Gender, Race/Ethnicity and Research Funding

Donna K. Ginther

Roy A. Roberts and Regents Distinguished Professor of Economics
Director, Institute for Policy & Social Research
University of Kansas
Research Associate, National Bureau of Economic Research

March 22, 2022

Abstract:

This paper examines recent data on gender and race/ethnicity differences in research funding. Data on gender differences in funding in the United States finds few instances where female investigators are disadvantaged. In fact, female investigators are more likely to receive research funding at the National Science Foundation and equally likely to be funded at the National Institutes of Health. In the 28 countries of the European Union, female investigators are 3% less likely to receive research funding. The data on funding rates by race/ethnicity is not reported by European funders nor by the National Institutes of Health. Piecing together data from a variety of sources indicates that Black or African American investigators are less likely to receive research funding at both the National Institutes of Health and the National Science Foundation. In addition, Asian investigators have the largest funding gap compared to white investigators at the National Science Foundation. Many studies have examined the gender gap in funding with the majority finding no significant funding gap. However, persistent gender gaps in funding were found in Canada, and the review process may adjust for gender bias in initial reviews in some European contexts. The study concludes with a discussion of bias in peer review. Four studies of the NIH funding process found limited evidence of bias against Black or African American investigators. However, in a study that anonymized proposals, white investigators received higher scores when their names were known to the reviewers.

Acknowledgements: Thanks to Sofia Berrospi Fernandez for research assistance and Carolyn Caine for editorial assistance. Any errors are my responsibility.

Many researchers have examined gender differences in science careers (see Kahn and Ginther 2018, Ceci et al. 2014, Cruz-Castro & Sanz-Menéndez 2020, Cruz-Castro, Ginther & Sanz-Menéndez forthcoming). These studies have mostly found that women publish fewer papers and receive fewer citations. However, the literature on gender differences in research funding is decidedly mixed. Far fewer researchers have examined race/ethnicity differences in science careers (National Center for Science and Engineering Statistics 2021, Ginther et al. 2011, Ginther et al. 2018). These studies focus exclusively on the United States and find that Black or African American (Black/AA) and Asian researchers are significantly less likely to receive research funding (Ginther et al. 2011). Since funding is essential for research in most scientific fields, having a deeper understanding of research funding gaps is warranted. This paper reviews the literature and provides new data on gender and race/ethnicity differences in research funding. This work also examines the impact of the peer review process on these funding gaps.

Throughout, I focus more on the United States and specifically on federal funding agencies. Some data and research on gender differences in funding from European countries will be presented; however, these countries do not systematically track data by race. Private funders such as foundations do not typically share their data on proposals and awards by demographic characteristics, and the review process differs significantly across these organizations. Thus, the role of foundations in funding gaps will not be considered.

Before a detailed discussion of the demographics of research funding, it is important to understand the details of the funding process. Here I focus on the National Science Foundation (NSF) and the National Institutes of Health (NIH). In fiscal year 2021, the NSF allocated $7.3 billion to research and related activities.[1] In the same year, the NIH allocated $32.3 billion on

[1] National Science Foundation. *FY 2021 Agency Financial Report (AFR) – List of Figures and Tables*. https://www.nsf.gov/pubs/2022/nsf22002/tables.jsp

56,794 competing and non-competing projects. That money funded 11,229 competing research project grants with a total expenditure of $23.3 billion.[2] To allocate this money, these agencies have developed a standardized peer review process.

An organization, typically a university, submits a proposal, and the NSF program officer in the relevant scientific division assigns peer reviewers to provide written reviews. The NSF review process evaluates the intellectual merit and broader impacts of the proposed research. While intellectual merit is about knowledge creation, NSF's broader impacts deserve additional explanation. These are defined by the NSF as: *"Broader impacts may be accomplished through the research itself, through activities that are directly related to specific research projects, or through activities that are supported by, but are complementary to, the project. NSF values the advancement of scientific knowledge and activities that contribute to the achievement of societally relevant outcomes."*[3] Broader impacts can range from training future scientists as research assistants or postdoctoral researchers on an NSF-funded project to providing scientific results that could benefit society. Proposals are evaluated using a scale of excellent, very good, good, fair, and poor in written reviews. A review panel of anonymous scientific peers discusses the reviews. After the panel discussion, proposals are ranked as competitive or not-competitive for funding. Using the panel feedback, the NSF program officer makes a funding recommendation.

The National Institutes of Health review system has a more elaborate structure. Based on the scientific content of an NIH proposal, the Center for Scientific Review (CSR) assigns a proposal to one of the NIH Institutes or Centers (ICs) or to a Study Section. NIH Study Section

[2] Lauer, M. 2022: https://nexus.od.nih.gov/all/2022/03/07/fy-2021-by-the-numbers-extramural-grant-investments-in-research/
[3] Wilbert and Tejasputra 2021: https://beta.nsf.gov/science-matters/nsf-101-five-tips-your-broader-impacts-statement

members are publicly reported, and individuals submitting a review proposal know their names and affiliations. The Study Section is the first step in the NIH review process. Here proposals are reviewed by Study Section and ad-hoc reviewers using five evaluation criteria: significance of the research question, qualifications of the investigators, innovation of the science, approaches used in the proposal, and the quality of the research environment. A reviewer will assign a score of 1 = exceptional to 9 = poor for each of these criteria. After compiling written reviews and criterion scores, proposals are divided into discussed and not-discussed categories. Discussed proposals will receive an overall impact score and ranking. Not-discussed proposals are of sufficiently low quality and will not be eligible for funding; this ends up being about half of the proposals submitted. At the end of the scoring process, the Study Section reports the review scores to the IC's selection process. During the second step of NIH review, those applications with the lowest (best) scores receive a secondary level of review by NIH IC staff. This step determines the funding decision. NIH program staff examine the applications' summary statements and review scores and weighs these factors against the needs and priorities of the IC. The IC director or their delegate makes the final decision as to whether to offer an award. Most NIH research proposals are funded in the order of their overall priority score and ranking[4] (Jacob and Lefgren 2011); however, in some NIH funding mechanisms, ICs have demonstrated significant levels of discretion (Ginther and Heggeness 2020).

In addition to the two-step review process, it is important to understand the various funding mechanisms offered by the NIH. NIH provides funding from training grants to research independence. Training grants to institutions (T32s) provide support for graduate students and postdoctoral researchers. The F31 mechanism is a mentored predoctoral research fellowship to

[4] At NIH low scores are the best scores.

graduate students and the F32 mechanism is a mentored postdoctoral research fellowship. Once a researcher is a faculty member, they qualify for mentored K awards to develop work for research independence. Research independence is achieved at NIH once individuals receive their Type 1 R01 (or equivalent) award.[5] A Type 1 R01 is for a new research project, whereas a Type 2 is a competitive renewal of a Type 1 award.[6]

At both the NIH and NSF, reviewers see the names of the principal investigators. Race is not reported on proposals, although reviewers can likely guess gender from the investigator's name. The investigator's biographical sketch includes institutions and years of degrees, current and previous institutional affiliations, a subset of the investigator's publications, and current and pending funding.

In a history of the peer review system, Baldwin (2018) indicates that peer review was a post-war phenomenon designed to boost the credibility of the scientific funding process in the face of Congressional opposition. She summarizes her argument as follows:

> The massive expansion of government funding for scientific research in the postwar United States led to more scrutiny of science—and to suggestions that scientists should be more accountable to the public. Scientists balked at the suggestion that their methods or conclusions might be vetted by scientific laymen but did not want to surrender the public status or funding opportunities they had gained. A controversy over peer review procedures at the National Science Foundation (NSF) in the 1970s highlighted the conflict between scientists' desire to keep scientific decisions in expert hands and the belief that public funds made scientists answerable to laymen and legislators. At a hearing about the NSF's review procedures, various stakeholders argued that peer review was the only acceptable method of selecting which proposals to fund, and the outcome of the controversy placed increased emphasis on referee opinions at the NSF. The episode both reflected and helped cement the view that peer review was central to the proper practice of science. (Baldwin 2018, p. 539)

---

[5] R01-equivalent funding mechanisms include DP1, DP2, DP5, R01, R37, R56, RF1, RL1, U01 and R35.
[6] You can read more about NIH training programs here: https://researchtraining.nih.gov/programs/career-development.

Despite being deemed "critically important to the proper practice of science," peer review has both its supporters and detractors (Fang and Casadevall 2016; Fang et al. 2016; Gallo et al. 2016; Li and Agha 2015; Li 2017; Pier et al. 2018; Ginther and Heggeness 2020). The key question that I will return to at the end is whether the peer review process exhibits gender or race/ethnicity bias. Before that discussion, it is important to establish some stylized facts about research funding in the United States and Europe.

The Applicant Pool

The grant submission process provides a useful approach to understanding demographic differences in funding. It starts with an application, proceeds to peer review, and concludes with funding decisions. The application process depends on the pool of scientists applying for funds. I focus on the applicant pool in the United States because it is for a single country as opposed to 28 countries in the European Union. A priori, the pool of faculty at colleges and universities would be one way to measure the size of the applicant pool. However, not all academic institutions engage in federally funded research. According to the Higher Education Research and Development Survey, the top 150 research universities in the United States received 90% of all research funding awarded in 2019. The top 300 research universities received 98% of all research funding. Thus, faculty at research institutions provide a better measure of the funding applicant pool.

The National Center for Science and Engineering Statistics (NCSES) is mandated by Congress to prepare the report *Women, Minorities and Persons with Disabilities in Science and Engineering* (2021). This report contains estimates of faculty by gender and race/ethnicity as well as self-reported federal research funding at four-year colleges and universities in the United

States in 2019.[7] I supplement this with additional data on faculty at research institutions collected by Academic Analytics.[8] Academic Analytics works closely with universities to ensure accurate lists of faculty and to identify departmental affiliations. Academic Analytics has collected data on individual faculty members since 2009. Information on faculty member publications, citations, and most importantly, federal research grants has been collected since 2004. For each faculty member I have measures of gender, academic institution, journal articles, citations, grants, faculty rank and academic department. Academic Analytics does not identify the race of faculty members. The 2018 data have data from 390 separate institutions in the US. I report these data by gender in Table 1.

Gender Differences in the Applicant Pool and Funding Rates

Estimates from the Survey of Doctorate Recipients (SDR), 2019 show that women make up one-third of science and engineering faculty (defined as holding assistant, associate, or full professor rank at a research university). However the female share varies by field: women are the majority of faculty in psychology, over 40% of faculty in social science, and 36% of faculty in life science. In mathematics and physical sciences one-quarter of faculty are female, but only 18% of computer science and 16% of engineering faculty are female. I report the same female share of faculty in research-intensive universities using the Academic Analytics faculty data. The share of female faculty is significantly lower in the fields of life science, mathematics, physical science, and psychology at research-intensive universities. Interestingly, the share of female computer scientists and engineers is higher at these institutions.

[7] These data are based on estimates from the 2019 wave of the Survey of Doctorate Recipients.
[8] Academic Analytics, LLC 2020: https://academicanalytics.com/aarc-scholarly-research/

The 2019 Survey of Doctorate Recipients asked respondents whether they received federal funding: "Was any of your work during 2018 supported by contracts or grants from the U.S. federal government?"[9] Note that this asks whether work was funded and not whether a person was principal or co-investigator. An SDR respondent could be funded as a subaward on someone else's project. If this happens often, then the SDR measure of funding would overstate the number of investigators leading research projects. In contrast, Academic Analytics matches faculty names to federal funding databases like *NIH REPORTER*[10] that list investigators and co-investigators. Thus, the grants attributed to faculty in the Academic Analytics database correspond directly to the investigators leading research projects. Using these data, I calculated the share of faculty who received at least one federal research grant between 2004 and 2018. This will tend to inflate the share with funding compared to the SDR which was whether or not a person had funding in 2018.

In Table 1, Assistant, Associate, and Full Professors at 4-Year Colleges and Universities by Gender, Broad Science Field and Research Funding, 2018, 39% of female science and engineering faculty report federal funding compared to 44% of male faculty. As expected, the numbers are higher at research institutions, with roughly half of female (48%) and male (51%) having research funding. Focusing on the research-intensive faculty, research funding varies dramatically across fields. More than half of faculty in life science, computer science and physical science fields are funded. About one-third of psychology faculty and just over 20% of social science faculty at research universities are funded.

