



March 17, 2015

The Honorable Eddie Bernice Johnson
Ranking Member
Committee on Science, Space, and Technology
House of Representatives

The Honorable Rosa DeLauro
House of Representatives

The Honorable Louise M. Slaughter
House of Representatives

Women in STEM Research: Federal Agencies Differ in the Data They Collect on Grant Applicants

Since the enactment of Title IX in 1972—which prohibits discrimination on the basis of sex in education programs and activities receiving any federal financial assistance¹—women have made significant gains in many academic fields. Nevertheless, recent research shows that women continue to lag behind men in academic and professional advancement in science, technology, engineering, and math (STEM) fields. In fiscal year 2014, colleges and universities around the country received nearly \$25 billion in federal funding for research in these fields.² You requested data regarding federal grant-making to women and men for research in STEM fields. This report provides the results of the first part of an ongoing review related to the representation of women in federal STEM research programs. Specifically, this report focuses on the extent to which federal agencies collect data that could be used to analyze differences, if any, in federal grants to women and men in STEM fields.³

We selected for review the six federal agencies that together funded approximately 90 percent of the federal government's investment in basic and applied research in fiscal year 2010 through 2012: the Department of Agriculture-National Institute of Food and Agriculture (USDA-NIFA), the Department of Defense (DOD), the Department of Energy (DOE), the National Aeronautics

¹Pub. L. No. 92-318, tit. IX, § 901, 86 Stat. 235, 373.

²This figure is reported annually by the National Science Foundation's National Center for Science and Engineering Statistics. The fiscal year 2014 figure is preliminary and includes the total amount of federal obligations for research at universities and colleges in the fields of computer science and mathematics, engineering, environmental sciences, life sciences, and physical sciences.

³Federal research grants are typically awarded to the institution (e.g., college or university) that employs the researcher and not directly to the individual researcher. In this report, we primarily focused on data that federal agencies collect about characteristics of the principal investigators that submit proposals for grant funding, not the institution that employs them.

and Space Administration (NASA), the National Institutes of Health⁴ (NIH), and the National Science Foundation (NSF).⁵ At these six agencies, we collected data from 16 different entities that administer grant programs for research in STEM fields.⁶ See figure 1 for the amount of research funding these six agencies provided to colleges and universities during fiscal year 2012, the most recent year for which complete agency-level data are available. We asked each agency to identify programs that fund research in STEM fields distributed through grants or cooperative agreements because these are the primary funding mechanisms by which colleges and universities receive research funds. A list of all the agencies and components included in this review can be found in enclosure I.

To conduct this work, we collected and reviewed relevant documentation from the agencies and components and interviewed agency officials. Specifically, we obtained and analyzed documentation about the data systems and data elements routinely tracked electronically on grant proposals and awards for research in STEM fields. We interviewed agency officials to understand how the data systems interface and how these data are maintained and used by the agencies for analysis. We assessed the reliability of the NSF data on the amount of research funding provided by federal agencies by reviewing documentation about how NSF's National Center for Science and Engineering Statistics collects and ensures the accuracy of these figures. We determined these data were sufficiently reliable for the purposes of this report.⁷ See enclosure I for more details on our scope and methodology.

We conducted this performance audit from March 2014 to March 2015 in accordance with generally accepted government auditing standards. These standards require that we plan and perform the audit to obtain sufficient, appropriate evidence to provide a reasonable basis for our findings and conclusions based on our audit objectives. We believe the evidence obtained provides a reasonable basis for our findings and conclusions based on our audit objectives.

Summary of Findings

Three of the six federal grant-making agencies we reviewed—the National Institutes of Health, the National Science Foundation, and the Department of Agriculture—routinely collect demographic information about the individual researchers who apply for and receive federal grants. Further, these agencies have data systems that could facilitate analysis of differences, if any, in grants made to women and men. The three agencies stated they use this demographic information for research and internal analysis of their applicant pool. The other three agencies in our review—the Department of Defense, the Department of Energy, and NASA—do not routinely collect demographic information about researchers who submit grant proposals and receive awards. These agencies or some of their components

⁴According to 2012 figures, NIH, an agency of the Department of Health and Human Services, was responsible for administering more than 90 percent of the department's research budget. Therefore, we focused on NIH in this report.

⁵Basic research concerns fundamental aspects of phenomena without specific applications toward products or processes in mind; applied research concerns gaining understanding necessary to determine the means by which a specific need may be met.

⁶The entities we reviewed within each agency vary and may be referred to as offices, commands, or components. For the purposes of this report, we will refer to these separate entities within a larger agency as "components".

⁷This report focused on whether the funding agencies have existing data systems and procedures in place to collect data. We did not evaluate the extent to which these data fields were populated, nor did we evaluate the quality or reliability of the data the agencies collected.

