

Diverse Identities are Irrelevant to Science

by John Staddon

“Identity diversity” (race, gender, sexuality, etc.) makes no contribution to science. The evidence comes from history, anecdote, and from science itself. Science was invented by a very non-diverse group of mostly white males. From Newton through Maxwell, Helmholtz, and Charles Darwin it was a homogeneity of culture, if not race, that led to the novel set of beliefs and practices we now call science. With a few Asian exceptions, such as Srinivasan Ramanujan, Jagadish Bose, and others, science is a product of a shared European Enlightenment culture.

Anecdotes and simple logic also show that a relatively homogeneous group focused on a widely recognized problem, especially a problem in a controversial area, is likely to interact more freely and creatively than a self-consciously “diverse” collection of people, constrained possibly by others’ sensitivities.

And finally, science. On April 14, 2020, the *Proceedings of the National Academy of Sciences* published a widely cited article by B. Hofstra and five co-au-

thors, supported by three National Science Foundation grants, a grant from a Dutch science organization and a Stanford University computing facility. The article purports to show the virtues of race and gender diversity for scientific productivity. But it may in fact demonstrate just the opposite.

In summary, the authors write:

Diversity breeds innovation, yet underrepresented groups that diversify organizations have less successful careers within them. Does the *diversity paradox* hold for scientists as well? We study this by utilizing a near-complete population of ~1.2 million US doctoral recipients from 1977 to 2015 and following their careers into publishing and faculty positions. [emphases added]

Since essentially every Ph.D. dissertation was included in this study, science is defined very broadly indeed.

It is worth noting that a “diversity paradox” exists only if we accept the unsupported claim that diversity always makes for better science. There is no hard data to support this idea and, as pointed out above, some contrary indications. The diversity paradox evaporates if the qualification criteria for minorities

are less stringent than for others—as often seems to be the case. If selected minorities are in fact on average less well qualified, the fact that they “have less successful careers” is no surprise. There is no “diversity paradox.”

The Hofstra et al. study claims to contradict this conclusion by showing that in science “underrepresented groups” are more innovative than others, so that their lesser career success is because “their novel contributions are discounted.” But before looking at the data, it is pretty obvious that there are logical problems with the whole approach.

“Innovation drives scientific progress” the authors proclaim, which is partially correct. Yes, scientific advances are innovative, but that doesn’t mean that all innovation represents a scientific advance (this is called “affirming the consequent,” a basic logical error: a scientific advance is innovative, but an innovation may not be a scientific advance). What favors discovery is variation: trying out many things in the hope that one of them (a novel one, obviously, but one of many) is correct. The bad guesses don’t count, are rarely published, and are not part of this study’s data set.

An exception is behaviorist B. F. Skinner in his important paper “A case history in scientific method.” The paper describes in entertaining detail the combination of trial, error, and luck that led to his discovery of the Skinner box and schedules of reinforcement and their many applications. Perhaps more familiar is the story of how the double-helix structure of DNA was discovered by

Watson and Crick, not by a flash of innovation or a suitably diverse team but by two clever and persistent people trying out lots of possibilities:

[It] meant taking on the arduous intellectual task of immersing themselves in all the fields of science involved: genetics, biochemistry, chemistry, physical chemistry, and X-ray crystallography. Drawing on the experimental results of others . . . [and] taking advantage of their complementary scientific backgrounds in physics and X-ray crystallography (Crick) and viral and bacterial genetics (Watson), and relying on their brilliant intuition, persistence, and luck.

Watson gives more details in his best-seller *The Double Helix* (1968). The two men were only diverse in the sense that they had complementary scientific backgrounds. So, yes, some expertise-diversity is useful to the progress of science. Much more important is persistence and a shared goal. Race and gender are irrelevant and bad guesses are rarely published.

There are many similar examples in the history of science (at least when it is written by historians who understand the science, which is something of a rarity these days). The common feature of all successful scientists is perseverance, a willingness to tolerate failure, rather than innovation per se. As evolution shows, most new mutations are negative. In discovery, most guesses are wrong. (The wrong guess of Crick and Watson’s brilliant competitor Linus Pauling is notorious.) Persistence is the thing.

