TESTIMONY

Dr. Lynda T. Carlson
Division Director
Division of Science Resources Statistics
National Science Foundation

Before the

Committee on Science and Technology
Subcommittee on Research and Science Education

United States House of Representatives

May 8, 2008

Chairman Baird, Ranking Member Ehlers, and Members of the Subcommittee, I am Lynda Carlson, Director of the Division of Science Resources Statistics (SRS) within the National Science Foundation (NSF). I appreciate the opportunity to testify on Representative Johnson's proposed legislation on gender biases and barriers. However, NSF cannot support the proposed legislation as its requirements will be excessive as they exceed current data collection capabilities.

NSF's Division of Science Resources Statistics (SRS)

The Division of Science Resources Statistics (SRS) is the federal statistical agency responsible for data collection and analysis related to the entire science and engineering (S&E) enterprise. The division's responsibilities include data collections and analyses related to the S&E workforce, the education of scientists and engineers, and research and development (R&D), including federal funding of R&D. We annually collect data on R&D in academe and industry, and we periodically collect data on R&D funding activities by states and nonprofits. SRS staff is responsible for writing and producing the biennial *Science and Engineering Indicators* report for the National Science Board, as well as the biennial report *Women, Minorities and Persons with Disabilities in Science and Engineering,* which is required under Section 37 of the Science and Engineering Equal Opportunities Act.

Data Collection on Grants Portfolio

NSF currently collects annual composite information on demographics, field, award type and budget request, review score and funding outcome for NSF proposals and awards. NSF publishes a summary of these data in the annual Merit Review Report, including principal investigator (PI) demographics on proposals and awards. However, PIs are not, nor can they be, required to provide demographic information because of the Privacy Act; therefore, the demographic information collected is incomplete. For example, the number of PIs who submitted proposals and did not declare a race/ethnicity in 2007 is nearly as large as the number who declared minority status. In the last ten years, the proportion of new PIs who choose to report their gender has been declining.

Furthermore, the process for collecting and correlating review scores across programs and directorates within NSF is complex. For example, differences in average review scores across programs and field of research are as likely to reflect different reviewer community norms as to reflect differences in the actual quality of proposals received. Given the variety of review processes and scoring systems used throughout federal government, coupled with the complexity of correlating scores even within agencies, it would be virtually impossible for SRS to provide a report to Congress with review scores that are in any way comparable across the federal science agencies.

Lessons Learned from Other SRS Surveys

Over the last several years, SRS has been in the process of redesigning two surveys that characterize R&D conducted in the federal sector: (1) "Survey of Federal Funds for Research and Development," or Federal Funds Survey; and (2) "Survey of Federal Science and Engineering Support to Universities, Colleges, and Nonprofit Institutions," or Federal Support Survey. The surveys are being redesigned to better reflect how R&D is actually conducted in today's economy. The redesign was guided in part by a 2005 study that SRS commissioned from the National Research Council's Committee on National Statistics (CNSTAT), entitled "Measuring Research and Development Expenditures in the U.S. Economy."

For the two aforementioned surveys, SRS must obtain data from S&E funding agencies on the type of activity supported and on the recipient of the funding, among other indicators. It has been increasingly difficult for SRS to obtain high quality data in a timely manner from the queried agencies. Moreover, agencies do not usually keep detailed information about the fields of S&E that they support. Further, even when the agency does maintain data by field, those data may not conform to SRS's data categorization system. Different agencies maintain their records in quite different ways to meet their particular needs and operating procedures. SRS may have to work with individual agencies for significant periods of time to obtain more comparable data. Because of poor data quality and incomplete agency reporting, data on field of S&E research has not been collected as part of the Federal Support Survey since 1999.

In response to the issues we have encountered in conducting these two surveys, SRS has commissioned CNSTAT to form a panel and hold a series of workshops to assist us in their revision. The panel, "Modernizing the Infrastructure of the NSF Federal Funds Survey," was recently formed, and the first workshop will be held in June 2008. The panel's report and recommendations, which may help streamline data collection for SRS, are expected to be released in early 2009.

As part of the redesign effort for another SRS survey entitled, "Survey of Research and Development Expenditures at Universities and Colleges," issues with field of study data have been elucidated. Recent site visits to sixteen academic institutions have indicated that many academic institutions do not capture research field of study at the proposal stage. Once a proposal has been funded, the ability to capture the field of study for individual proposals varies considerably across institutions from easy to quite difficult. Institutions have indicated that it would require some effort to educate faculty on how to code their research by field, as the methods are not straightforward, especially as more and more research is interdisciplinary.

