market research on behalf of the joint group, says that the response rate was much higher than expected. “Even if people didn’t feel entirely well represented, they were willing to answer. They didn’t need perfection,” she says.

Some respondents who were concerned about giving their race or ethnicity said they didn’t feel it necessary to disclose because they believed science was a meritocracy; others, however, worried about how the data would be used. The publisher group has since changed the wording of its questions to make clearer why it is collecting the



SOMETIMES YOU REALLY HAVE TO SHAKE THE SYSTEM TO FORCE IT TO CHANGE.”



Raymond Givens has tallied editor ethnicities.

data and how they will be used and stored. The information will not be visible to peer reviewers, and although collected through editorial management systems, will be stored separately, with tightly controlled access, Falk-Krzesinski says.

Publishers will meet next month to vote on endorsing the schema to roll it out into editorial management systems; they declined to share the final list of questions and categories publicly until they had reached a consensus.

The American Psychological Association (APA) in Washington DC, which publishes 90 journals, has forged its own path outside the joint group. Last year, it updated its electronic manuscript system, which had previously only invited users to give gender information and the option to answer ‘yes’ or ‘no’ for minority or disability status. Now, users can choose from 11 options describing race and ethnicity (similar to, but not the same as, US census categories), and from a wider slate of descriptors around gender identity. A blog post on this initiative noted that the data will help to set goals to develop more representative pools of authors and editorial board members (see go.nature.com/3uwkab7). In the longer term, researchers hope to study

acceptance rates for authors with various demographics to examine potential biases in peer review.

From data to policy

Babdor is not surprised it has taken publishers so long to agree on standards to collect data, because of the complexity and the fact that it has not been done before. “Every country has its own rules about how to talk about these issues,” he says.

He says that the data should be freely available so that everyone can analyse and discuss them – and that it will be crucial to look at the compounding effects of intersectionality, such as how disparity affects Black women and Black disabled individuals.

Keletso Makofane, a public-health researcher and activist at the Harvard T.H. Chan School of Public Health in Boston, Massachusetts, says that the efforts of publishers are a fantastic start. He sees a use for the data in his work – a project to track the networks of researchers who are studying structural racism. Understanding the race and ethnicity of the scientists involved in this type of work is important, he says. But it’s not just about authors and reviewers. “It’s important to look at the people who make the higher-level decisions about policies of the journals,” he says.

To engage the historically marginalized populations they hope to reach, Lerback says, publishers (and researchers studying how ethnicity affects scholarly publishing) must commit to engaging with these groups beyond simply asking for data. Most importantly, she adds, they should build trust by following up findings with action.

In the wake of her AGU study, for instance, the organization changed its article submission system with the aim of increasing the diversity of peer reviewers. It now points out to both authors and editors that the process of recommending or finding reviewers can be biased – and invites them to expand their peer-review networks.

“Data is the currency of which policy gets implemented,” Lerback says.

Holly Else reports for *Nature* from London.

Jeffrey M. Perkel is *Nature*’s Technology Editor.

1. Lerback, J. C., Hanson, B. & Wooden, P. *Earth Space Sci.* **7**, e2019EA000946 (2020).
2. Day, A. E., Corbett, P. & Boyle, J. *Chem. Sci.* **11**, 2277–2301 (2020).
3. Ye, J. et al. *Proc. 2017 ACM Inf. Knowl. Mgmt* 1897–1906 (2017).
4. Le, T. T. et al. *Cell Syst.* **12**, 900–906 (2021).
5. Davidson, N. R. & Greene, C. S. Preprint at *bioRxiv* <https://doi.org/10.1101/2021.06.21.449261> (2021).
6. Freeman, R. B. & Huang, W. *Nature* **513**, 305 (2014).
7. Bertolero, M. A. et al. Preprint at *bioRxiv* <https://doi.org/10.1101/2020.10.12.336230> (2020).
8. Kozłowski, D., Larivière, V., Sugimoto, C. R. & Monroe-White, T. *Proc. Natl Acad. Sci. USA* **119**, e2113067119 (2022).
9. Hippen, A. A., Davidson, N. R. & Greene, C. S. *Nature Hum. Behav.* <https://doi.org/10.1038/s41562-021-01279-2> (2022).
10. Wang, N. C. *J. Am. Heart Assoc.* **9**, e015959 (2020); retraction **9**, e014602 (2020).

NATHAN BAJAR/AMY/REDUX/EYEVINE