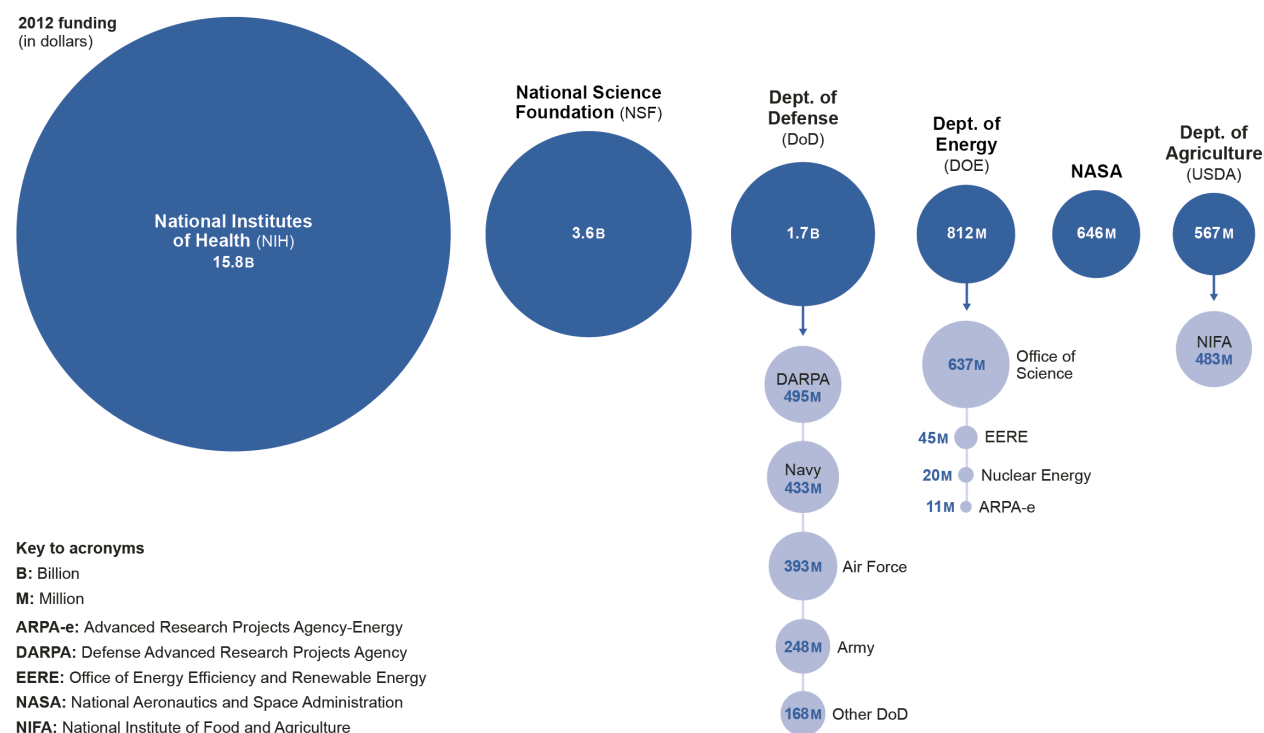
reported they were uncertain of the legal constraints around collecting these data and that their agencies had no internal purpose for such information.

Background

Federal Agencies That Fund Research in STEM Fields

The six federal agencies included in our analysis provide billions of dollars annually for research at colleges and universities in STEM fields. This report will focus on federal funding for research in the fields of engineering; life sciences (e.g., agricultural, biological, and environmental sciences); physical sciences (e.g., chemistry, earth sciences, and physics); computer and information technology; and math and statistics; where our prior work has shown women are still underrepresented.⁸ Figure 1 details the total amount of research funding provided in these STEM fields by each agency in fiscal year 2012.

Figure 1: Fiscal Year 2012 Federal Obligations for Select Areas of STEM Research Performed at Universities and Colleges



Source: National Science Foundation. | GAO-15-291R

Notes: The funding totals presented in figure 1 include only research obligations in fields that fall within the scope of our review (i.e., computer and information technology, engineering, environmental sciences, life sciences, math and statistics, and physical sciences), and may differ from total agency research funding obligations for FY 2012. For example, NSF's total STEM research obligations include funding for additional STEM fields such as psychology and social sciences. FY 2012 is the most recent year for which complete data are available for the agencies and components included in this review.

⁸GAO, *Science, Technology, Engineering, and Mathematics Education: Assessing the Relationship between Education and the Workforce*, [GAO-14-374](#) (Washington, D.C.: May 8, 2014).

Data to Calculate Success Rates and Analyze Differences between Women and Men in Grant Awards for Research in STEM Fields

The success rate is one statistic that allows for a comparison of federal grant-making between different groups of applicants, such as women and men.⁹ The success rate is the number of grants awarded divided by the number of grant proposals received.¹⁰ This calculation is important because simply comparing the total number of grants awarded to women to the total number of grants awarded to men does not take into account that there are fewer women participating in STEM fields.¹¹ In order to calculate success rates for women and men, data on the sex of the grant applicant—or principal investigator (PI)—needs to be collected and retained for both proposals and awards.¹² Ideally, each grant proposal would be assigned a unique ID that would also be assigned to any award that resulted from the proposal. This linkage between the record of the grant proposal and award ensures that for every grant award in the numerator of the success rate, there is a corresponding grant proposal in the denominator of the success rate. A data system that integrates or electronically links proposal and award data facilitates this type of analysis.

Calculating the success rate alone, however, does not produce complete and accurate information about what differences may exist between women and men receiving federal funding for research in STEM fields. First, it is also important to compare the size of grants awarded to women and men. Second, there may be other characteristics of the female and male applicants for research grants that influence the likelihood of success or award size. To determine if sex alone has any effect on whether an applicant is awarded a grant or on the size of the grant, one needs to compare women and men with similar educational and career characteristics. Therefore, federal agencies would need to collect data on these characteristics in order to accurately determine whether an individual's sex seems to have a significant effect on the likelihood of receiving a grant.¹³ Including variables on these types of characteristics in any analysis could be used to explain the source of differences, if any, between funding to women and men and could also clarify what proportion of any differences remains unexplained after controlling for educational and career characteristics.

