

The Uneven Spread of Citations

Minority physicists are significantly undercited compared to their white, male counterparts, which can have serious career consequences. A new tool is allowing researchers to reduce this imbalance.

By **Katherine Wright**

In early 2017, Dani Bassett was approached by then graduate student Jordan Dworkin with some intriguing data. Dworkin had delved into the reference lists of papers from a few top neuroscience journals, finding that they contained significantly fewer papers authored by women researchers than he'd expected based on the field's overall gender diversity. Dworkin was hoping to partner on a full-fledged study with Bassett, a statistical physicist at the University of Pennsylvania who uses the pronouns they/them. But Bassett was unsure about taking on the project, which at the time was not in an area that they focused on. "It's not something I'm proud of, but I hesitated for several months about working on this topic," Bassett says.

Ultimately, Bassett changed their mind, and they, along with several iterations of students and collaborators, have spent the last five years collecting and analyzing citation data from papers published in different fields of science, the most recent being



Several recent studies of citation statistics find a negative bias in the number of cited researchers from unrepresented groups.

Credit: sebra/stock.adobe.com

physics, to see if that initial finding held. Spoiler alert: it does. And these results are not isolated. The last year has seen a flurry of reports on citation practices from multiple, independent research groups, with the studies showing reduced citation rates for women and Black scientists, as well as for those from outside of a few "core" scientific countries, of which the US is included [1–3]. In some cases, papers authored by underrepresented groups are cited 14% less than expected from data-informed statistical models.

This trend indicates an underattribution of the contributions of minority scientists, Bassett says, which they note is unsurprising given that minority scientists receive proportionally fewer speaking invitations and proportionally fewer awards than white, male scientists. But whereas a small number of scientists choose who speaks at a conference or who wins a Nobel Prize, every scientist has some say in deciding who to cite in their papers, Bassett notes. Citations can also have very real and immediate consequences for every scientist's career progression. "Citation metrics are increasingly becoming the key measure of a scholar's standing in the field," says Perry Zurn, a political philosopher at the American University, Washington, D.C., a collaborator of Bassett, and Bassett's twin. "It's therefore critical to pay attention to how those [citations] are being distributed," he says.

Efforts to address citation inequalities are proliferating, with more and more researchers starting to analyze their own citation practices to search for holes. Journals are also beginning to take note of those analyses, allowing researchers to include so-called citation diversity statements (CDSs) at the ends of published papers. Initial results suggest that those efforts are working—an analysis by Kendra Oudyk of McGill University, Montreal, of a year's worth of data from journals that

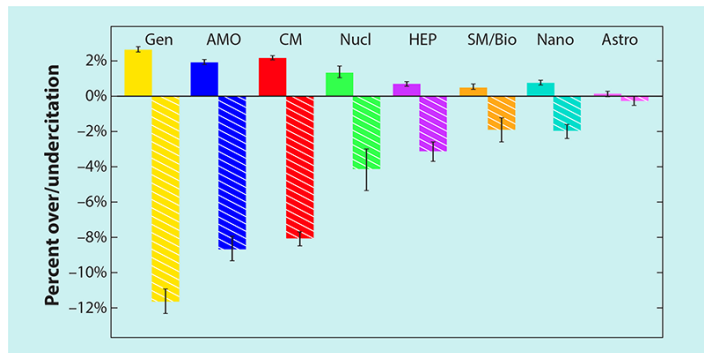
publish CDSs shows that, on average, papers that include such statements have no citation imbalances [4]. Bassett hopes that the trend of CDSs being adopted by journals and by individuals will continue and that citation inequalities will decline as that happens. “These disparities are driven by individuals and can be mitigated by individuals,” Bassett says. “We can all change how we cite.”

The Fame Factor

Bassett’s most recent study, which was led by Erin Teich of Wellesley College, Massachusetts, analyzes the citations of over one million papers that appeared in 35 physics journals between 1995 and 2020 [1]. The team fed the author list of each paper into an algorithm that used information from birth registries to determine the most likely gender of the first and last authors from the researchers’ forenames. The next step was to divide papers into different groups based on those author genders and compare the average number of citations in each group to the predicted number assuming that citation practices are uninfluenced by gender. The team used a data-informed statistical model to make these predictions.

Teich, Bassett, and their colleagues found that papers that had a woman as either the first or last author, or as both, received, on average, 3% fewer citations than those expected from the model, while those with men in those two authorship positions received 1% more. This imbalance led to a citation gap of 4%. This gap was strongest—about 14%—among the citations that appeared in so-called general physics journals, such as *Nature Physics* and *Physical Review Letters*, and weakest—less than 1%—among citations in astrophysics journals.

Next, the team looked at who is citing whom, and found that papers authored by men overcited papers by men and undercited papers by women. Teich says that this pattern suggests that the imbalances are driven largely by the citation practices of men. However, she notes that the team found that all researchers, regardless of their gender, were more likely to cite papers authored by men when citing work from outside of their immediate subfield. Teich links this finding to the fame factor; because more men win physics prizes and more men publish in “household name” journals, researchers are more likely to be aware of work done by men in other fields, she says. “Men just have more name recognition.”