[9] NCSES 2019 Survey of Doctorate Recipients: https://www.nsf.gov/statistics/srvydoctoratework/surveys/srvydoctoratework-2019.pdf
[10] NIH RePORTER: https://reporter.nih.gov/

To get a deeper sense about gender funding gaps, I use the Academic Analytics data to compare whether a faculty member was ever an investigator on a grant between 2004 and 2018, the total number of federal grants during that time, and the cumulative dollars in funding. Table 2 shows these estimates by broad field of science. I find that female faculty are significantly more likely to have been investigators than male faculty in physical science, social science and engineering. The positive gender gap in physical science is eight percentage points. Male faculty are significantly more likely to have received grants in psychology. In other fields, there are no significant differences. Remarkably, there are few significant differences in the total number of grants. Female faculty have about one-third fewer grants in the life sciences and small but significant (0.13) fewer grants in psychology. Finally, female faculty have received significantly less money in life science (-$750,000), computer science (-$250,000), and physical science (-$370,000).

The Academic Analytics data show few significant gender gaps in research funding. As a final thought experiment, I estimate the application rate by using the number of applications to the NSF science directorates of Biological Science, Engineering, Social Behavioral and Economic Sciences, and the combination of Geosciences and Mathematical and Physical Sciences. I use the number of research faculty in life science, engineering, social and behavioral science, and mathematics and physical sciences from research-intensive universities as the denominators (from Table 1). The gender differences in application rates are presented in Figure 1. In every NSF Directorate considered, the female application rate is higher. Female faculty in engineering are five percentage points more likely to apply for funding. In the combined directorates of Geoscience and Mathematical and Physical Sciences, the female-led application rate is a huge 17 percentage points higher.

Gender Differences in Funding Rates

The Association of Women in Science (AWIS) was founded in 1971 to advocate on behalf of women in science. One of its first acts was to sue the NIH to compel the agency to assign more women to NIH's committees. According to the AWIS website, "*AWIS et al. v Richardson and Marston*—National Institutes of Health (NIH) is required to freeze appointments to study sections and advisory groups until women are given the chance to suggest names for the open positions. AWIS files suit against the NIH for underrepresentation of women."[11] AWIS won this lawsuit later that decade. It has been fifty years since AWIS advocated for greater female access to NIH funding. I now consider whether female investigators have made progress at NIH.

The analysis begins with a review of funding rates for research project grants (RPGs) by gender over time at NIH (Figure 2). Funding rates are the ratio of awards to applications. In 1998 at the start of the doubling of the NIH budget (Couzin and Miller 2007), female investigators had funding rates that were two percentage points lower than males. This gap persisted through the NIH doubling and beginning in 2003 the rates converged. Since 2017, female funding rates have been slightly higher than male rates. Figure 3 shows data from the NSF for male and female investigators. Starting in 1999, female investigators were equally likely to receive NSF funding compared to male investigators. Every year since, female funding rates exceeded rates for male investigators at NSF.

The European Commission reports on gender and research and innovation periodically in the *She Figures* report (European Commission 2021). I show the funding rates by gender in 24

[11] Association for Women in Science 2021: https://www.awis.org/about-awis/awis-history/

European countries as well as the combined EU 28 countries for 2019 (Figure 4). In 11 of the 24 countries shown, female investigators have higher or the same funding rates as male investigators. For the EU 28 countries combined, female investigators had a 3 percentage point lower funding rate.

Since the distribution of female investigators and male investigators across fields differs, I also provide data on funding by broad field of science. The NSF provides gender funding rates by scientific directorate, and these data are show in Figure 4. In every directorate (field of science) women have higher funding rates than men in 2019. The European Commission reports funding by broad field of science. In Europe, male and female investigators have the same funding rates in engineering, and women have slightly higher funding rates in agriculture. In the remaining fields, male funding rates are higher.

As mentioned earlier, funding rates are calculated based on an investigator's decision to apply. It could be that female investigators are positively selected in funding because only the best apply. There is some evidence for that given the relatively low share of female applicants in the US. Using data from NSF for 2019, the female share of applications in physical sciences is 20%, engineering is 17%, and social science is 30%. Data from the NIH for the same year shows that the female share of applications is 34%. The female share of EU-28 applications is much higher across all fields at 34% for natural (physical) science, 25% for engineering, 45% for medical/health, and 49% for social science.

Taken together the data from the US indicate that female investigators are equally successful at NIH and are more likely to apply for and receive funding at NSF. Using Academic Analytics data, I found a limited number of significant differences that point to female disadvantage in ever receiving a grant, number of grants and even cumulative dollars. There is a gender funding gap in Europe of about 3 percentage points, but in 11 out of 24 countries considered, female investigators are doing the same or somewhat better than their male colleagues. It is interesting and worthy of additional work to understand why the female share of grant applications by field is much higher in the EU-28 countries than in the United States.

Race/Ethnicity Differences in Research Funding

While the data point to few gender gaps in research funding, the same is not true for race/ethnicity. The Academic Analytics data does not have a measure of race/ethnicity, nor does Europe collect research funding data by race. It is helpful to review the race/ethnicity applicant pool using data from the SDR.

Race/ethnicity Applicant Pool

Table 3 shows the share and count of faculty by race/ethnicity for non-Hispanic whites (whites), non-Hispanic Asians (Asian), non-Hispanic Black or African American (Black/AA), and Hispanic or Latino of any race (Hispanic). The table omits Native Americans, Native Hawaiians, other race, and multiple race because in several instances, counts of assistant, associate or full professor by science field was suppressed because the numbers were so low that they would compromise confidentiality. The numbers are sobering. The share of white faculty is 70% or more in all fields of science except for computer science, physical science and engineering. In those fields Asian faculty and white faculty together make up over 90% of the faculty. Black/AA faculty are less than 4% of total faculty in every field except for psychology and social science. As of 2019, there are 140 Black/AA computer scientists who are ranked faculty members, 240 mathematical scientists, and 620 physical scientists. The numbers look very similar for Hispanic faculty, with only 60 in computer science. The share of Hispanic faculty in other fields range from 4 to 6%.

The share of faculty reporting funding by race is also instructive. Sixty percent of white, 70% of Asian, 45% of Black/AA and 53% of Hispanic faculty in life science fields report funding. Except for social science and psychology fields, Asian faculty report the highest level of funding followed by white, Hispanic and Black/AA faculty. Since these are based on the SDR self-reported data, they may overstate funding because of subawards and senior/key personnel roles.

Race/Ethnicity Differences in Funding Rates

A significant lack of data hampers our ability to show trends in race/ethnicity differences in research funding. NSF provides funding rates by race/ethnicity, but NIH does not. Figure 5 shows funding rates by race/ethnicity at NSF from 2009 to 2019. White investigators have the highest funding rates across time, followed by Hispanic, Black/AA and Asian investigators. Although Asian investigators report having the highest level of funding in the SDR, their funding success rates are consistently 5-10 percentage points lower than whites. This suggests that Asian investigators are either submitting a higher number of applications, are more likely to have subawards, or both.

Using NSF data from 2019, I make a (perhaps ill-advised) attempt to compare funding by race/ethnicity at NIH and NSF. Instead of providing funding rates, the NIH reports the number of principal investigators in a given year by race who had research grants.[12] That said, I will attempt to show a comparison of funding rates by race/ethnicity at NSF and NIH. Using data from NIH's UNITE report (NIH Advisory Committee 2021) for the year 2020, I will compare funding by

[12] These data are not comparable to the information reported by gender by NIH.

race with several caveats. The funding rates reported in the UNITE report (NIH Advisory Committee 2021) are the number of awards divided by the number of applicants. This is not the same as NIH funding rates by gender (awards divided by applications) shown in Figure 2. The denominator is smaller (by focusing on applicants instead of applications) because many investigators submit multiple applications. Thus the NIH award rates by applicant are not directly comparable to funding rates.

Nevertheless, I show NSF funding rates by race/ethnicity and NIH award rates by applicant in Figure 7. Some familiar patterns emerge. White investigators are more likely to be funded at NIH and NSF than any other race/ethnicity. Hispanic investigators (any race) have slightly lower but remarkably similar funding rates at NSF and NIH. Black/AA and Asian investigators have lower funding rates at both agencies. Black/AA applicants have an award rate of 24% at NIH compared to 31% for white applicants. The NSF funding rate is somewhat higher at 26% for Black- or AA-led applications. Asian applicants have lower funding rates at 27% at NIH, but the funding rate for Asian-led applications at NSF is the lowest of any race/ethnicity at 23%. These data show that the NIH is not alone in having funding gaps by race/ethnicity.

While gender gaps in funding are small or non-existent, the racial diversity of the applicant pool and the significant race/ethnicity differences in funding success underscore problems beyond those documented at NIH (Ginther et al. 2011). I will now review some relevant literature on gender and race/ethnicity gaps in research funding as well relevant work on peer review.

Explanations for the Research Funding Gap

In what follows, I provide a summary of research findings on gender and race/ethnicity differences in research funding. While there is an extensive literature, the methods used can be classified into the categories of descriptive analysis of awards, anonymization studies, analysis of applications, and analysis of applications and peer review.

The first part of this paper engaged in descriptive analysis of funding success rates. Data on success rates by gender or race/ethnicity tell only a small part of the story. Some studies like Oliveria (et al. 2019) make gender comparison on the share of awards to female investigators, and differences in the amount of money that awardees receive. While this ex-post analysis is interesting, we are interested in whether the review process treats equivalent proposals the same. Typically, administrative data is required to perform these kinds of studies. While NIH has used its administrative data to study the peer review process (Pohlhaus et al. 2011, Ginther et al. 2011, Lauer et al. 2021), NSF has only shared its administrative data a handful of times (Broder 1993).

The ideal research design uses anonymization. In this case, identical proposals would be submitted with random assignment of names, race, and gender to a review panel and then compare the peer-review evaluations of “male” and “female” (“white” or “Black/African American”) (see Forscher et al. 2019 as an example). A second approach would anonymize proposals and send them through review panels with and without names attached (see Nakamura et al. 2021 as an example). Both of these approaches would identify the causal effect of gender or race/ethnicity on the evaluation of the proposal. Both of these approaches are also time-consuming and expensive (Nakamura et al. 2021).

This literature review will focus on analysis of applications and analysis of applications coupled with the peer review process. Analysis of applications requires the use of administrative data and control variables associated with the investigator’s biographical sketch to evaluate the review process (see Ginther et al. 2011, Hoppe et al. 2019 for example). The goal of this approach is to explain whether differences in observable characteristics contribute to gender and race/ethnicity differences in research funding. An alternative approach that does not require administrative data uses the stages of NIH funding from training to research independence (receiving an R01 research award) to study gender differences in research funding (Ley and Hamilton 2008, Jagsi et al. 2009). Finally, some researchers have access to both administrative data and the text of peer reviews (Bol, de Vaan & van de Rijt 2022, for example). Clearly, without access to administrative data on both proposals and awards, it is very difficult to obtain credible explanations for observed differences in research funding.

Explanations for Gender Differences in Research Funding

The publication of Wennerås & Wold (1997) launched the literature on using application data to evaluate gender differences in research funding. They used applications to a postdoctoral fellowship program in Sweden and found that female applicants were disadvantaged. To review the many studies that followed, I used these criteria to identify studies to discuss: 1) studies should use data from applications for their analysis; 2) studies should have been published since 2005 in order to address the more recent situation for women; and 3) meta analyses are excluded because different countries and contexts make it difficult to interpret the results. A list of these studies appears in Table 4. In addition to the citation, the table lists the country and funding agency, the study population and method used (award data, applications, or applications and peer review), and a brief summary of findings.