Previous studies that have analyzed the factors affecting federal grant funding have also made use of data on researcher characteristics. Specifically, such studies have used demographic characteristics of principal investigators such as age, race, and sex; educational characteristics, such as highest degree and discipline of degree; and career characteristics, such as faculty position, years of experience, tenure status, and type and/or rank of the institution from which

⁹Total number of awards, total dollar amounts of awards, and average dollar amount of awards are also relevant statistics to compare funding differences between women and men.

¹⁰To calculate the number of grant proposals used to calculate the success rate, revisions to a grant are counted as part of the same grant application. The "award rate" is the term used if the number of grants awarded is divided by number of applications received where revisions are counted as separate applications.

¹¹According to NSF data, in fiscal year 2012, women made up around 35 percent of the graduate student and post-doctoral fellow population in STEM fields, and women were awarded 38.8 percent of doctoral degrees in STEM fields.

¹²Depending on the agency, grants can have multiple principal investigators, or co-investigators may also be designated. Previous research has suggested that ideally, the sex of all principal investigators and co-investigators would be recorded for every grant proposal received. For example, see Susan D. Hosek, et al, *Gender Differences in Major Federal External Grant Programs* (Santa Monica: RAND Corporation, 2005).

¹³The same information would be required to determine whether other researcher characteristics (e.g., age, disability, race, etc.) have an effect on the likelihood of receiving a grant.

the investigator is applying.¹⁴ For example, in 2005, the RAND Corporation completed a study on behalf of NSF that analyzed sex differences in federal grant-making using such variables to explain differences in grants awarded to women and men at NIH, NSF, and USDA.¹⁵ In their report, RAND cited the lack of adequate data collection and other limitations in the data collected by federal agencies that fund research in STEM fields as a major impediment to conducting their analyses.¹⁶ See enclosure I for a list of key variables that we identified as important to studying potential factors affecting grant funding differences between women and men from our review of previous studies.

Three of Six Agencies Routinely Collect Data That Allow for Analysis of Potential Differences in Federal Grant-Making between Women and Men

Three of Six Agencies Collect Data on the Demographic Characteristics of Applicants and Awardees

Three of the six agencies we reviewed routinely collect information on researcher characteristics that could be used to analyze differences in federal grants and cooperative agreements between women and men for research in STEM fields.¹⁷ While some components in all six agencies collect some data on researcher characteristics that would be helpful in conducting this type of analysis, only NIH, NSF, and USDA-NIFA routinely collect data on the sex of applicants and grantees, as well as a variety of other demographic and career characteristic variables that could be used to explain the gaps, if any, between women and men.¹⁸ While these three agencies do not require applicants to identify themselves as male or female, the percentages of individuals who voluntarily provide this information in response to demographic questions in the application are typically high.¹⁹ Both NIH and NSF maintain separate

¹⁴Academic institutions can be categorized using different classification systems. For example, the Carnegie classification system categorizes all colleges and universities in the U.S. based on types of degrees granted, population served and other characteristics. Other systems classify institutions by research rankings, such as those developed by The Center for Measuring University Performance, which divides research universities into different tiers based on research output and other characteristics.

¹⁵The National Science Foundation Authorization Act of 2002 (Pub. L. No. 107-368, § 18(c), 116 Stat. 3034, 3061), directed NSF to sponsor a study to “examine differences in amounts requested and awarded, by gender, in major Federal external grant programs”. See Susan D. Hosek, et al, *Gender Differences in Major Federal External Grant Programs* (Santa Monica: RAND Corporation, 2005).

¹⁶Both the RAND report and the National Academies of Science report *Beyond Bias and Barriers: Fulfilling the Potential of Women in Academic Science and Engineering* (Washington, D.C.: The National Academies Press, 2006) identified lack of data as an impediment to research in this area. Other studies of federal grant-making in STEM fields focused exclusively on NIH, which does collect data on researcher characteristics. See Pohlhaus, Jennifer Reineke, et al, “Sex Differences in Application, Success, and Funding Rates for NIH Extramural Programs.” *Academic Medicine*, vol. 86, no. 6 (2011). See also Ginther, Donna K., et al, “Race, Ethnicity, and NIH Research Awards”, *Science*, vol. 333, no. 1015 (2011).

¹⁷These three agencies accounted for approximately 85 percent of the \$25 billion in total federal STEM research funding to colleges and universities.

¹⁸We developed a list of variables needed to analyze factors that influence grant success by reviewing the methodology of relevant studies. For example, demographic variables include age, race, and sex, and career characteristic variables include discipline or highest degree attained. A list of key variables identified in these studies is included in enclosure I.