Physics papers authored by men (solid bars) receive, on average, more citations than expected, whereas those that have a woman (striped bars) as either the first or last author, or as both, receive, on average, fewer citations. This citation gap is strongest for papers published in general physics journals (yellow). The gaps are also shown for more specialized journals in atomic, molecular, and optical physics (dark blue); condensed-matter physics (red); nuclear physics (green); high-energy physics (purple); soft matter/biophysics (orange); nanophysics (light blue); and astrophysics (pink).

Credit: E. G. Teich *et al.* [1]; adapted by APS/Carin Cain

In another study that came out in August 2022, a different team found significant disparities between the citation patterns of men and women scholars [2]. In that study, Kristina Lerman of the University of Southern California and her colleagues constructed citation networks for scientists elected to the US National Academy of Sciences. They found that women scholars were more likely to reciprocate citations—if one researcher cited another, then that first researcher was more likely to be cited by the second researcher in return. Lerman interprets the data as showing that women are “tightly embedded in their research communities and have very strong interactions with their peers.” She adds, “men can get away with just being famous—people cite them, but they return the favor only rarely. But women need community support to get recognition.”

Citing Like It’s 1995

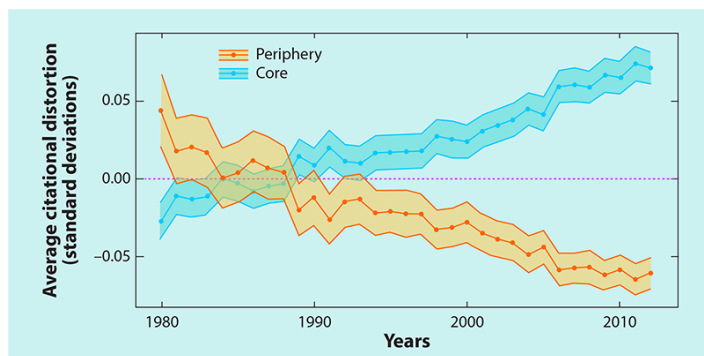
Together, these findings paint a stark picture of the citation landscape for women scientists, Bassett says. But they note that the news was not all bad. For example, Teich, Bassett, and their colleagues found that physics journals that publish more papers by women and those that publish longer reference lists tend to contain papers that cite closer to the levels expected if

citations were uninfluenced by gender. Bassett finds these correlations striking as they suggest that if researchers cited more papers or if journals published more papers by women, the citation gender imbalance could disappear. “Those are correlative findings, but they are findings we could test for causal relations,” Bassett says.

While those causal tests have yet to be carried out, the team has homed in on possible causes of the current disparities. One of those is what Bassett calls the “1995 effect.” The team’s analysis shows that most physics papers cite as if the diversity of the field has remained static for the past 25 years, which they note is not the case. Since 1995, the percentage of women professors has jumped from 3% to over 13% and women PhD students from under 6% to over 16%. The number of women at other career levels has also risen, but their representation in the most cited works has not. If that trend continues, it means that over time the imbalance will get worse, “which is what we are seeing in the data,” Bassett says.

The 1995 effect suggests that there is a significant lag between a person entering a field and their work being cited, says Christopher Lynn of Princeton University, who worked with Teich and Bassett on the physics study. “We’re still citing papers from the 1970s or 1980s or even before and missing out on the newer papers, which are more likely to be authored by women,” he says. Lynn thinks that this effect may have been exacerbated over recent years by the increased reliance of researchers on search engines, such as Google Scholar, for searching through the relevant literature. That idea is backed up by a 2008 study that found that when journals moved from being available only in print to being predominantly online, fewer unique papers were cited [5]. That finding suggests that the transition caused citation diversity to collapse, Bassett says.

When ranking papers in a list, Google Scholar accounts for factors such as the closeness of the words in a paper’s title to the search terms and the number of citations a paper has received. Older papers are more likely to have more citations, because they have been around longer, and thus get pushed to the top of the first page of the search results. The proliferation of scholarly publications also makes it difficult to keep up with all the work going on in the field, says Erika Andersson, a quantum physicist at Heriot-Watt University, UK, and an editor for *Physical Review A*. Andersson recently became aware of



Papers authored by researchers in a handful of core scientific countries (blue) receive significantly more citations than those published in the rest of the world (orange). Positive citational distortions correspond to overcitation, while negative values correspond to undercitation.

Credit: C. Gomez *et al.* [3]; adapted by APS/Carin Cain

Bassett’s work at a workshop. “It’s easy to find that you are only following the famous people and missing other work that is often more interesting,” she says.