Many studies have compared the share of awards to female and male investigators and examined aspects of the gender differences. I highlight two examples here. Ley and Hamilton used the entire NIH awards database and assigned gender to the investigators. They then compared the percentage of female applicants and success rates of female applicants to each career stage of the NIH funding process: loan repayment programs, training grants (K01, K23, K08), first research grant (NIH R01 Type 1), and R01 awards to experienced investigators. Three main findings emerged from Ley and Hamilton's analysis. First, female investigators had similar success rates to male investigators for most funding mechanisms. Second, the female share of applicants to NIH programs dropped precipitously after training grants. Third, female investigators were less successful as experienced investigators, usually the Type 2 R01 renewals of previously funded R01 projects. Many researchers have used this approach in other contexts (see Cheng et al. 2016 or Head et al. 2013). Oliveira et al. (2019) provides another example. They analyzed 53,903 NIH awards to first-time recipients and compared the amount of research funding awarded by gender. Female investigators were disadvantaged by $39,106 compared to males. In the 10 highest funded mechanisms, this funding disadvantage dropped to $10,527. Both of these studies demonstrate how awards data can be used to gain a deeper understanding of the dynamics of gender differences in funding, but this approach does not provide explanations for these gaps, nor do they identify a causal effect of gender on funding outcomes.

Other papers have used the career path of funding at NIH to examine gender differences in the transition from training grants to research independence (an R01-equivalent award). Jagsi et al. (2009) identified 2,784 K awardees from 1997-2003 and examined the share that achieved research independence. They found that female NIH trainees were nine percentage points less likely to receive R01 funding. Kalyani et al. (2015) took a similar approach with more-recent NIH data and found that female K-awardees were significantly more likely to reach research independence than males.

Table 4 shows 15 studies that use application data to examine gender differences in research funding. Countries include the United States, United Kingdom (UK), Sweden, Switzerland, Canada, Austria, and Belgium. Ten studies found no significant gender gap in research funding in the UK (2 studies), US (5 studies), Sweden (1 study), Belgium (1 study), and Switzerland (1 study). The remaining four studies are Wennerås & Wold (1997) and three studies from Canada (Tamblyn et al. 2018, Witteman et al. 2019 and Burns et al. 2019). The results for Canada all focus on the Canadian Institutes of Health Research and are remarkably consistent in showing a 3 to 4% funding disadvantage for female investigators.

While some studies found no evidence of a gender gap in research funding (e.g. Ginther, Kahn and Schaffer 2016), others have more nuanced findings. Beck and Halloin (2017) found no gender differences in funding, but they also found that males were more likely to apply and receive higher credit for their previous research. Others found that females were less likely to apply in the US (Waisbren et al. 2008) and the UK (Boyle et al. 2015). While Pohlhaus et al. (2011) found no gender differences in funding success rates for most NIH mechanisms, they did find that success rates were lower for female investigators in the Type2 R01 mechanism. They also found that female investigators received larger R01 funding amounts but had fewer NIH R01 awards than male investigators.

Studies that Examine Gender and Peer Review

The next type of study combines information from applications with an analysis of peer review outcomes. Here the findings are more mixed. Mutz, Bornmann and Daniel (2012) examined 8,496 proposals with 23,977 reviews. They found no significant gender differences in funding. Interestingly, scores were lower for both male and female investigators when the reviewers were majority female. Both van der Lee & Ellemers (2015) and Bol, de Vaan & van de Rij (2022) analyzed applications and reviews from the Netherlands Organization for Scientific

Research Talent Program (NOW). Van der Lee & Ellemers (2015) found that female investigators were 2.8 percentage points less likely to be funded. They then analyzed the review panel's evaluation of applicants during the interview phase. Female investigators were rated lower on "quality of researcher" despite having no significant gender differences in the "quality of the proposal." This study did not control for an applicant's prior research record. Albers (2015) and Volker & Steenbeek (2015) criticized van der Lee & Ellemers (2015) findings as being a statistical artifact of having a large number of female applicants in social science fields and that also had lower award rates.

Bol, de Vaan & van de Rij (2022) revisit the NOW Talent Program with the goal of unpacking gender differences in evaluation throughout the review process. The process begins with pre-selection where panelists review proposals when the number of applications is over four times the number of grants available. Based on these pre-review scores, the top proposals (equal to exactly four times the number of grants available) are sent to external reviewers. Once external reviews are complete, the applicants have an opportunity to submit a two-page rebuttal. The external reviews, rebuttals and proposals next go to the review panel that uses this information to collectively score the proposals. At this stage, two times the number of grants available are invited for an interview with the panel. At the end of the interview stage, a final priority score (1 = best; 9 = worst) is assigned to the proposals. Bol, de Vaan & van de Rij (2022) found that female investigators receive 40% of a standard deviation lower scores during the external review process. During the final panel discussion, the panel adjusts for these gender differences, and this results in no gender gap in funding. The authors estimate that the panels redistribute approximately 64 million euros in funding from male to female investigators.

Van den Besselaar & Mom (2020) found similar results when examining the 2014 Starting Grant of the European Research Council (ERC). They control for prior merit measured by publications and associated bibliometrics, quality of the collaboration network, number of coauthors and previous grants. After controlling for merit, female investigators received lower scores in the first selection process than males. After the first selection, only 25% of the top-scoring applications proceed to the second step where funding decisions are made. At the second stage, grants from female investigators have a higher probability of being funded. Thus, while there is not a gender gap in overall funding, the bias against female investigators in the first stage and in favor of female investigators in the second stage "nets out."

Finally, three papers from related research teams have evaluated the summary statements of the NIH review process (Kaatz et al. 2015, Kaatz et al. 2016 and Magua et al. 2017). These papers featured text analysis of reviews and summary statements of NIH applications from the University of Wisconsin-Madison. Kaatz et al. (2015) found more praise and expressions of competence for proposals from female investigators. Kaatz et al. (2016) built on their prior study to examine whether the description of male and female investigators was associated with their NIH criterion scores. They found that female proposals had worse priority, approach, and significance scores in Type 2 R01 applications, despite receiving high praise in the written reviews. Using a set of reviews for Type 2 R01 applications at the University of Wisconsin, Magua et al. (2017) found that male investigators were characterized as "highly innovative" while female PIs were characterized as having "expertise." Male investigators received better scores that were associated with higher levels of productivity. Taken together, these authors argue that these differences in the descriptions of investigators suggest that gender stereotypes are influencing the peer review process. However, it is important to recall that there is only evidence of a gender gap in NIH Type 2 R01 awards and not in other NIH funding mechanisms.

Explanations for the Race/Ethnicity Gap in Funding

Far fewer studies have examined race/ethnicity differences in research funding, and all of them are focused on the NIH. This research began with Ginther et al. (2011) that used over 83,000 NIH Type 1 R01 applications from 2000 to 2006. They found that Black/AA researchers were half as likely to receive an NIH Type 1 R01 award as white investigators. After controlling for educational background, country of origin, training, previous research awards, publication record, and employer characteristics, the funding gap fell to one-third (Ginther et al. 2011). Asian investigators also were significantly less likely to receive NIH funding, although this gap was being driven by foreign-born Asian investigators. The funding gap for Hispanic investigators was fully explained by observable characteristics.

Building on these findings, (Ginther et al. 2012) considered whether the funding gap was similar among Black/AA and white MDs in medical schools. They found that the gap was narrower for MDs in medical schools, and much of it could be explained by proposals that included human subjects (Ginther et al. 2012). Next, Ginther, Kahn and Schaffer (2016) examined the intersectional effects of both gender and race/ethnicity on NIH funding. They found that white female investigators were slightly more successful than white males at receiving NIH R01 awards, but that Black/AA, Hispanic and Asian female investigators faced disadvantages of the same magnitude as male investigators of those race/ethnicities.

Finally, Ginther et al. (2018) revisited the impact of prior research funding, publications and associated bibliometrics, and NIH training on the Black/AA/white NIH funding gap. After hand-coding 2,400 biosketches that accompanied NIH proposals from 2003-2006, they found that prior NIH training and research funding explained only a small portion of the Black/AA/white funding gap. In contrast, with improved measures of publications, they found that Black/AA researchers published fewer papers, had fewer coauthors and were cited less and that these factors explained 50% of the Black/AA/white funding gap.

Researchers funded or employed by NIH have further examined the sources of the Black/AA funding gap and the NIH review process. However, the results of this work cannot be directly compared to the previous studies (Ginther et al. 2011, 2012, 2015, 2018) because of differences in how the sample of R01applications is defined.[13] Hoppe et al. (2019), examined each stage of the NIH review process to evaluate whether grant topic choice could explain the Black/AA/white funding gap. They found that Black/AA researchers chose topics that were less likely to receive funding. However, their study found that topic choice was only salient once the analysis is limited to those proposals that are discussed and received a priority score. Thus, the funding gap investigated in previous work by Ginther et al. (2011, 2012, 2015, 2018) and the funding gap where topic choice had explanatory power were not the same.

Lauer et al. (2021) reexamined the topic choice result controlling for the success rates at NIH Institutes and Centers (ICs) that received these proposals. They found that Black/AA investigators were more likely to be reviewed by ICs that had significantly lower funding rates. This suggests that part of the Black/AA funding gap may be explained by Simpson's paradox as was argued in the case of gender differences in funding in the Netherlands (Albers 2015 and Volker & Steenbeek 2015).

[13] The Ginther et al. (2011) and subsequent papers limited the sample to the last submission of a Type 1 R01 application. Thus, a proposal that was submitted twice, with the first time being rejected and the resubmission being funded would be counted as a funded proposal. Other researchers have combined Type 1 and Type 2 applications or "double-counted" proposals that were resubmitted.

Other researchers have used institutional data to study race/ethnicity funding gaps. Lewit et al. (2021) matched faculty at the University of Tennessee Health Science Center to the NIH Reporter Data to examine gender and race/ethnicity differences in surgical departments. They found that no Black/AA or Hispanic women received R01-equivalent grants in 2019.

Studies that Examine Race/Ethnicity and Peer Review

Four papers have examined whether the NIH review process is subject to bias. Erosheva et al. (2020) examined whether the process of combining criterion scores into overall impact scores (know as commensuration), was biased against Black/AA NIH investigators. They found that criterion scores are not subject to commensuration bias and that criterion scores completely explain overall impact scores. In their study, Black/AA investigators received worse scores on all five review criteria—Significance, Investigator, Innovation, Approach, and Environment than white investigators. An earlier study by Eblen et al. (2016) found a similar result—that criterion scores explained overall impact scores at NIH. Gender and race/ethnicity differences in overall scores were fully explained by criterion scores in Eblen et al. (2016).

Two studies used anonymization to evaluate race/ethnicity bias in NIH funding. Forscher et al. (2019) used 48 NIH proposals and then created white male, white female, Black male and Black female versions of each proposal. They then solicited over 1200 reviews of these proposals from 412 scientists. They found little evidence of gender or race bias in the scores assigned to the proposals. They also used an approach developed by Kaatz et al. (2015) to examine whether written critiques of the proposals differed by race or gender, and found no evidence to support these differences. Overall Forscher et al. (2019) found little evidence of gender or race bias in an experiment designed to replicate the first-stage of NIH reviews. They conclude: “Nevertheless, our evidence does suggest some good news: any name-based race or gender discrimination that is present in the initial review of R01 grant proposals is probably small, below half a point.”

In a second anonymization study, Nakamura et al. (2021) used 1200 NIH grant applications from Black/AA and white investigators and redacted all information about the identity of the investigator. Next, they sent these applications to over 2000 reviewers who produced over 7000 reviews. Their results showed that proposals from white investigators scored better than those from Black/AA investigators. Some of this can be explained by combining Type 1 and Type 2 proposals in their experiment. They note: “Competitive renewals of previously funded projects [Type 2 R01s] and resubmissions of previously reviewed applications scored better. These effects are well-known and tend to favor applications from White PIs (because the representation of White PIs among established investigators is higher)” (p. 19). Next, they compared the scores that redacted proposals received to the unredacted versions. Interestingly, they found that Black/AA redacted and unredacted proposals received the same score, but redacted proposals from white investigators scored worse. They conclude: “The data reveal little evidence of systematic bias based on knowledge of, or perceptions of PI race per se” (p. 19). However, it appears that when the names of white investigators are known to the reviewers, they receive a scoring premium.

Taken together these studies found limited to no evidence of bias in peer review. Erosheva et al. (2020) and Eblen et al. (2016) found that criterion scores explained the overall impact sore, and that there was no evidence of bias race/ethnicity in the overall impact scores. The Forscher et al. (2019) study found no evidence of race or gender bias in the review of NIH proposals where the race and gender were anonymized. Nakamura et al. (2021) used actual proposals from Black/AA and white investigators and found that scores of anonymized Black/AA-led proposals did not change but those of white investigator-led proposals got a bit worse. The Nakamura et al. (2021) results show evidence of white privilege or perhaps a halo effect from affiliation and prior research accomplishments.