¹⁹According to agency officials at USDA-NIFA, in FY 2013, approximately 86 percent of grant records have the gender field populated; officials at NIH and NSF said the percentage is between 90-95 percent and 85-90 percent, respectively, over the past several years.

researcher records in their database with a unique researcher identifier to allow the individual's profile to be associated with multiple proposals and awards over time. Demographic data are not used as a part of the proposal review process; rather, their purpose is primarily to allow the agencies to analyze their own applicant and grantee pool. For example, NSF recently published a study of its own merit review process that analyzes the applicant and grantee pool by demographic characteristics.²⁰ In addition, these data have been used for other research purposes. For example, NIH data have been used in previous studies on both the effect of a primary investigator's sex and race on the likelihood of receiving a grant, and NIH regularly reports on the demographic breakdown of its grant applicants and recipients.²¹ Although USDA was included in the 2005 RAND study of this topic, USDA-NIFA has not recently published its own statistics on the characteristics of its applicants and grantees.

In contrast, the other three agencies—DOE, DOD, and NASA—do not collect data on whether grant applicants are male or female and vary in the extent to which they collect other researcher characteristics. Specifically, one of four components at DOE and three of eight components at DOD included in our review do not retain any electronic data on applicants, making calculation of a grant success rate for these components impossible. On the other hand, NASA and several components at DOE and DOD do collect data on some researcher characteristics (e.g., degree and discipline).²² Table 1 lists some of the relevant data agencies collect on researcher characteristics.

²⁰National Science Foundation, *Report to the National Science Board on the National Science Foundation's Merit Review Process Fiscal Year 2013*, NSB-14-32 (May 12, 2014).

²¹Ginther, Donna K., et al, "Race, Ethnicity and NIH Research Awards", *Science*, vol. 333, no. 1015 (2011). Pohlhaus, Jennifer Reineke, et al, "Sex Differences in Application, Success, and Funding Rates for NIH Extramural Programs", *Academic Medicine*, vol. 86, no. 6, (June 2011). Ley, Timothy J. and Barton H. Hamilton, "The Gender Gap in NIH Grant Applications", *Science*, vol. 332 (2008). See also National Institutes of Health, *NIH Data Book*, <http://report.nih.gov/nihdatabook/>.

²²NASA, through its Office of the Chief Scientist, is taking steps to pursue the feasibility of collecting additional demographic data (including age, race, and sex) about applicants to its competitive grant solicitations. NASA officials anticipate rolling out the implementation of this data collection effort in the fall of 2015.

Table 1: Researcher Information Collected by Agency Administrative Data Systems

| Variables collected | U.S. Dept. of Agriculture ^a | Dept. of Defense ^b | Dept. of Energy ^c | NASA | National Institutes of Health | National Science Foundation |
|---|--|-------------------------------|------------------------------|------|-------------------------------|-----------------------------|
| Sex | ✓ | | | | ✓ | ✓ |
| If tracked, approx. % of records with sex recorded ^d | 86% | | | | 90-95% | 85-90% |
| Race | ✓ | | | | ✓ | ✓ |
| Age/year of degree | ✓ | 1 of 8 components | | | ✓ | ✓ |
| Institution | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Highest degree | ✓ | 2 of 8 components | | | ✓ | ✓ |
| Discipline (or subject area of grant) | ✓ | 4 of 8 components | 3 of 4 components | ✓ | ✓ | |
| Co-Investigator information ^e | ✓ | 2 of 8 components | | ✓ | ✓ | |

Source: GAO analysis of agency documents. | GAO-15-291R

Notes: We did not evaluate the extent to which these data fields were populated (other than for the data on applicants' sex, where collected), nor did we evaluate the quality or reliability of the data the agencies collected.

^aThe only component included in our review at USDA was the National Institute of Food and Agriculture (NIFA).

^bThree of eight components included in this review at DOD do not retain proposal data for grant applicants once awards are made.

^cOne of four components included in this review at DOE does not retain proposal data for grant applicants once awards are made.

^dDemographic information such as sex and race are collected on a voluntary basis only; applicants are not required to provide this information to the agency. The percentages in this table were reported to us by the agencies. NIH and NSF provided us a range based on the past several fiscal years and USDA-NIFA provided us the percentage from their most recent full year of data, FY 2013.

^eIn this row, agencies received a checkmark if they collected any information regarding co-investigators, such as name or institution.