These citation practices mean that papers authored by men are more likely to get read—and then get more citations—Lynn says. “It’s a snowball effect. Any paper that has a slight edge in the [citation] game will rise to the top.” Knowing that biases can creep in, Lynn says that he has changed how he searches the literature. Lynn no longer picks the first relevant study on the search list to back up some statement in his paper’s introduction. Instead, he clicks through, looking for other studies, perhaps with more specific data. He then also includes those studies in the citations. In doing so, he naturally sees his references cite more diverse researchers. “I hold myself accountable for each and every reference I cite in a paper,” Lynn says. “We are all busy people, but this [action] doesn’t take that much effort.”

Physicist-turned-sociologist Charles Gomez of the University of Arizona thinks that journal editors could also hold researchers accountable for their reference lists, checking to see which papers researchers included. But he acknowledges that “there is no silver bullet to fix citation inequities.” Gomez published a study that came out in May 2022 that found that papers authored by researchers at institutions in core scientific

countries—which include the US, Canada, China, Japan, South Korea, Israel, and some countries in western Europe—are significantly overcited compared to those in other places (Latin America, Africa, the rest of Asia, and the Middle East) [3]. “Modern science is thought of as an international enterprise that transcends national boundaries, but citations overwhelmingly go to researchers in just a handful of countries,” he says.

Looking Inward

Not long after Bassett was first approached by Dworkin about studying citation biases, Bassett decided to check their own work. Bassett asked Dworkin to take papers they had published since 2004 and calculate the gender diversity of the citations. “The results were horrifying. I had undercited women more than the average person,” Bassett says. Phil Chodrow, a mathematician at Middlebury College, Vermont, says he was similarly dismayed when he recently carried out a related analysis on his own work. “It’s disappointing to go through a paper that you have written and see that the majority of papers you cited were [authored] by men.”

Chodrow was spurred to investigate his citation stats after learning about CDSs. These statements provide a framework for researchers to measure who they cite and then share that information with others. The practice is advocated by all the citation-data researchers contacted for this story. Zurn says that putting a CDS together prompts him to take stock of who he has cited. “I consider myself to have my heart in the right place. But every time I get to the end of writing a paper, I’ll look and realize that I need to do some rebalancing,” he says. He adds that it’s not about beefing up or slimming down numbers for different groups. “It’s about being mindful and asking, What lineup do I have here?” The practice is no different than that of conference organizers revisiting lists of invited speakers to make sure that a variety of voices is represented, Zurn says. Lerman agrees: “It’s about being deliberate about our choices.”

The possibility of increasing diversity in citations has led journals, such as those published by Cell Press and the Biomedical Engineering Society (BMES), to provide tools for authors to put together CDSs as part of wider statements on inclusivity and diversity. Those statements can then be published alongside the papers. Sheba Agarwal-Jans and Benedicte Babayan, who work on inclusion and diversity

initiatives at Cell Press, write in a joint email that currently around a third of authors of accepted papers put together Cell’s inclusivity and diversity statements, and one quarter of those people then opt to make those statements public. Stefan Duma, the editor in chief of the *Annals of Biomedical Engineering* quoted a similar number for BMES journals.

While researchers have been “overwhelmingly positive” about Cell Press’ decision to publish CDSs, Agarwal-Jans and Babayan write that the publisher has received negative feedback about the practice. This feedback claimed that the CDSs and related tools introduce bias rather than eliminate it. In response, the publisher updated their inclusivity-and-diversity-statement form and associated materials to better explain the inequities in current citation statistics. Recognizing the contributions of minority scientists does not create boundaries; it adds “equity in a system that is inherently biased already,” Agarwal-Jans and Babayan write.

Chodrow has heard similar comments. He thinks of a CDS as a transparency measure and one tool in the toolbox to shift the culture of physics. “If people read my papers a little less and those of scholars with more diverse perspectives a little more, that’s a net positive,” he says. “And if I lose some citations because of that, that is totally fine.” Zurn says that he likes to remind people that reflecting on citations isn’t about telling researchers who to cite and who not to cite. Rather, “it’s an invitation to remember that we don’t know everything,” he says. “There are more scholars, more papers, and more projects going on than we are each aware of; taking stock of citations is about getting curious about what’s out there.”

Katherine Wright is the Deputy Editor of *Physics Magazine*.

REFERENCES

1. E. G. Teich *et al.*, “Citation inequity and gendered citation practices in contemporary physics,” *Nat. Phys.* **18**, 1161 (2022).
2. K. Lerman *et al.*, “Gendered citation patterns among the scientific elite,” *Proc. Natl. Acad. Sci. U.S.A.* **119** (2022).
3. C. J. Gomez *et al.*, “Leading countries in global science increasingly receive more citations than other countries doing similar research,” *Nat. Hum. Behav.* **6**, 919 (2022).
4. K. Oudyk *et al.*, GitHub repository, <https://github.com/koudyk/leanBibImpact> (2021).
5. J. A. Evans *et al.*, “Electronic publication and the narrowing of science and scholarship,” *Science* **321**, 395 (2008).