Conclusions and Directions for Future Research

This paper analyzed gender and race/ethnicity differences in public research funding. The data and research findings vary dramatically, so I will discuss conclusions about the gender gap in funding, followed by the race/ethnicity gap. The data is more extensive on gender differences in research funding, and it reveals limited evidence of gender bias in the funding process. This study used data on success rates from NIH, NSF and the European Commission, as well as federal grants matched to faculty at research-intensive universities in the United States compiled by Academic Analytics. The analysis shows that faculty at research-intensive universities are the most likely to receive research funding. Comparing Academic Analytics data for these faculty, I showed that female investigators were equally, and in some cases, more likely to have received federal research funding between 2004 and 2018. With the exception of life science and psychology, female investigators received the same number of grants as males. However, female investigators received less total funding in life science, computer science and physical science fields.

I used the count of faculty by discipline at research-intensive universities as a proxy for the applicant pool. Comparing the share of female applicants to female faculty in disciplines served by NSF Directorates, I found that females had higher application rates. Finally, female funding success rates are higher than males at NSF for the past decade. Since the end of the NIH doubling in 2003, funding success rates at NIH have been very similar for female and male investigators. A review of gender differences in funding found a gender gap in the 28 countries of the European Union of about 3%. A review of the literature on gender gaps in funding mostly found no significant gaps. Exceptions were found in Canada, Switzerland, and the Netherlands. In some European reviews, there were significant differences in how proposals were scored (where female investigators were penalized), but upon review in the panel, funding rates for female investigators reached parity with males. This suggests that a deeper investigation of peer review is warranted. Nevertheless, in the majority of contexts evaluated, and at US science funding agencies, there is little evidence of gender bias in research funding.

The situation by race/ethnicity is altogether different. The data comparisons and research literature has been hampered by a lack of access to data on funding success rates by race at the National Institutes of Health. Academic Analytics does not have data by race; thus, it is difficult to observe the race/ethnicity characteristics of faculty at research-intensive institutions. Europe does not collect data by race. Data on all faculty at US colleges and institutions paint a sobering picture. Over 90% of faculty in broad science fields are white or Asian. In some cases, the number of Black/AA or Hispanic faculty in a discipline number in the hundreds. Although NIH does not publish numbers on funding success rates by race/ethnicity the NSF does. When I compared funding rates by race/ethnicity at NIH and NSF, the Black/AA/white funding gap at NIH is the largest, but the Asian/white funding gap at NSF is even larger. Clearly, NIH is not alone in having unexplained funding gaps by race/ethnicity.

Since Ginther et al. (2011) researchers have attempted to explain the Black/AA/white funding gap. Ginther et al. (2018) found that lower publication rates by Black/AA investigators explained 50% of the Black/AA/white funding gap. Lauer et al. (2021) found that Black/AA investigators were more likely to apply at ICs with lower funding rates. Two anonymization studies found no evidence of bias against Black/AA investigators in the review process. If anything, Nakamura et al. (2021) found that white investigators got “extra credit” when their names were revealed to the reviewers. This suggests that a “Matthew effect” (Merton 1967) for white researchers may be contributing to the Black/AA/white funding gap at NIH.

Suggestions for Future Data and Research Questions

Management consultant Peter Drucker reportedly said, “If you can’t measure it, you can’t improve it.” While data on funding rate success by gender is readily available, that is not the case for funding by race/ethnicity. Although the NIH pledged to publish data by race (NIH Advisory Committee 2021), their current effort of reporting the number of principal investigators funded by race falls far short of ideal. NSF has published funding success rates by race for over a decade. It is long past time for NIH to do so.

That said, NIH has made application data available to a select few researchers. NSF has only done so for one study (Broder, 1993). I recommend that NSF and NIH place their administrative data on grant applications and awards in a Federal Research Data Center so that more researchers can examine the factors that explain gender and racial disparities in research funding. The findings in Bol, de Vaan & van de Rij (2022) and Van den Besselaar & Mom (2020) suggest that there is much to be gained by a careful review of how proposal reviews and overall evaluations evolve during the process and arrive at a funding decision.

Institutional data holds promise for studying the impact of gender and race/ethnicity on funding outcomes. Studies such as Waisbren et al. (2008), Warner et al. (2017), Lewit et al.

(2021) and Kaatz (2015, 2016) suggest that information can be gleaned from the review process by looking at applications from investigators at a sufficiently large institution.

This review also points to new research questions. First, do panels "correct" for gender bias in the review process as was suggested by Bol, de Vaan & van de Rij (2022) and Van den Besselaar & Mom (2020). Given the complexity of the review process at NOW and ERC, does this happen in other review contexts? Second, how much of the funding disadvantage for Black/AA researchers is the result of applying to NIH ICs with lower funding rates? The results in Lauer et al. (2021) are suggestive but not comparable to Ginther et al. (2018). It would be beneficial to have the estimation sampled defined the same way to make comparisons across the two studies. Third, why are Asian investigators hugely disadvantaged at NSF? The Asian funding gap at NSF is comparable to the Black/AA/white funding gap at NIH first identified in Ginther et al. (2011). Fourth, why are female application rates so low in the United States relative to Europe? Cross-national comparisons of applications by scientific field may shed light on why in the US female investigators have reached parity in research funding, but in Europe female investigators are less likely to receive funding.

References

Academic Analytics, LLC (2020). *Academic Analytics Research Center.* https://academicanalytics.com/aarc-scholarly-research/

Albers, C. J. (2015). Dutch research funding, gender bias, and Simpson's paradox. *Proceedings of the National Academy of Sciences of the United States of America, 112*, E6828–E6829. https://doi.org/10.1073/pnas.1518936112

Association for Women in Science. (2021). *AWIS History*. https://www.awis.org/about-awis/awis-history/

Baldwin, M. (2018). Scientific Autonomy, Public Accountability, and the Rise of "Peer Review" in the Cold War United States. *Isis*, *109*(3), 538–558. https://doi.org/10.1086/700070

Beck, R. & Halloin, V. (2017) Gender and research funding success: Case of the Belgian F.R.S.-FNRS. *Research Evaluation, 26*(2), 115–123, https://doi.org/10.1093/reseval/rvx008

Bol, T., de Vaan, M., & van de Rijt, A. (2022). Gender-equal funding rates conceal unequal evaluations. *Research Policy, 51*(1), 104399. https://doi.org/10.1016/j.respol.2021.104399

Boyle, P., Smith, L., Cooper, N., Williams, K., & O'Connor, H. (2015). Women are funded more fairly in social science. *Nature*, *525*, 181-183. https://doi.org/10.1038/525181a

Broder, I. E. (1993). Review of NSF Economics Proposals: Gender and Institutional Patterns. *The American Economic Review*, *83*(4), 964–970.

Burns, K.E.A., Straus, S.E., Liu, K., Rizvi, L. & Guyatt, G. (2019). Gender differences in grant and personnel award funding rates at the Canadian Institutes of Health Research based on research content area: A retrospective analysis. *PLOS Medicine, 16*(10), e1002935. https://doi.org/10.1371/journal.pmed.1002935

Ceci, S. J., Ginther, D. K., Kahn, S., & Williams, W. M. (2014). Women in Academic Science: A Changing Landscape. *Psychological Science in the Public Interest, 15*(3), 75-141. https://doi.org/10.1177/1529100614541236

Cheng, M.Y., Sukhov, A., Sultani. H., Kim, K., & Maverakis, E. (2016). Trends in National Institutes of Health Funding of Principal Investigators in Dermatology Research by Academic Degree and Sex. *JAMA Dermatol, 152*(8), 883-888. https://doi.org/10.1001/jamadermatol.2016.0271

Couzin, J., & Miller, G. (2007). Boom and Bust. *Science, 316*(5823), 356-361. https://doi.org/10.1126/science.316.5823.356

Cruz-Castro, L., & Sanz-Menéndez, L. (2020). Grant Allocation Disparities from a Gender Perspective: Literature Review. Synthesis Report. *GRANteD Project D.1.1.* https://doi.org/10.20350/digitalCSIC/10548

Cruz-Castro, L., Ginther, D.K., & Sanz-Menéndez, L. (forthcoming). Gender and Underrepresented minorities differences in research funding. In B. Lepori, B. Jongbloed, & D. Hicks. (Eds.), *Handbook of public funding of research*. Edward Elgar Publishing.

Eblen, M.K., Wagner, R.M., RoyChowdhury, D., Patel, K.C., & Pearson, K. (2016). How Criterion Scores Predict the Overall Impact Score and Funding Outcomes for National Institutes of Health Peer-Reviewed Applications. *PLOS ONE, 11*(6), e0155060. https://doi.org/10.1371/journal.pone.0155060

Erosheva, E. A., Grant, S., Chen, M.-C., Lindner, M. D., Nakamura, R. K., & Lee, C. J. (2020). NIH peer review: Criterion scores completely account for racial disparities in overall impact scores. *Science Advances*, *6*(23), eaaz4868. https://doi.org/10.1126/sciadv.aaz4868

European Commission, Directorate-General for Research and Innovation, (2021). She figures 2021. Gender in research and innovation: statistics and indicators. *Publications Office.* https://data.europa.eu/doi/10.2777/06090

Fang, F.C., & Casadevall, A. (2016). Research Funding: The Case for a Modified Lottery. *mBio. 7*(2), e00422-16. https://doi.org/10.1128/mBio.00422-16

Fang, F.C., Bowen, A., & Casadevall, A. (2016). NIH Peer Review Percentile Scores Are Poorly Predictive of Grant Productivity. *eLife 5*, e13323. https://doi.org/10.7554/eLife.13323

Forscher, P. S., Cox, W. T. L., Brauer, M., & Devine, P. G. (2019). Little race or gender bias in an experiment of initial review of NIH R01 grant proposals. *Nature Human Behaviour*, *3*(3), 257–264. https://doi.org/10.1038/s41562-018-0517-y

Gallo, S.A., Sullivan, J.H., & Glisson, S.R. (2016). The Influence of Peer Reviewer Expertise on the Evaluation of Research Funding Applications. *PLOS ONE, 11*(10), e0165147. https://doi.org/10.1371/journal.pone.0165147

Ginther, D. K., Schaffer, W. T., Schnell, J., Masimore, B., Liu, F., Haak, L. L., & Kington, R. (2011). Race, Ethnicity, and NIH Research Awards. *Science*, *333*(6045), 1015–1019. https://doi.org/10.1126/science.1196783

Ginther, D.K., & Heggeness, M.L. (2020). Administrative discretion in scientific funding: Evidence from a prestigious postdoctoral training program. *Research Policy*, *49*(4), 103953. https://doi.org/10.1016/j.respol.2020.103953

Ginther, Donna K., Basner, J., Jensen, U., Schnell, J., Kington, R., & Schaffer, W. T. (2018). Publications as predictors of racial and ethnic differences in NIH research awards. *PLOS ONE*, *13*(11), e0205929. https://doi.org/10.1371/journal.pone.0205929

Ginther, Donna K., Haak, L. L., Schaffer, W. T., & Kington, R. (2012). Are Race, Ethnicity, and Medical School Affiliation Associated with NIH R01 Type 1 Award Probability for Physician Investigators? *Academic Medicine*, *87*(11), 1516–1524. https://doi.org/10.1097/ACM.0b013e31826d726b

Ginther, D.K., Kahn, S., & Schaffer, W.T. (2016). Gender, Race/Ethnicity, and National Institutes of Health R01 Research Awards: Is There Evidence of a Double Bind for Women of Color? *Academic Medicine, 91*(8), 1098-107. https://doi.org/10.1097/ACM.0000000000001278.

Head, M.G., Fitchett, J.R., Cooke, M.K., Wurie, F.B., & Atun, R. (2013). Differences in research funding for women scientists: a systematic comparison of UK investments in global infectious disease research during 1997-2010. *BMJ Open, 3*(12), e003362 https://doi.org/10.1136/bmjopen-2013-003362.