Lastly, SRS is also revising the existing taxonomy(s) of Fields of Science in order to capture new and emerging fields. SRS is developing a schema to revise the taxonomy in a manner that would allow it to be updated on a continuous basis. We expect this project to be finalized in two to three years. We will engage in significant consultation with the other science funding agencies as part of this activity.

Lessons Learned from a Study of Grants by Gender

The NSF Authorization Act of 2002 required NSF to "examine differences in amounts requested and awarded, by gender, in major federal external grants." SRS contracted with the RAND Corporation to conduct the survey, and the results were published in a 2005 report entitled, "Gender Differences in Major Federal External Grant Programs."

The report covered several federal science agencies, or federal agencies responsible for at least 2% of federal R&D obligations to universities. We had intended that the study collect data on grants by gender from NSF, the Department of Defense (DoD), the Department of Energy (DoE), the Department of Agriculture (USDA), and the National Institutes of Health (NIH). Data collection was only feasible from NSF, NIH and USDA; adequate data on grant applications and awards were not available from DoD or DoE. According to the report:

> "[There are] numerous limitations in the information collected in federal agencies' grant application and award data systems. Such limitations hinder the ability to track gender differences in federal grant funding. Better tracking of gender differences in such funding would require that all agencies awarding significant grant funding do the following:
>
> • Maintain a data system that stores information on all grant applications and investigators, including co-investigators. Ideally, each agency would have a single data system rather than separate systems for each sub-agency or grant program and the agencies would agree on a common list of key data elements.
>
> • Include in the application form key personal characteristics for each investigator, including gender, race and ethnicity, institution (in a way that can be easily categorized), type of academic appointment for investigators in postsecondary education, discipline, degree, and year of degree.
>
> • Fill in missing personal information, including gender, where possible from other applications by the same investigator.

- Record the amount requested and awarded for each proposal and any score assigned to it by the peer reviewers.

- Clearly identify initial proposals and awards, supplements that involve new funding, and amendments that involve no new funding."

Cost of Survey Implementation

Current, simple federal surveys conducted by SRS cost approximately $800,000 annually to implement. The costs are incurred by a survey firm contracted to collect and process the data. This expense does not include the cost of SRS staff, who provide oversight and administration of the survey efforts, or the costs of collection and reporting incurred by each of the individual federal agencies.

If NSF were tasked to expand its data collection efforts to include the more complex project-specific and demographic data envisioned in the proposed legislation, SRS would require additional funding, or we would have to reduce other ongoing survey efforts. These costs do not include the additional SRS staff time and resources that would be required to facilitate the data surveys, nor the additional costs that would be incurred by other federal agencies in setting up the requisite data systems and annually reporting the data to SRS.

Conclusion

Mr. Chairman, I hope that I have been able to articulate NSF's unique role in gathering and analyzing data about the nation's S&E enterprise. I hope my comments help feed the discussion about how to collect indicators adequately to help our nation measure our progress in ensuring that there is no gender bias in science and technology.

In summary, however, SRS does not have the ability to require funding agencies to maintain such records. If Congress seeks to require such a collection, the Grants Policy Committee, which is charged with overseeing government-wide grants policy initiatives and making policy recommendations to the Office of Management and Budget, might be able to provide additional insight.

SRS does welcome the opportunity, however, to continue to be involved in discussions on this important draft legislation, as we are constantly striving to improve our contribution to the policy process. NSF looks forward to collaborating with our sister agencies and the broader S&E community to more effectively collect and report on important data related to innovation and competitiveness. Thank you for the opportunity to appear before you, and I am happy to answer any questions.

About Dr. Lynda T. Carlson

Since 2000, Dr. Lynda Carlson has been the Director of the National Science Foundation's Division of Science Resources Statistics. In that role, she is responsible for all activities of the Division, a federal statistical agency within NSF. Prior to coming to NSF, Dr. Carlson was at the Energy Information Administration (EIA) of the Department of Energy where she held a variety of positions over 23 years. She is internationally known for the design and development of the nation's energy consumption surveys, including the development of a unique statistical sampling frame of commercial buildings. Dr. Carlson's last position at EIA was that of Director of the Statistical Methods Group with responsibility for all statistical activities throughout EIA.

Dr. Carlson received her M.A. and Ph.D. from the University of Illinois at Urbana-Champaign in Political Science and her B.A. from Brooklyn College, CUNY. She is a Fellow of the American Statistical Association and member of various groups such as AAPOR and AAAS. She has served on a series of OMB committees, is a member of the Federal Committee on Statistical Methodology, and has served on several NAS committees. In 2000, she received the highest departmental award from the Department of Energy for her service to that agency.

Dr. Carlson has written on energy consumption, survey methodology, and the science and engineering enterprise.