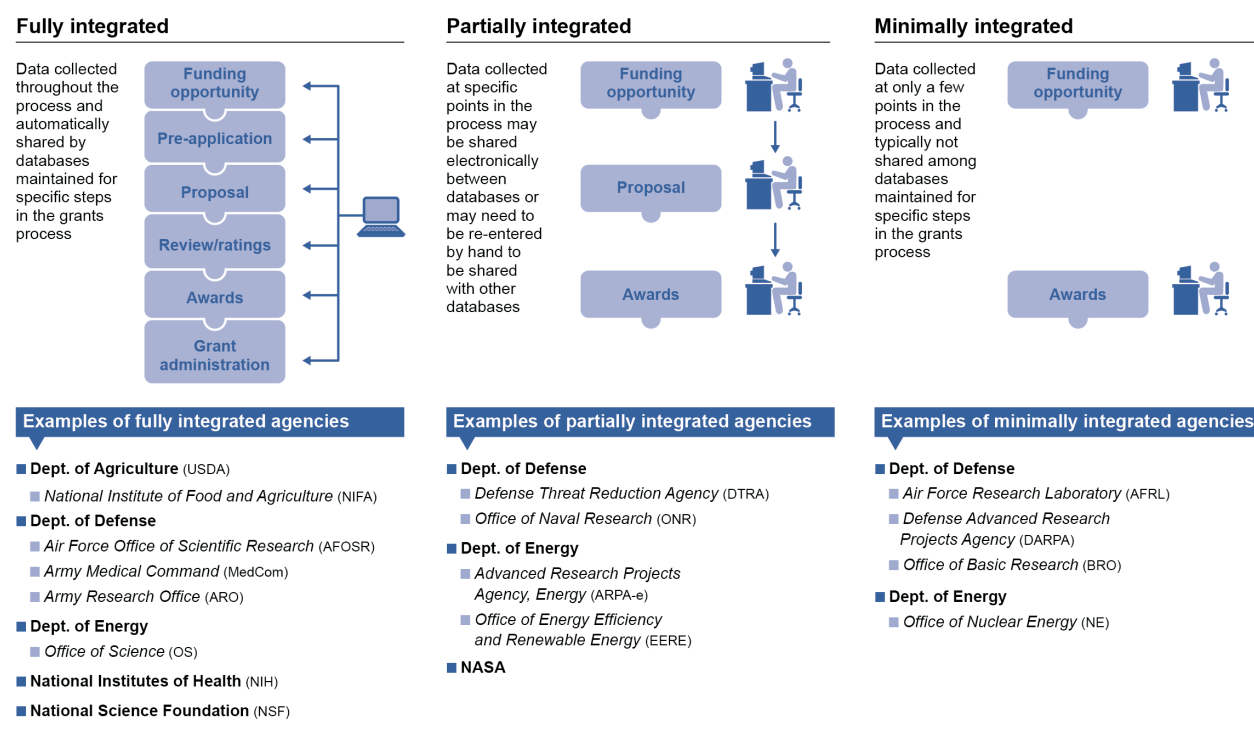
The three agencies we reviewed that routinely collect demographic data—NIH, NSF, and USDA-NIFA—each have an integrated data system that enables the agency to track the complete lifecycle of a single grant, and to link a single proposal with the award decision. Having linked proposal and award data enables the agency to compute and analyze success rates. Consequently, these agencies have data systems that could facilitate analysis of differences, if any, in grants made to women and men.

The remaining three agencies that do not routinely collect demographic data across all of their components—DOD, DOE, and NASA—have data systems that vary in the extent to which they integrate proposal and award data. In some cases, these data systems only partially or minimally integrate proposal and award data.²³ This makes it difficult to reliably link a single proposal with the award decision, and therefore to calculate a reliable grant success rate. For

²³Some components at DOD have integrated data systems. For example, though it does not routinely collect demographic information about its grant applicants, in 2014, the Air Force Office of Scientific Research at DOD began tracking proposals, awards, and grants in one system.

example, NASA and some components at DOD and DOE use separate databases that are not electronically linked, some of which retain data on proposals and others for grants awarded. In addition, some components at DOD and DOE use minimally integrated systems that retain no electronic data at all on applicants and track only minimal data on awardees. At DOE for example, the Office of Nuclear Energy does not retain data on grant applicants, only on awardees. At DOD, the Basic Research Office within Acquisition, Technology & Logistics (AT&L) also does not retain applicant data.²⁴ Without information on the number of applicants and information about those applicants, one cannot calculate a success rate overall or by gender or award type. See figure 2 for types of systems agencies use to collect data.

Figure 2: Types of Agency Data Systems Used to Collect Data on Grant Proposals and Awards



Source: GAO analysis of data from the agencies. | GAO-15-291R

Note: Relevant components are only listed in this graphic where data systems differ within the agency.

Agencies Cited a Variety of Reasons for Why They Do or Do Not Collect Demographic Data

Agencies reported several reasons for why they do or do not routinely collect these types of data. For example, four agencies and seven components told us they did not collect this information because there was no legal requirement to collect demographic data on applicants or grantees.²⁵ Further, one agency and five components cited concerns about whether they are

²⁴Officials at DOD told us that the Basic Research Office (in AT&L) makes approximately 10 awards per year. Minimal data on successful grantees is tracked in a spreadsheet; however, because of the small number of awards, officials said it would be feasible for the program manager in the Basic Research Office to provide detailed data on the grantees without the need for a formal database.

²⁵Title IX, the federal statutory provision prohibiting discrimination on the basis of sex in education programs and activities receiving any federal financial assistance, does not require federal agencies to collect any data on the sex of grant applicants or grantees. According to officials at the Department of Education's Office for Civil Rights and the

allowed to collect demographic data. Officials at one agency stated that they had been advised by their agency's general counsel not to ask demographic questions on their grant applications.

In contrast, the three agencies that have been collecting these data for a number of years—NIH, NSF, and USDA-NIFA—cited the desire to monitor their award processes as the basis for collecting these data. Both USDA-NIFA and NSF use the same personal data collection form approved by the Office of Management and Budget to collect demographic data, which provides the following guidance to applicants:

“The Federal Government has a continuing commitment to monitor the operation of its review and award processes to identify and address any inequities based on gender, race, ethnicity, or disability of its proposed PDs/PIs [Project Directors/ Principal Investigators] and co-PDs/PIs Submission of the requested information is voluntary and is not a precondition of award Upon receipt of the application, this form will be separated from the application. This form will not be duplicated, and it will not be a part of the review process.”