Hoppe, T. A., Litovitz, A., Willis, K. A., Meseroll, R. A., Perkins, M. J., Hutchins, B. I., … Santangelo, G. M. (2019). Topic choice contributes to the lower rate of NIH awards to African-American/black scientists. *Science Advances*, *5*(10), eaaw7238. https://doi.org/10.1126/sciadv.aaw7238

Jacob, B.A., & Lefgren, L. (2011). The Impact of Research Grant Funding on Scientific Productivity. *Journal of Public Economics*, *95*(9-10), 1168–1177. https://doi.org/10.1016/j.jpubeco.2011.05.005

Jagsi, R., Motomura, A.R., Griffith, K.A., Rangarajan, S., & Ubel, P.A. (2009). Sex differences in attainment of independent funding by career development awardees. *Annals of Internal Medicine, 151*(11), 804-811. https://doi.org/10.7326/0003-4819-151-11-200912010-00009

Kaatz, A., Magua, W., Zimmerman, D.R., & Carnes, M. (2015). A quantitative linguistic analysis of National Institutes of Health R01 application critiques from investigators at one institution. *Academic Medicine, 90*(1), 69-75. https://doi.org/10.1097/ACM.0000000000000442.

Kaatz, A., Lee, Y.G., Potvien, A., Magua, W., Filut, A., Bhattacharya, A., … Carnes, M. (2016). Analysis of National Institutes of Health R01 Application Critiques, Impact, and Criteria Scores: Does the Sex of the Principal Investigator Make a Difference? *Academic Medicine, 91*(8), 1080-8. https://doi.org/10.1097/ACM.0000000000001272.

Kahn, S., & Ginther, D. (2018). Women and Science, Technology, Engineering, and Mathematics (STEM): Are Differences in Education and Careers Due to Stereotypes, Interests, or Family? In S. L. Averett & L. M. Argys (Eds.), *The Oxford Handbook of Women and the Economy*. Oxford University Press. https://doi.org/10.1093/oxfordhb/9780190628963.001.0001

Kalyani, R.R., Yeh, H.C., Clark, J.M., Weisfeldt, M.L., Choi, T., & MacDonald, S.M. (2015). Sex Differences Among Career Development Awardees in the Attainment of Independent Research Funding in a Department of Medicine. *Journal of Women’s Health, 24*(11), 933-939. https://doi.org/10.1089/jwh.2015.5331

Lauer, M. S., Doyle, J., Wang, J., & Roychowdhury, D. (2021). Associations of topic-specific peer review outcomes and institute and center award rates with funding disparities at the National Institutes of Health. *eLife*, *10*, e67173. https://doi.org/10.7554/eLife.67173

Lauer, M. (2022, March 7). *FY 2021 By the numbers: Extramural grant investments in research*. National Institutes of Health Office of Extramural Research. https://nexus.od.nih.gov/all/2022/03/07/fy-2021-by-the-numbers-extramural-grant-investments-in-research/

Lewit, R.A., Black, C.M., Camp, L., Goldstein, A.M., Matthews, J.B., Emamaullee, J., & Gosain, A. (2021). Association of Sex And Race/Ethnicity with National Institutes of Health Funding of Surgeon-Scientists. *JAMA Surgery, 156*(2), 195-197. https://doi.org/10.1001/jamasurg.2020.5016

Ley, T. J., & Hamilton, B. H. (2008). The Gender Gap in NIH Grant Applications. *Science*, *322*(5907), 1472–1474. https://doi.org/10.1126/science.1165878

Li, D., & Agha, L. (2015). Big Names or Big Ideas: Do Peer-Review Panels Select the Best Science Proposals? *Science, 348*(6233), 434–438. https://doi.org/10.1126/science.aaa0185.

Li, D. (2017). Expertise versus Bias in Evaluation: Evidence from the NIH. *American Economic Journal: Applied Economics, 9*(2), 60-92. https://doi.org/10.1257/app.20150421.

Magua, W., Zhu, X., Bhattacharya, A., Filut, A., Potvien, A., Leatherberry, R., & Kaatz, A. (2017). Are Female Applicants Disadvantaged in National Institutes of Health Peer Review? Combining Algorithmic Text Mining and Qualitative Methods to Detect Evaluative Differences in R01 Reviewers' Critiques. *Journal of Women's Health, 26*(5), 560-570. https://doi.org/10.1089/jwh.2016.6021

Merton, R. K. (1968). The Matthew Effect in Science: The reward and communication systems of science are considered. *Science*, *159*(3810), 56–63. https://doi.org/10.1126/science.159.3810.56

Mulvey, M.R., West R.M., Cotterill L.A., Magee, C., Jones D.E.J., Harris-Joseph, H., … Hewison J. (2022). Ten years of NIHR research training: who got an award? A retrospective cohort study. *BMJ Open, 12*(1), e046368. https://doi.org/10.1136/bmjopen-2020-046368

Mutz, R., Bornmann, L., & Daniel, H.-D. (2012). Does gender matter in grant peer review? An empirical investigation using the example of the Austrian science fund. *Zeitschrift für Psychologie/Journal of Psychology*, *220*(2), 121–129. http://dx.doi.org.ezproxy1.lib.asu.edu/10.1027/2151-2604/a000103

Nakamura, R., Mann, L.S., Lindner, M.D., Braithwaite, J., Chen, M.-C., Vancea, A., … Reed, B. (2021). An Experimental Test of the Effects of Redacting Grant Applicant Identifiers on Peer Review Outcomes. *eLife*, *10*, e671368. https://doi.org/10.7554/eLife.71368

National Science Foundation. *FY 2021 Agency Financial Report (AFR) – List of Figures and Tables.* https://www.nsf.gov/pubs/2022/nsf22002/tables.jsp

NCSES 2019 Survey of Doctorate Recipients. OMB No: 3145-0020. National Center for Science and Engineering Statistics (NCSES) within the National Science Foundation (NSF) and the National Institutes of Health (NIH). https://www.nsf.gov/statistics/srvydoctoratework/surveys/srvydoctoratework-2019.pdf

NIH Advisory Committee. (2021*). The Advisory Committee to the Director Working Group on Diversity, Racism in Science Report.* https://www.acd.od.nih.gov/documents/presentations/02262021DiversityReport.pdf

NIH RePORTER. 2022. Initial public release of Modernized NIH RePORTER, version 2020.9. reporter.nih.gov.

National Center for Science and Engineering Statistics. (2021). *Women, Minorities, and Persons with Disabilities in Science and Engineering: 2021. Special Report NSF 21-321.* https://ncses.nsf.gov/wmpd.

National Science Foundation. (2019). *Merit Review Process Fiscal Year 2019 Digest.* https://www.nsf.gov/nsb/publications/2020/merit_review/FY-2019/nsb202038.pdf

National Science Foundation. (2009). *Merit Review Process Fiscal Year 2009 Digest.* https://www.nsf.gov/nsb/publications/2010/nsb1027.pdf

Oliveira, D. F. M., Ma, Y., Woodruff, T. K., & Uzzi, B. (2019). Comparison of National Institutes of Health Grant Amounts to First-Time Male and Female Principal Investigators. *JAMA*, *321*(9), 898–900. https://doi.org/10.1001/jama.2018.21944

Pier, E. L., Brauer, M., Filut, A., Kaatz, A., Raclaw, J., Nathan, M. J., Ford, C. E., & Carnes, M. (2018). Low agreement among reviewers evaluating the same NIH grant applications. *Proceedings of the National Academy of Sciences*, *115*(12), 2952–2957. https://doi.org/10.1073/pnas.1714379115

Pohlhaus, J. R., Jiang, H., Wagner, R. M., Schaffer, W. T., & Pinn, V. W. (2011). Sex Differences in Application, Success, and Funding Rates for NIH Extramural Programs. *Academic Medicine*, *86*(6), 759–767. https://doi.org/10.1097/ACM.0b013e31821836ff

Reinhart, M. (2009). Peer review of grant applications in biology and medicine. Reliability, fairness, and validity. *Scientometrics. 81*, 789-809. https://doi.org/10.1007/s11192-008-2220-7.

Sandström, U., & Hällsten, M. (2008). Persistent nepotism in peer-review. *Scientometrics, 74*, 175–189. https://doi.org/10.1007/s11192-008-0211-3

Severin, A., Martins, J., Heyard, R., Delavy, F., Jorstad, A., & Egger, M. (2020). Gender and other potential biases in peer review: Cross-sectional analysis of 38 250 external peer review reports. *BMJ Open*, *10*(8), e035058. https://doi.org/10.1136/bmjopen-2019-035058

Tamblyn, R., Girard, N., Qian, C. J., & Hanley, J. (2018). Assessment of potential bias in research grant peer review in Canada. *Canadian Medical Association Journal*, *190*(16), E489–E499. https://doi.org/10.1503/cmaj.170901

van den Besselaar, P., & Mom, C. (2020). Gender differences in research grant allocation - a mixed picture. *Preprint.*

van der Lee, R., & Ellemers, N. (2015). Gender contributes to personal research funding success in The Netherlands. *Proceedings of the National Academy of Sciences*, *112*(40), 12349–12353. https://doi.org/10.1073/pnas.1510159112

Volker, B., & Steenbeek, W. (2015). No evidence that gender contributes to personal research funding success in The Netherlands: A reaction to van der Lee and Ellemers. *Proceedings of the National Academy of Sciences*, *112*(51), E7036–E7037. https://doi.org/10.1073/pnas.1519046112

Waisbren, S.E., Bowles, H., Hasan, T., Zou, K.H., Emans, S.J., Goldberg, C., … Christou, H. (2008). Gender differences in research grant applications and funding outcomes for medical school faculty. *Journal of Women's Health, 17*(2), 207-214. https://doi.org/10.1089/jwh.2007.0412

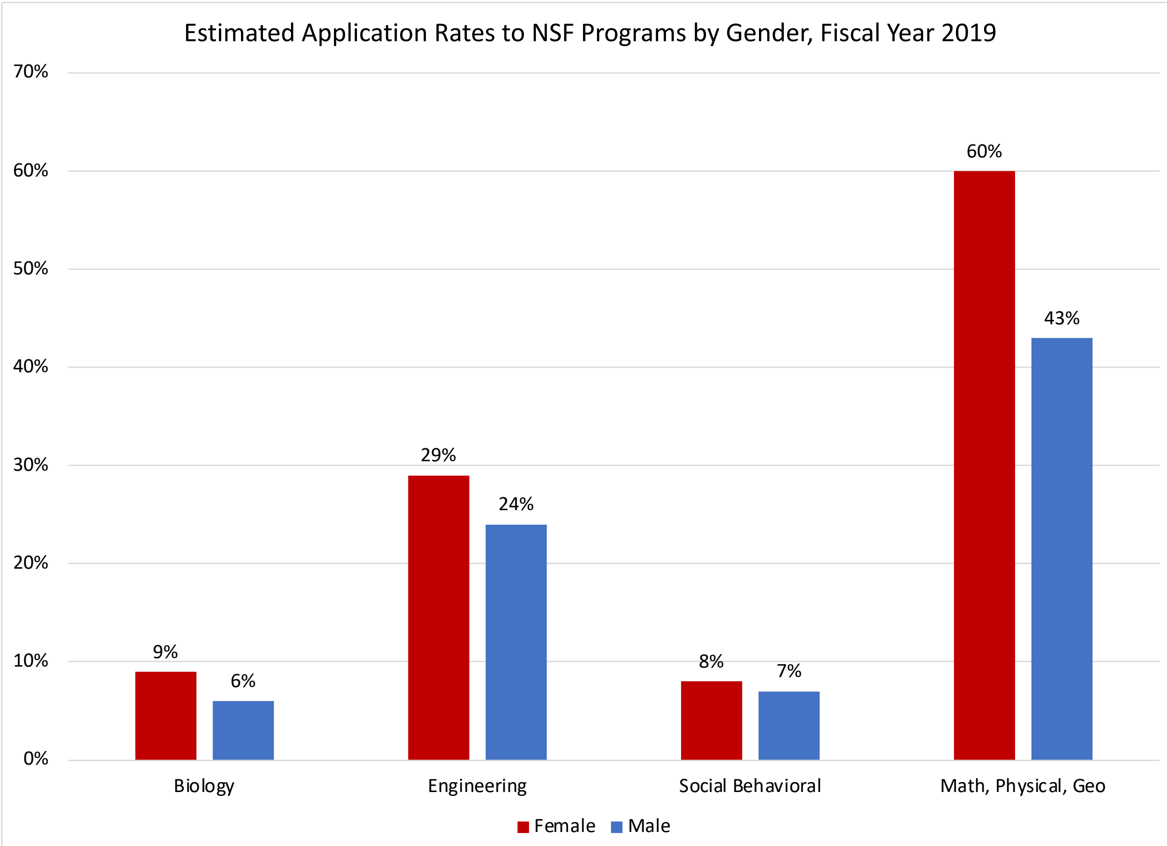
Warner, E.T., Carapinha, R., Weber, G.M., Hill, E.V., Reede, J.Y. (2017). Gender Differences in Receipt of National Institutes of Health R01 Grants Among Junior Faculty at an Academic Medical Center: The Role of Connectivity, Rank, and Research Productivity. *Journal of Women's Health, 26*(10), 1086-1093. https://doi.org/10.1089/jwh.2016.6102