Innovation is necessary but far from sufficient for scientific discovery. Nev-

ertheless, innovation is the focus of the Hofstra et al. study. What exactly is innovation? Given a corpus of 1.2 million doctoral thesis abstracts, in every possible academic area, their method, which is basically statistical, presumes that innovation in, say, history, has statistical properties that are essentially the same as innovation in biology or physics.

To compare the performance of groups they needed to identify the gender and race of authors. Since this information is usually not directly given, they had to use statistical methods to identify the race and gender of dissertation authors: “The ProQuest dissertation corpus does not contain records of gender and race of students that filed their theses. Therefore, we predict the race and gender of students based on their first (gender) and last (race) names.” So their basic categorization is inferential. Things get worse when it comes to innovation itself.

The set of novel scientific concepts in these theses was identified using “natural language processing techniques of phrase extraction and structural topic modeling.” These concepts were not identified by their semantic meaning but by structural topic models, a statistical technique that has as its input the texts of thesis abstracts:

Structural topic modelling is an unsupervised learning technique that represents texts within a corpus as a mixture of latent thematic dimensions without a priori knowledge of what these dimensions might be. STMs rely on co-occurring words within documents . . . the model produces an optimal representation of every document as a mixture of topics.”

The outcome is what the authors term “new links” defined as “the number of unique recombinations of scientific concepts.” Whether these novel combinations are true—scientifically valid—is not measured.

The final step is to see how much impact a thesis has: “we measure how often a thesis’s new conceptual linkages are adopted in ensuing documents of each [subsequent] year.” They find that:

[T]he more students are underrepresented genders ($P < 0.001$) or races ($P < 0.05$) in their discipline, the more they are likely to introduce novel conceptual linkages (# new links) . . . Women and nonwhite scholars introduce more novelty (both $P < 0.001$) but have less impactful novelty (both $P < 0.05$) when compared to men and white students . . . demographic diversity breeds novelty and, especially, historically underrepresented groups in science introduce novel recombinations, but their rate of adoption by others is lower, suggesting their novel contributions are discounted.

(The significance levels are high because the sample is large.) Possibly more of the “new links” associated with men are of greater validity? Again, we don’t know, since scientific truth is not considered in the study.

This is a very complex study; the models used will be incomprehensible to any but an expert. No effort is made to compare the “conceptual linkages” identified by STMs with the semantic structure of the analyzed document or the validity of its conclusions.

The basic finding seems to be that (a) the topics identified by their models are more numerous for “demographically

underrepresented” (usually female and black) authors than the majority (usually white males); (b) white male topics have a longer life—“are taken up by other scholars” at higher rates. Minority contributions are “less impactful” because they are “discounted.” The same asymmetry is reflected in the generally less successful careers of the “discounted” groups.

The obvious question raised by the study is: which data are a more valid measure of scientific value, the statistically extracted estimates of innovation, or the success or failure of the “innovative” topics? In other words, are minority innovations “discounted” for good reasons? This question is not addressed by the study. Since the definition of “topic” is opaque and the identity of authors is subject to error, it is hard to give the results much credence. It is at least as likely that the fate of “innovations” in years after the dissertation is a more reliable measure of their scientific quality than the results of this tortuous analysis. In effect, the study is perfectly consistent with the counterclaim that the scientific innovations of racial and gender minorities are “discounted” because they are false or less fruitful than others.

It is fashionable to believe that minorities are “underrepresented” because of race or gender bias, (which are unmentioned but implied in the study) rather than for the lesser quality of their scientific work. Hence this flawed study has gone unchallenged, even though the data are in fact consistent with the idea that the lesser success of minority

ideas (“links”) reflects lesser quality rather than racial or gender bias. The most parsimonious conclusion is that there is no “diversity paradox.” But the study is cited because it looks “scientific,” sits on a prestigious perch, and conforms to the prevailing ideology.

John Staddon is James B. Duke Professor of Psychology, and Professor of Biology emeritus at Duke University; jers@duke.edu. Three of his most recent books are *The New Behaviorism: Foundations of behavioral Science*, 3rd edition (Psychology Press, 2021), *Unlucky Strike: Private Health and the Science Law and Politics of Smoking*, 2nd edition, (2022 PsyCrit Press), and *Science in an Age of Unreason* (Regnery, 2022). He last appeared in these pages in winter 2022 with “*Stratification Economics: How Social Science Fails.*”