Similarly, NIH offers the following guidance to applicants concerning their voluntary submission of demographic data:

“Federal Agencies have a continuing commitment to monitor the operation of its [sic] review and award processes to detect, and deal appropriately with, any instances of real or apparent inequities Personal demographic data on PD/PIs [Project Directors/Principal Investigators] and those with a postdoctoral role is vital to comply with these requirements.”

Examples of data collection forms from these three agencies are included in enclosure II.

Agencies also varied in their capacity and willingness to collect data they deem peripheral to the grant decision making process. Two agencies and three components from a third agency indicated that they lacked the capacity in existing systems to collect these data. For example, officials at one component stated that the standard forms available from Grants.gov—a centralized electronic portal for applicants to find information about government funding opportunities—provided no mechanism for asking applicants for demographic information. However, other agencies include demographic data forms or links to such forms on Grants.gov.²⁶ Lastly, when we asked agencies and components why they did not collect certain data elements, the most common response from agencies and components was that they had no internal use for such data.

Agency Comments

We provided a draft of this report to the departments of Agriculture, Defense, Energy, Health and Human Services (specifically, the National Institutes of Health), the National Aeronautics

Department of Justice's Civil Rights Division, while Title IX does not require federal agencies to collect this type of data, it also does not prohibit it.

²⁶An E-Government initiative under the governance of the Office of Management and Budget, Grants.gov contains a repository of standard data collection forms for agencies' use in building proposal packages for funding opportunities, and agencies are able to customize these application packages, which can interface with their existing grant management systems. In addition, federal agencies can post funding opportunities and receive proposals through Grants.gov.

and Space Administration, and the National Science Foundation. DOE, NASA, and NSF provided technical comments, which we incorporated as appropriate. USDA, DOD, and NIH had no comments on the draft.

As agreed with your offices, unless you publicly announce the contents of this report earlier, we plan no further distribution until 30 days from the report date. At that time, we will send copies of this report to the Secretaries of Agriculture, Defense, and Energy, the Administrator of the National Aeronautics and Space Administration, the Directors of the National Institutes of Health and the National Science Foundation, appropriate congressional committees, and other interested parties. In addition, the report will be available at no charge on the GAO website at <http://www.gao.gov>.

If you or your staff have any questions about this report, please contact Melissa Emrey-Arras at (617) 788-0534 or emreyarrasm@gao.gov. Contact points for our Offices of Congressional Relations and Public Affairs may be found on the last page of this report. Major contributors to this report were Erin M. Godtland (Assistant Director), Nora Boretti (Analyst in Charge), and Jessica K. Rider. Also contributing to this report were James Bennett, Deborah Bland, David Chrisinger, William Egar, Ashley McCall, Amy Moran Lowe, Lindsay Read, James Rebbe, Douglas Sloane, Walter Vance, and David Watsula.



Melissa Emrey-Arras
Director, Education, Workforce,
and Income Security

Enclosures – 2

Enclosure I: Scope and Methodology

We reviewed the total amount of federal research and development funding from fiscal year 2010 through 2012 and selected the top six federal funding agencies to include in our review: the Department of Agriculture-National Institute of Food and Agriculture (USDA-NIFA), the Department of Defense (DOD), the Department of Energy (DOE), the National Aeronautics and Space Administration (NASA), the National Institutes of Health (NIH), and the National Science Foundation (NSF).²⁷ These six agencies funded more than 90 percent of basic and applied research in STEM fields in 2012. We assessed the reliability of the NSF survey data on federal funds for research and development by reviewing documentation about how the National Center for Science and Engineering Statistics collects these data from federal agencies and ensures the accuracy of these figures. We determined these data were sufficiently reliable for the purposes of our review. See table 2 for a list of all of the federal agencies and components we included.

Table 2: Agencies and Components in Our Review

| Agency | Agency Component |
|--|--|
| Department of Agriculture (USDA) | National Institute of Food and Agriculture |
| Department of Defense (DOD) | Air Force Office of Scientific Research Air Force Research Laboratory Army Medical Command Army Research Office Basic Research – Office of Assistant Secretary of Defense for Research and Engineering Defense Advanced Research Projects Agency Defense Threat Reduction Agency Office of Naval Research |
| Department of Energy (DOE) | Advanced Projects Agency (Energy) Office of Energy Efficiency and Renewable Energy Nuclear Energy Office of Science |
| National Aeronautics and Space Administration (NASA) | |
| National Institutes of Health (NIH) | |
| National Science Foundation (NSF) | |

Source: GAO. | GAO-15-291R

At each of these agencies and components, we provided a data collection instrument that asked officials to identify all programs that met the following inclusion criteria for our review:

²⁷According to 2012 figures, NIH, an agency of the Department of Health and Human Services, was responsible for administering more than 90 percent of the department's research budget. Therefore, we focused on NIH in this report.

1. Research funding in core STEM fields only: computer and information technology, engineering, math, physical and life sciences, and statistics.
2. Funding for basic and applied research at the graduate, post-graduate, and principal investigator or full professor level.
3. Funds distributed through external grants or cooperative agreements.

We limited our scope to external grants or cooperative agreements because these are the primary funding methods by which colleges and universities receive research funds. Officials at each agency and component we included identified the relevant STEM research programs at the agency level and, across these six agencies, we identified 16 different entities that administered these research grant-making programs with varying degrees of independence, depending on the agency's structure.