Wennerås, C., & Wold, A. (1997). Nepotism and sexism in peer-review. *Nature*, *387*, 341–343. https://doi.org/10.1038/387341a0

Wilbert, T., Tedjasaputra, V. (2021, May 3). *NSF 101: Five tips for your Broader Impacts statement*. National Science Foundation Science Matters. https://beta.nsf.gov/science-matters/nsf-101-five-tips-your-broader-impacts-statement

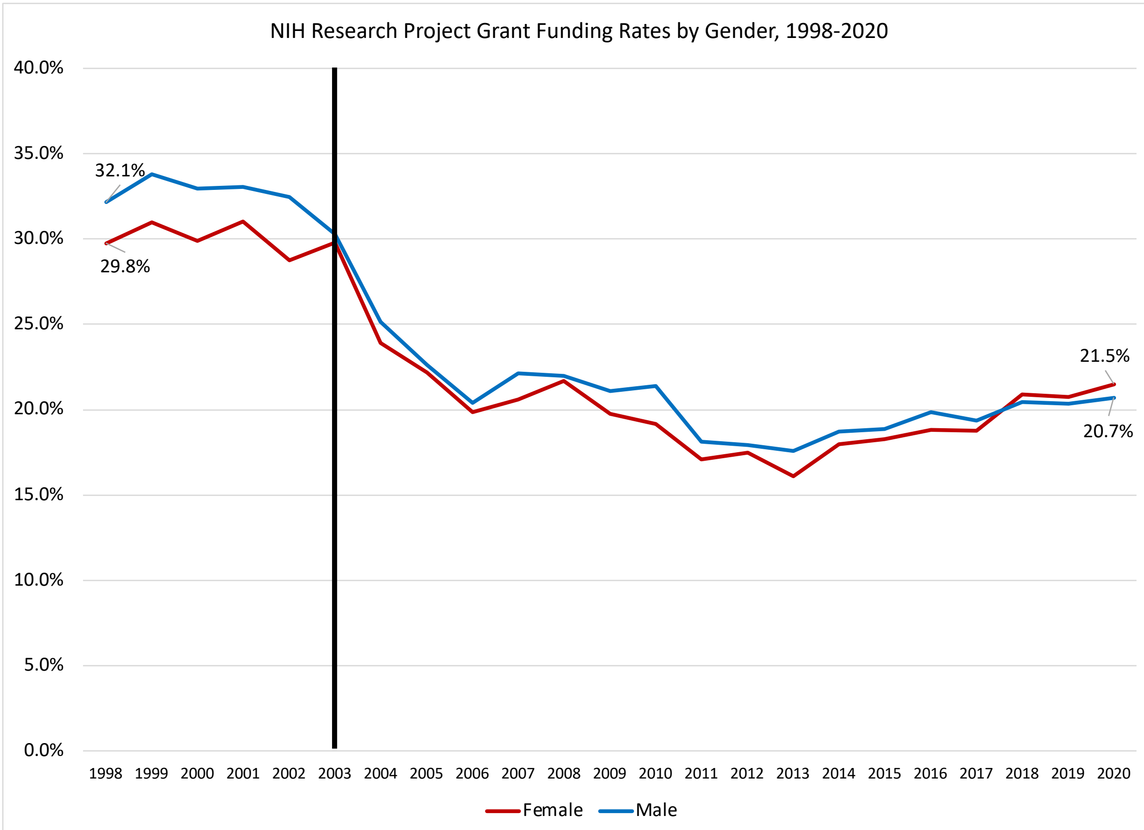
Witteman, H. O., Hendricks, M., Straus, S., & Tannenbaum, C. (2019). Are gender gaps due to evaluations of the applicant or the science? A natural experiment at a national funding agency. *The Lancet*, *393*(10171), 531–540. https://doi.org/10.1016/S0140-6736(18)32611-4

Figure 1



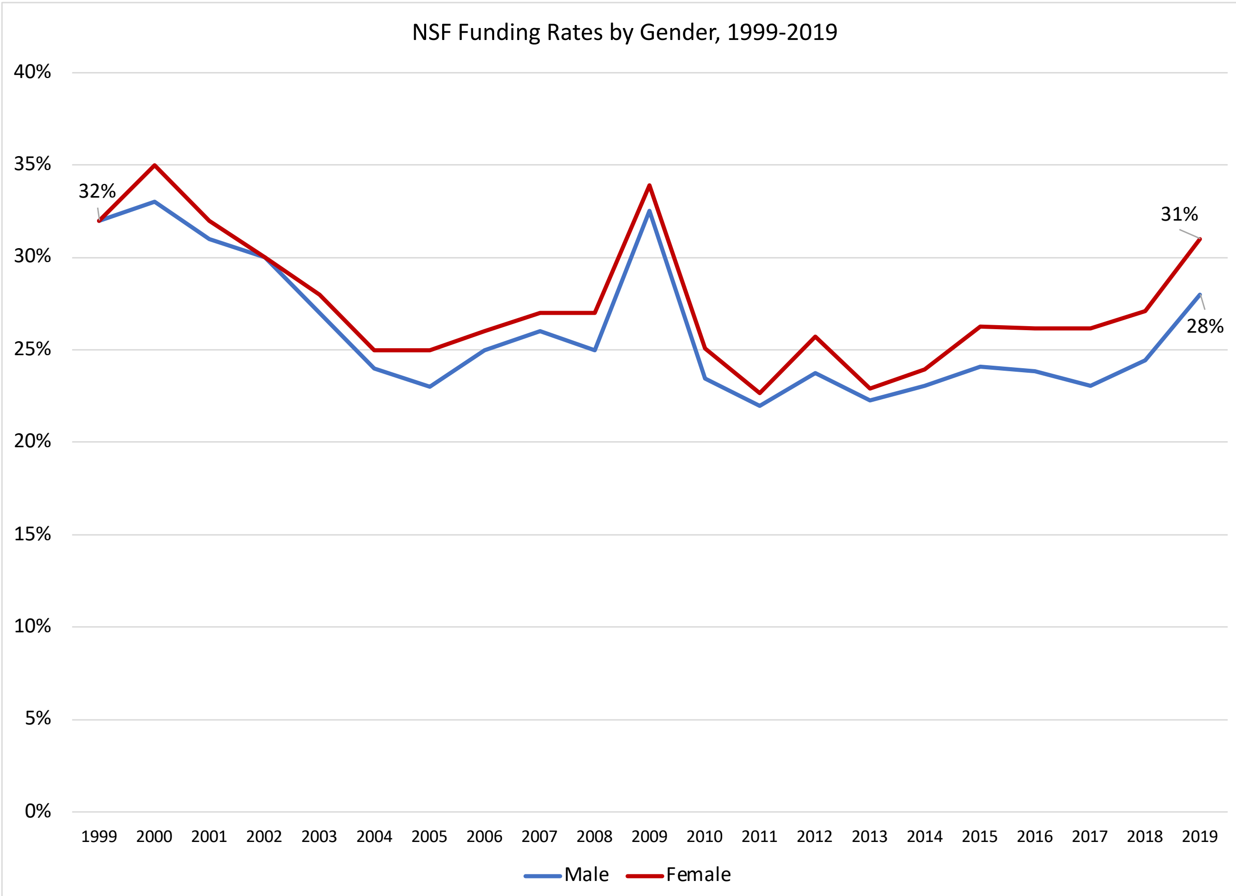
Source: National Science Foundation Merit Review Process Fiscal Year 2019 Digest and Academic Analytics 2018.

Figure 2



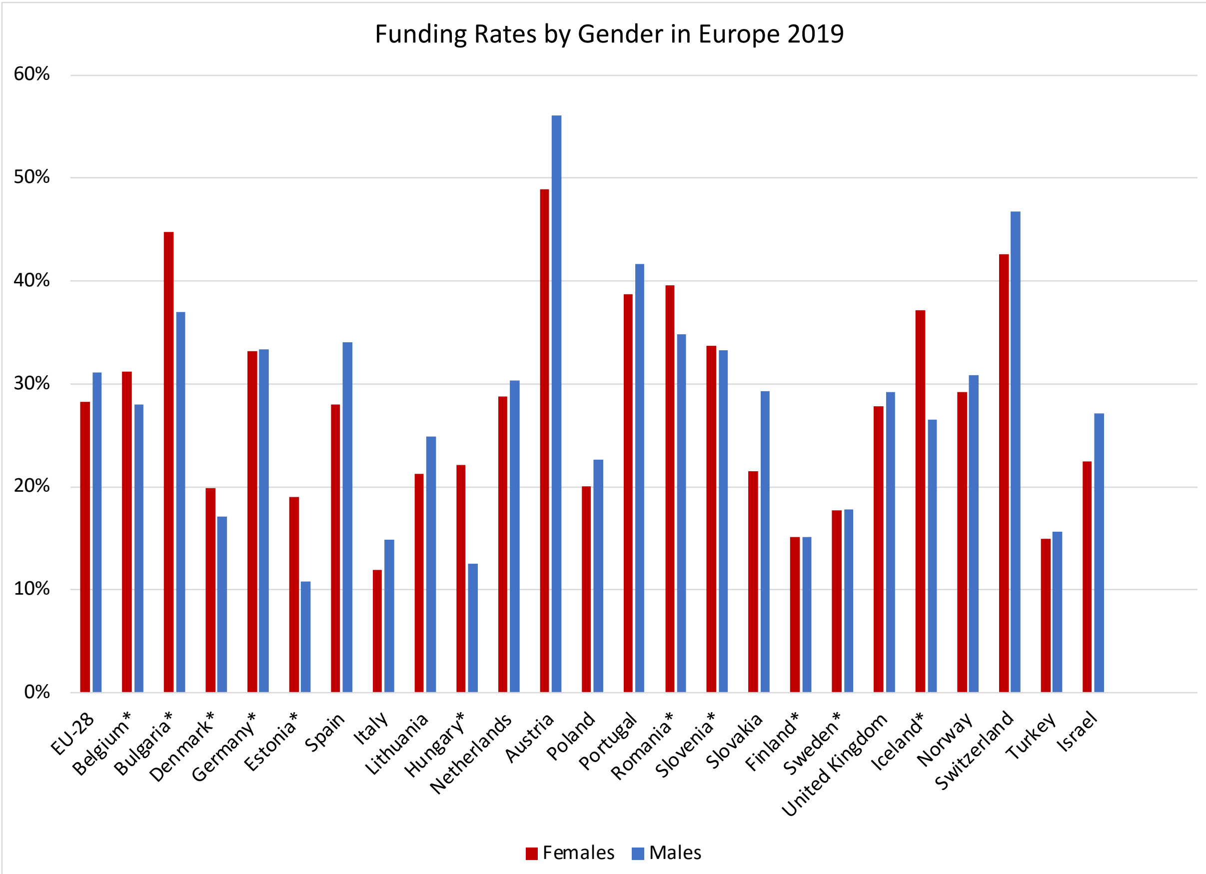
Source: NIH REPORTER.

Figure 3



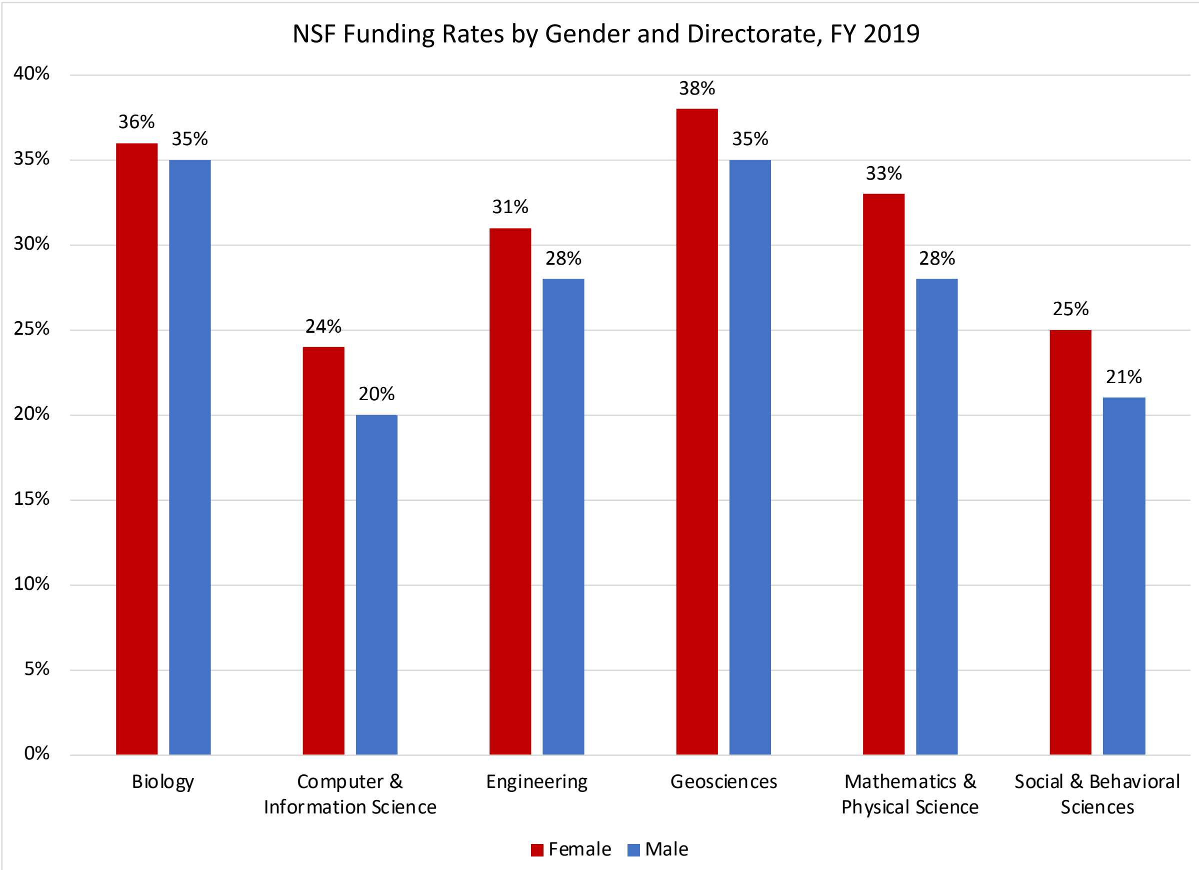
Source: National Science Foundation Merit Review Process Fiscal Year 2019 and 2009 Digests.

Figure 4.



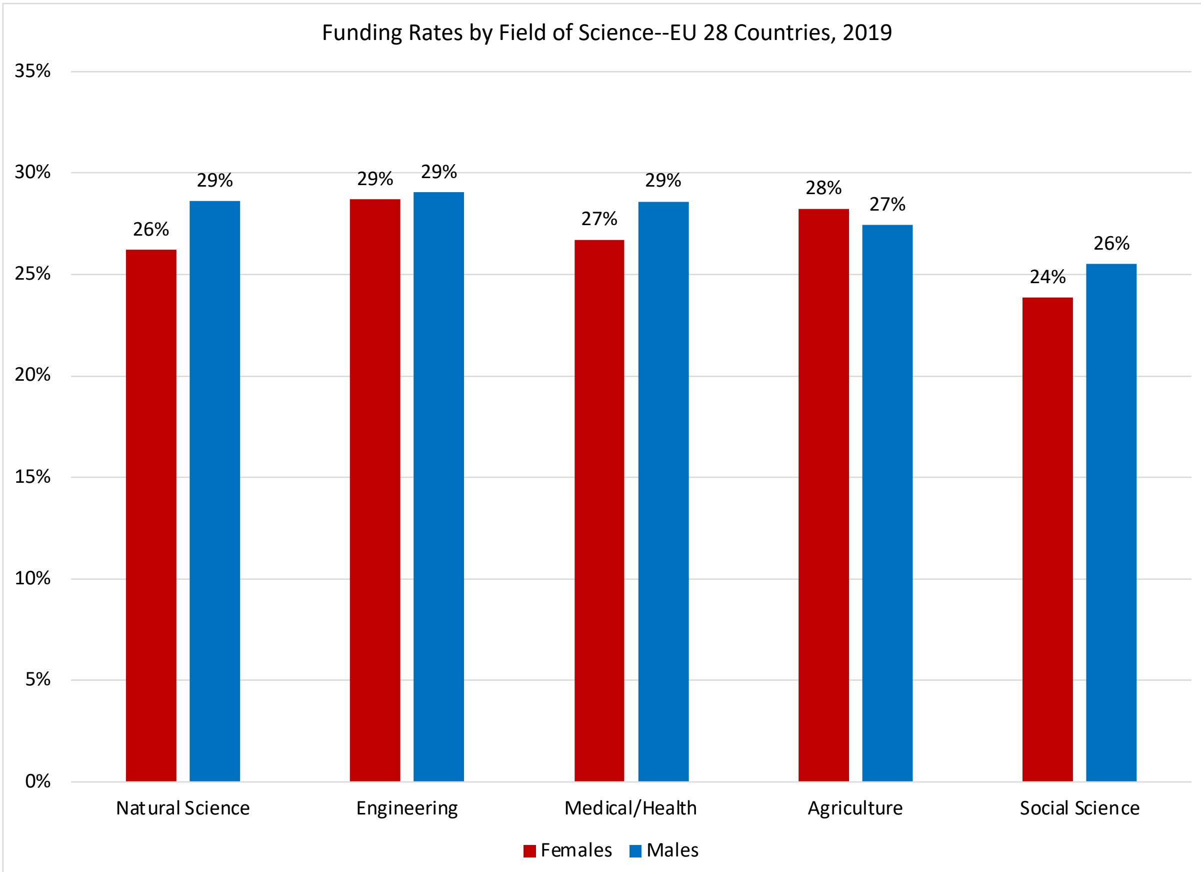
Source: European Commission *She Figures* 2021.

Figure 5



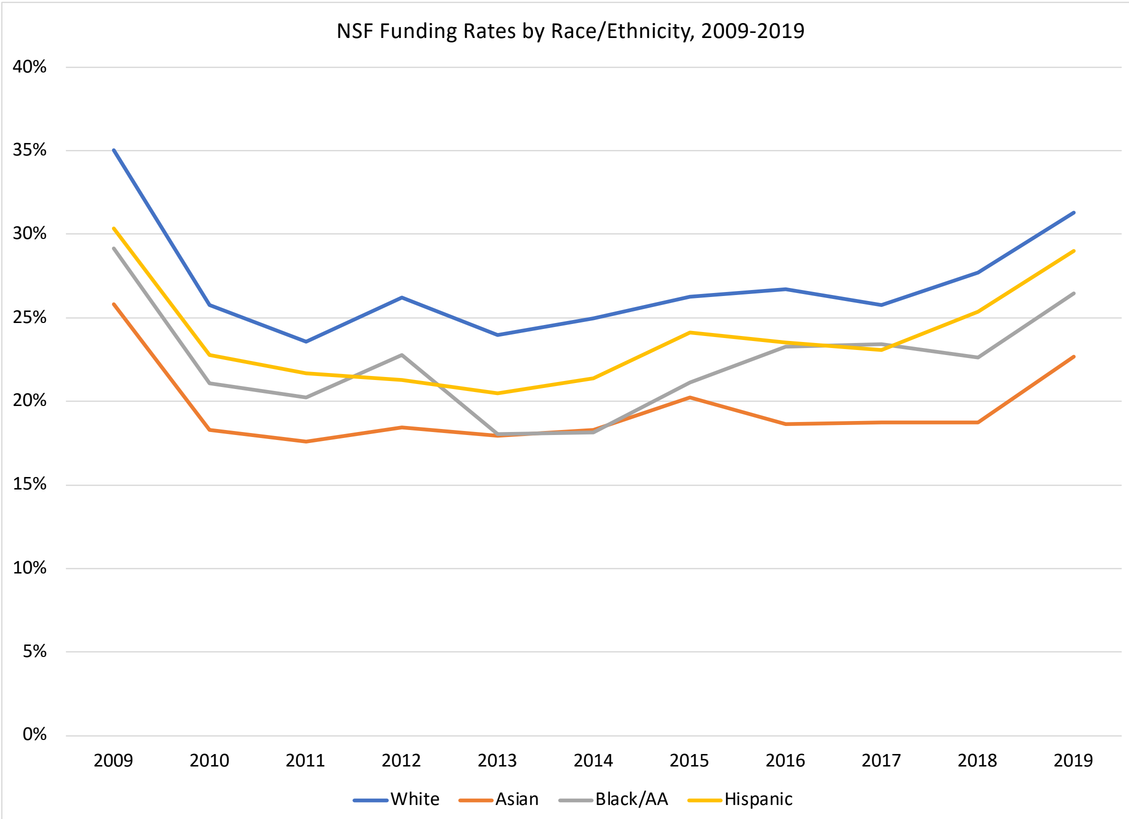
Source: National Science Foundation Merit Review Process Fiscal Year 2019 Digest.

Figure 6



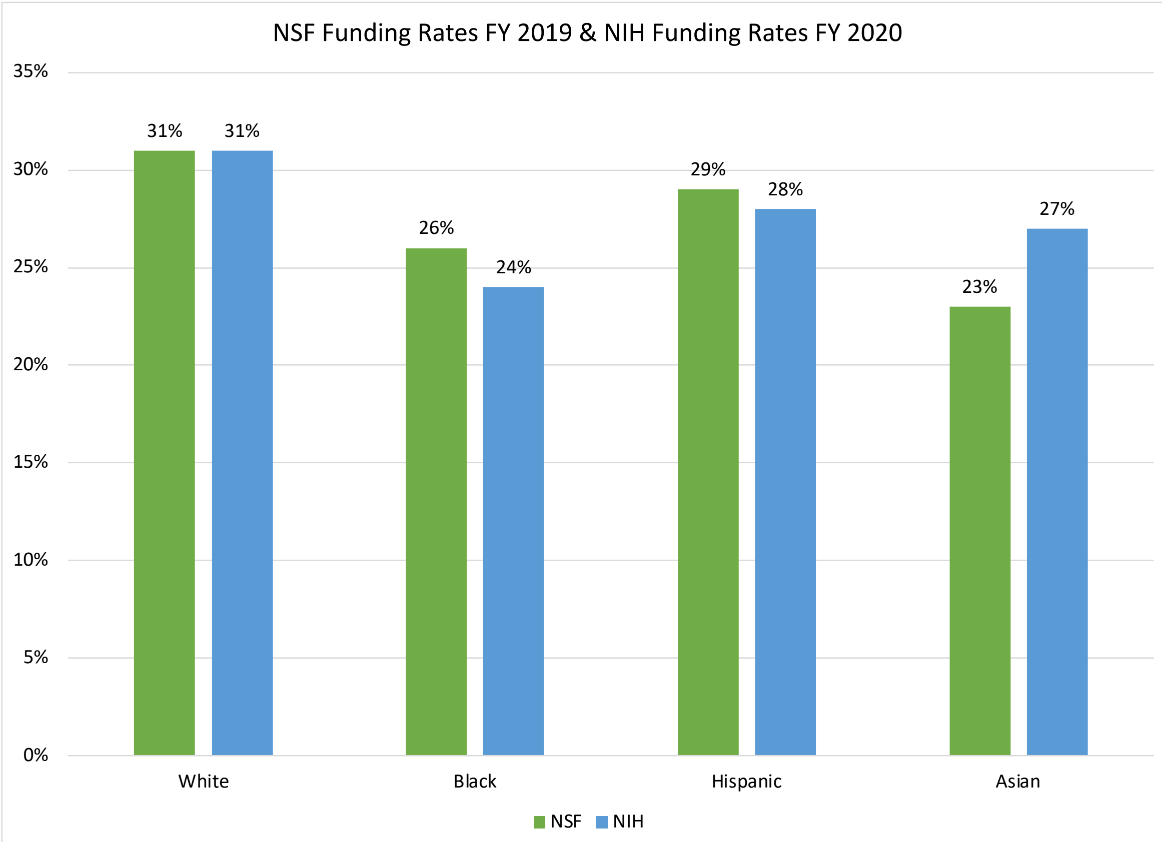
Source: European Commission *She Figures* 2021.