We conducted an extensive literature review to identify previous studies that have analyzed factors affecting whether a researcher receives federal funding. Such studies typically include in their analysis demographic characteristics such as age, race, and sex; educational characteristics such as highest degree, discipline of degree; and career characteristics such as tenure status, rank of current institution, and years of experience. Based on these studies and our own analysis, we developed a list of variables that would be useful in calculating success rates for women and men and to analyze what factors, including the sex of the principal investigator, could possibly affect whether a grant proposal is selected for funding at a federal agency. See table 3 for a list of key variables we identified through our own analysis and a review of the relevant literature.²⁸

²⁸See, for example, Susan D. Hosek, et al, *Gender Differences in Major Federal External Grant Programs* (Santa Monica: RAND Corporation, 2005) and Ginther, Donna K., et al, "Race, Ethnicity, and NIH Research Awards", *Science*, vol. 333, no. 1015 (2011).

Table 3: Key Variables Useful in Analyzing Success Rates and Explaining Any Differences in Success Rates between Women and Men

| Variables Used in Previous Studies | Purpose |
|--|---|
| Age | Experience proxy |
| Amount awarded | Main dependent variable |
| Amount requested | Main dependent variable |
| Award status (approved, declined) | Main dependent variable |
| Begin date of award | To calculate fiscal year |
| Degree | Explanatory variable |
| End date of award | To calculate fiscal year |
| Experience | Explanatory variable |
| Fiscal year of award | Explanatory variable |
| Institution | To identify institutional rank |
| Institutional rank | Explanatory variable |
| Institution type (college vs. research university) | Explanatory variable |
| Nonacademic applicant | Explanatory variable |
| Not PhD | Explanatory variable |
| Proposal receipt date | To calculate fiscal year |
| Discipline | Explanatory variable |
| Sex | Focus of research question: to analyze if the sex of applicant is a significant determinant |
| Tenure track/status | Explanatory variable |
| Type of proposal (reapply, modification, initial) | To filter out out-of-scope awards |
| Identified by Previous Studies as a Meaningful Explanatory Variable | |
| Award type | Explanatory variable, to filter/analyze awards by type |
| Co-Investigator data | Explanatory variable |
| Ethnicity | Explanatory variable |
| Race | Explanatory variable |
| Reviewer score | Explanatory variable |
| Type of academic appointment | Explanatory variable |
| Year of degree | Explanatory variable |
| Further Data Useful For Summary Grant Statistics | |
| Amount awarded for each corresponding award, gender | Total and mean amount awarded, total and by gender |
| Amount requested in each proposal, gender | Total and mean amount requested, total and by gender |
| Award unique ID, link to unique proposal ID, sex | To calculate success rates |
| Award type, unique award ID | All of the above by award type |
| Proposal unique ID | To calculate success rates |

Source: GAO analysis and review of literature. | GAO-15-291R

In order to determine whether agencies were collecting information on these characteristics, we interviewed officials at the agency level and at each of the relevant components to understand their grants management and procurement databases. We verified this information by reviewing

documents from each agency and component that explained these data systems and what data their systems collected. We reviewed documents such as:

- Data dictionaries
- Database handbooks
- Database table schema
- Forms used to collect data

Reviewing these documents, we identified the relevant data fields each agency or component collected and compared this information with the list of key variables we identified in our analysis of previous studies. Based on data fields we observed to be absent from data dictionaries and other documents, we created lists of relevant data each agency and component did not collect. We then confirmed the list of data not collected by a particular agency or component with officials at that particular agency or component. We also asked officials from each agency and component to provide the main reasons why they do or do not collect these data. It is important to note that this report focused on whether the funding agencies have existing data systems and procedures in place to collect these data and we did not evaluate the extent to which these data fields were populated, nor did we evaluate the quality or reliability of the data the agencies collected.

Enclosure II: Examples of Forms Agencies Use to Collect Demographic Data

Three agencies collect sex and other demographic characteristics of grant applicants. Both the National Science Foundation and the Department of Agriculture-National Institute of Food and Agriculture use a standard form approved by the Office of Management and Budget (OMB) for collecting these data as part of their proposal process. See figure 3.

Figure 3: OMB Form Used by the National Science Foundation and the United States Department of Agriculture-National Institute of Food and Agriculture for Collecting Applicant Demographic Data

OMB Number: 4040-0001
Expiration Date: 6/30/2016

| RESEARCH & RELATED PERSONAL DATA | | |
|--|----------------------|---|
| Project Director/Principal Investigator and Co-Project Director(s)/Co-Principal Investigator(s) | | |
| <p>The Federal Government has a continuing commitment to monitor the operation of its review and award processes to identify and address any inequities based on gender, race, ethnicity, or disability of its proposed PDs/PIs and co-PDs/PIs. To gather information needed for this important task, the applicant should submit the requested information for each identified PD/PI and co-PDs/PIs with each proposal. Submission of the requested information is voluntary and is not a precondition of award. However, information not submitted will seriously undermine the statistical validity, and therefore the usefulness, of information received from others. Any individual not wishing to submit some or all the information should check the box provided for this purpose. Upon receipt of the application, this form will be separated from the application. This form will not be duplicated, and it will not be a part of the review process. Data will be confidential.</p> | | |
| Project Director/Principal Investigator | | |
| Prefix: | * First Name: | Middle Name: |
| <input type="text"/> | <input type="text"/> | <input type="text"/> |
| * Last Name: | Suffix: | |
| <input type="text"/> | <input type="text"/> | |
| Gender: | <input type="text"/> | |
| Race (check all that apply): | Ethnicity: | Disability Status (check all that apply): |
| <input type="checkbox"/> American Indian or Alaska Native | <input type="text"/> | <input type="checkbox"/> Hearing |
| <input type="checkbox"/> Asian | | <input type="checkbox"/> Visual |
| <input type="checkbox"/> Black or African American | | <input type="checkbox"/> Mobility/Orthopedic Impairment |
| <input type="checkbox"/> Native Hawaiian or Other Pacific Islander | | <input type="checkbox"/> Other |
| <input type="checkbox"/> White | | <input type="checkbox"/> None |
| <input type="checkbox"/> Do Not Wish to Provide | | <input type="checkbox"/> Do Not Wish to Provide |
| Citizenship: | | |
| <input type="text"/> | | |
| <input type="button" value="Next Person"/> | | |