Figure 7



Source: National Science Foundation Merit Review Process Fiscal Year 2019 Digest.

Figure 8



Source: National Science Foundation Merit Review Process Fiscal Year 2019 Digest and The Advisory Committee to the Director Working Group on Diversity, Racism in Science Report, 2021.

Table 1: Assistant, Associate and Full Professors at 4-Year Colleges and Universities by Gender, Broad Science Field and Research Funding, 2018.

	Total	Female share of total	Research-Intensive Universities	Female Share Research Intensive	Female Funded	Research-Intensive Female Funded	Male Funded	Research-Intensive Male Funded
Science & Engineering	201,700	33%	95,890	28%	39%	48%	44%	51%
Science	172,100	36%	77,263	30%	38%	46%	41%	50%
Biological/ agricultural/ other life scientist	48,450	36%	28,345	30%	57%	61%	62%	61%
Computer/ information scientist	10,550	18%	4,630	21%	49%	56%	39%	58%
Mathematical scientist	17,900	26%	6,595	19%	36%	52%	30%	49%
Physical scientist	31,050	25%	12,249	19%	47%	75%	51%	67%
Psychologist	20,600	57%	7,474	48%	29%	32%	34%	35%
Social scientist	43,500	41%	17,970	37%	19%	23%	16%	21%
Engineering	29,550	16%	18,627	19%	55%	56%	55%	52%

Notes: Data from NCSES Report, *Women, Minorities and Persons with Disabilities in Science and Engineering*, 2021.
Data on total from the 2019 Survey of Doctorate Recipients.
Data on Research-Intensive Universities from 2018 Academic Analytics. Total faculty are those working at 4-year colleges or universities with the academic rank of assistant, associate or full professor.

Table 2: Assistant, Associate and Full Professors at Research-Intensive 4-Year Colleges and Universities by Gender, Broad Science Field and Research Funding, 2018.

	Ever Received Federal Funding, 2004-2018			Total Grants Awarded, 2004-2018			Cumulative Dollars (Millions), 2004-2018		
	Male	Female	Diff	Male	Female	Difference	Male	Female	Difference
Life Science	61%	61%	1%	2.44	2.11	**-0.33****	2.782	2.033	**-0.75****
Computer Science	58%	56%	-1%	2.38	2.24	-0.14	1.304	1.055	**-0.25***
Mathematics/Statistics	49%	52%	3%	1.52	1.39	-0.13	0.504	0.523	0.02
Physical Science	67%	75%	**8%****	3.01	3.02	0.01	2.184	1.816	**-0.37***
Psychology	35%	32%	**-2%***	0.89	0.76	**-0.13***	0.898	0.784	-0.11
Social Science	21%	23%	**2%***	0.51	0.48	-0.03	0.222	0.198	-0.02
Engineering	52%	56%	**4%****	2.17	2.18	0.01	1.322	1.235	-0.09

Notes: Data from 2018 Academic Analytics. Total faculty are those working at research-intensive 4-year colleges or universities with the academic rank of assistant, associate or full professor. T-test of significant difference by gender: ** p<.001, * p<.05.

Table 3: Assistant, Associate and Full Professors at Research-Intensive 4-Year Colleges and Universities by Race/Ethnicity, Broad Science Field and Research Funding, 2018.

		Race and ethnicity											
	Total	White Number	Share White	Funded White	Asian Number	Share Asian	Funded Asian	Black or African American Number	Share Black or African American	Funded Black or African American	Hispanic or Latino[a] Number	Share Hispanic or Latino[a]	Funded Hispanic or Latino[a]
Science & Engineering	201,700	143,200	71%	42%	38,850	19%	48%	7,250	4%	26%	9,600	5%	38%
Science	172,100	126,600	74%	40%	28,250	16%	46%	6,400	4%	24%	8,250	5%	36%
Biological/ agricultural/ other life scientist	48,450	35,800	74%	59%	8,400	17%	70%	1,200	2%	45%	2,350	5%	53%
Computer/ information scientist	10,550	6,150	58%	41%	3,800	36%	42%	140	1%	33%	60	1%	45%
Mathematical scientist	17,900	11,800	66%	32%	4,750	27%	32%	240	1%	33%	670	4%	30%
Physical scientist	31,050	23,800	77%	49%	5,000	16%	57%	620	2%	37%	1,200	4%	50%
Psychologist	20,600	16,600	81%	33%	1,600	8%	29%	1,050	5%	17%	1,100	5%	29%
Social scientist	43,500	32,650	75%	19%	4,800	11%	12%	2,600	6%	9%	2,650	6%	17%
Engineering	29,550	16,550	56%	55%	10,600	36%	55%	900	3%	38%	1,300	4%	52%

Notes: Data from NCSES Report, *Women, Minorities and Persons with Disabilities in Science and Engineering*, 2021. Data source 2019 Survey of Doctorate Recipients. [a] Hispanic or Latino may be any race; race categories exclude Hispanic origin.

Table 4: Studies of Gender Differences in Research Funding 2005-2022

Citation	Country/Agency	Study Population & Type	Funding Gap
Raphaël Beck, Véronique Halloin (2017)	Belgium, Fonds de la Recherche Scientifique - FNRS	N=6,393 Applications	No significant gender gap. Men are more likely to apply and receive higher evaluation for research credit.
Boyle, Smith, Cooper, Williams & O'Connor (2015)	UK, ESRC	Open call social science resarch grants. 2008-2013, N=2085 Applications	No significant gender gap success. Women less likely to apply. No difference in funding amounts.
Burns, Straus, Liu, Rizvi, Guyatt (2019)	Canada, Canadian Institutes of Health Research	Applications to Canadian Institutes of Health Research, N=55,700 Applications	Women 3.2% less likely to receive research funding; Women 6.6% less likely to receive personnel funding.
Ginther, Kahn & Schaffer (2016)	USA, NIH	NIH Type 1 R01, 2000-2006, N=97877 Applications.	No significant gender gap
Jagsi, Motomura, Griffith, Rangarajan, Ubel (2009)	US, NIH K Award Winners	2,784 K awardees 1997-2003. Award data.	Female K awardees were 9 percentage points less likely to reach research independence (R01 award) than males.
Kaatz, Magua, Zimmerman, Carnes (2015)	NIH reviews from University of Wisconsin-Madison	454 reviews from 67 R01 investigators at the University of Wisconsin-Madison unfunded applications subsequently funded between December 2007 and May 2009.	Reviews showed differences due to applicant sex despite similar application scores or funding outcomes: more praise for applications from female investigators and more negative evaluation words for applications from male investigators
Kaatz, Lee, Potvien, Magua, Filut, Bhattacharya, Leatherberry, Zhu (2016)	NIH reviews from University of Wisconsin-Madison	739 reviews for grants awarded to 125 PIs at the University of Wisconsin-Madison between 2010 and 2014.	Female investigators had worse scores on Type 2 applications despite receiving words of praise.
Ley & Hamilton (2008)	USA, NIH	LRP, K08, K01, K23, R0 1-1, R0 1-E grant success rates; Award Data	Women disadvantaged as careers progresses in applications. No significant gender gap for funding success rate, women less likely to apply for R01s. Women disadvantaged as experienced investigators for R01s.
Magua, Zhu, Bhattacharya, Filut, Potvien, Leatherberry, Lee, Jens, Malikireddy,Carnes & Kaatz (2017)	USA, University of Wisconsin-Madison NIH proposals.	241 critiques from 79 Summary Statements for 51 R01 renewals from University of Wisconsin-Madison	Male PI's characterized as "highly innovative" while female PI's characterized as having "expertise". Male investigators received better scores and this was associated with higher productivity
Kalyani, Yeh, Clark, Weisfeldt, Choi, MacDonald. (2015)	US, NIH K Award Winners	N=92 K Award winners at John Hopkins Department of Medicine 1999-2008	Female K awardees were more likely to reach research independence (R01 award) than males.
Mulvey, West, Cotterill Magee, Jones, Harris-Joseph, Thompson, Hewison (2022)	UK, National Institute for Health Research (NIHR) Training grants	UK, NIHR Training grants 2007-2016, N=4,388 Applications	No significant gender gap

Citation	Country/Agency	Study Population & Type	Funding Gap
Mutz, Bornmann, & Daniel (2012)	Austrian Science Fund	N=8,496 research proposals from 1999-2009 with 23,977 reviews.	No significant gender gap in funding. Scores are lower when majority of reviewers are women.
Oliveira, Ma, Woodruff & Uzzi (2019)	USA, NIH	53,903 NIH awards to First-time recipients. Awards data.	Women disadvantaged by $39 106 in all types, by $10 527 in 10 highest-funded types, by $81 711 in Big Ten universities
Pohlhaus, Jiang, Wagner, Schaffer & Pinn (2011)	USA, NIH	NIH Type 1 R01, 2000-2006, N=97877 Applications.	No significant gender gap except for Type 2 awards.
Reinhart M. (2009)	Switzerland, Swiss NSF	496 applications from 1998	No significant gender gap
Sandström, U., Hällsten, M. (2008)	Sweden	Replication of Wennerås & Wold (1997)	Women receive 10-15% advantage in scores. Attributed to neopotism.
Severin, Martins, Heyard, Delavy, Jorstad & Egger (2020)	Switzerland, SNSF	N=38250 Peer-review external reports on12,294 grant applications.	Women disadvantaged by 11% in favorable evaluation, by 11% in scores
Tamblyn, Girard, Qian, & Hanley (2018)	Canada, Canadian Institutes of Health Research	N=11,624 applications between 2012--2014.	Women disadvantaged by 0.09 points and applied sciences disadvantaged by 0.3 points in final application scoring
Bol, de Vaan, van de Rij (2022)	Netherlands Organization for Scientific Research	N=20,360 applications , 2005-2016.	No gender differences in funding. However, women receive 40% of a standard deviation lower evaluation scores than men. Panels adjust for lower scores.
van den Besselaar & Mom (2020)	EU, European Research Council	N=3,030 applicants for 2014 Starting Grant	Women disadvantaged by 0.0068/0.0067 in First Step and 0.037/0.031 in Second Step of CV/Project scoring. No significant gender gap in decision.
van der Lee & Ellemers (2015b)	Netherlands Organization for Scientific Research (NWO).	N=2,823 applications and reviews, 2010-2012.	Women 2.8 ppt less likely to be funded. Men received bonus for "quality of researcher" but there was no gender difference in "quality of proposal."
Volker & Steenbeek (2015)	Netherlands, NOW	Replication of Lee & Ellemers (2015)	No significant gender gap
Albers (2015)	Netherlands Organization for Scientific Research (NWO).	Replication of Lee & Ellemers (2015)	Accounting for multiple comparisons, no signifcant gender gap.
Waisbren et al. (2008)	US, Harvard Medical School	N = 6,319 faculty applications from 2,480 faculty	Women less likely to submit, requested fewer years and lower amounts of funding. Grant success same after controlling for academic rank.
Warner, Carapinha, Weber, Hill, Reede (2017)	US, Harvard Medical School	N=5,445 instructors and assistant professor faculty, 2008-2015	No gender difference in R01 awards. Only 7.6% received R01 funding.
Wennerås & Wold (1997)	Sweden, Medical Research Council	Peer-review applications for postdoctoral fellowships. N=114 from 1990s.	Women disadvantaged by 3.2 points in final scores

Citation	Country/Agency	Study Population & Type	Funding Gap
Witteman, Hendricks Straus & Tannenbaum (2019)	Canada, Canadian Institutes of Health Research	N=23918 applications 2011 d016.	Women disadvantaged by 0.9% in traditional program and new program with focus on the proposed science, by 4% in new program with focus on the caliber of investigator.