Source: GUSDA-NIFA Application Form. | GAO-15-291R

Source: USDA-NIFA Application Form. | GAO-15-291R

The National Institutes of Health collects personal data through the Electronic Research Administration (eRA) Commons, its electronic data collection and grants administration tool. Specifically, data are collected through the Personal Profile. The data are provided by the individual through a secure, electronic system; an example form is shown in Figure 4.

Figure 4: National Institutes of Health Form for Collecting Applicant Demographic Data

U.S. Department of Health & Human Services

eRA Commons
Sponsored by National Institutes of Health

JANE AUSTEN
AUSTEN
SAMPLE UNIVERSITY
PI IAR

NIH National Institutes of Health
Office of Extramural Research

Home Admin Institution Profile **Personal Profile** Status RPPR Internet Assisted Review xTrain Admin Supp eRA Partners

Personal Profile
JANE AUSTEN
Roles:
TRAINEE - Trainee
IAR - Internet Assisted Review User
PI - Principle Investigator
Person ID: 1234567

PERSONAL PROFILE SUMMARY
✓ NAME AND ID
✓ DEMOGRAPHICS
✓ EMPLOYMENT
✓ REVIEWER INFORMATION
✓ TRAINEE INFORMATION
✓ EDUCATION
✓ REFERENCE LETTERS
✓ PUBLICATIONS

Profile updated:
06/27/2013
[Change your password by:](#)
12/31/2013
eRA Commons Help Desk

Personal Profile Summary ?

NAME AND ID ? EDIT VIEW + View All

DEMOGRAPHICS ? EDIT CLOSE

✓ All questions answered.

Gender: Female Gender, ethnicity, race and disability are used for statistical reporting only.

Ethnicity: Non-Hispanic

Race: White

Disability:

Disadvantaged: Intentionally withheld

Federal Debt: No

EMPLOYMENT ? EDIT VIEW

Source: NIH eRA Commons Personal Profile Module User Guide. | GAO-15-291R

This is a work of the U.S. government and is not subject to copyright protection in the United States. The published product may be reproduced and distributed in its entirety without further permission from GAO. However, because this work may contain copyrighted images or other material, permission from the copyright holder may be necessary if you wish to reproduce this material separately.

GAO's Mission

The Government Accountability Office, the audit, evaluation, and investigative arm of Congress, exists to support Congress in meeting its constitutional responsibilities and to help improve the performance and accountability of the federal government for the American people. GAO examines the use of public funds; evaluates federal programs and policies; and provides analyses, recommendations, and other assistance to help Congress make informed oversight, policy, and funding decisions. GAO's commitment to good government is reflected in its core values of accountability, integrity, and reliability.

Obtaining Copies of GAO Reports and Testimony

The fastest and easiest way to obtain copies of GAO documents at no cost is through GAO's website (<http://www.gao.gov>). Each weekday afternoon, GAO posts on its website newly released reports, testimony, and correspondence. To have GAO e-mail you a list of newly posted products, go to <http://www.gao.gov> and select "E-mail Updates."

Order by Phone

The price of each GAO publication reflects GAO's actual cost of production and distribution and depends on the number of pages in the publication and whether the publication is printed in color or black and white. Pricing and ordering information is posted on GAO's website, <http://www.gao.gov/ordering.htm>.

Place orders by calling (202) 512-6000, toll free (866) 801-7077, or TDD (202) 512-2537.

Orders may be paid for using American Express, Discover Card, MasterCard, Visa, check, or money order. Call for additional information.

Connect with GAO

Connect with GAO on [Facebook](#), [Flickr](#), [Twitter](#), and [YouTube](#). Subscribe to our [RSS Feeds](#) or [E-mail Updates](#). Listen to our [Podcasts](#). Visit GAO on the web at www.gao.gov.

To Report Fraud, Waste, and Abuse in Federal Programs

Contact:

Website: <http://www.gao.gov/fraudnet/fraudnet.htm>

E-mail: fraudnet@gao.gov

Automated answering system: (800) 424-5454 or (202) 512-7470

Congressional Relations

Katherine Siggerud, Managing Director, siggerudk@gao.gov, (202) 512-4400, U.S. Government Accountability Office, 441 G Street NW, Room 7125, Washington, DC 20548

Public Affairs

Chuck Young, Managing Director, youngc1@gao.gov, (202) 512-4800 U.S. Government Accountability Office, 441 G Street NW, Room 7149 Washington, DC 20548

