GAO
Accountability * Integrity * Reliability

United States General Accounting Office
Washington, DC 20548

May 9, 2001

The Honorable John McCain
Chairman
Committee on Commerce, Science,
and Transportation
United States Senate

The Honorable Bill Frist
United States Senate

Subject: Research and Development Funding: Reported Gap Between Data From Federal Agencies and Their R&D Performers Results From Noncomparable Data

In a January 2000 report, the National Science Foundation (NSF), responsible for disseminating science and engineering information, identified a $5 billion gap between the amount of money that federal agencies reported as research and development (R&D) support and the amount of money that the performers of the R&D work reported as spent in 1998. According to NSF, federal agencies obligated about $72 billion for R&D support in 1998, while the performers of federal R&D (including industries, universities, and other nonprofit organizations) reported spending about $67 billion. This reported gap, which was first identified in the 1980s, has created perceptions that data-quality problems exist or that performers are not receiving or spending all the federal R&D funds obligated to them.

Legislation establishing NSF included, among its various responsibilities, the requirement to collect, analyze, and disseminate a variety of R&D data. NSF sponsors a series of surveys of federal agencies and R&D performers to collect data on the financial and human resources devoted to R&D in the various sectors of the U.S. economy. The results of these surveys are published in various individual reports, including NSF's biennial *Science and Engineering Indicators* report. Such data have been used as a barometer to help gauge the overall health and vitality of the nation's R&D enterprise. Because of concerns about whether Members of Congress can rely on NSF's data to fulfill their oversight and legislative responsibilities, you asked us to review the procedures for collecting and reporting federal R&D funding data. Through discussions with your offices, we agreed to determine why a gap exists between federal agency-reported data and performer-reported data. We briefed your staff on the results of our review on April 4, 2001, using the enclosed briefing slides. (See enc. I.)

The following summarizes our findings:

The gap results primarily from annually comparing two separate and distinct types of financial data—federal obligations and performer expenditures—that are not comparable. Obligations are estimates of payments to be made by federal agencies without regard to when the payments may (or may not) ultimately be made. These obligations are made on the basis of the availability of appropriations and on the assumption that the performers will meet their expectations. Expenditures represent actual performer cost or expense data without regard to when the federal obligation was made. These expenses can occur years after the obligation.

Data collection and reporting issues may also contribute to the gap, as shown in the following examples:

R&D funding data are collected on a yearly basis, but, in reporting, the period that defines a year can vary. Federal agencies report by the federal fiscal year—October 1 to September 30. In contrast, performers report by the calendar year or a different fiscal year. For example, universities generally follow a July 1 to June 30 fiscal year, while industries use a variety of fiscal and calendar years.

Agencies and performers do not always agree on what type of activities fall under the category of R&D. While federal agencies often consider program management as an R&D activity, performers may not. Furthermore, for multitask contracts where the tasks are split between procurement and R&D activities, different performers may report the entire contract as R&D or procurement.

Expenditure data from industrial performers are composed of both reported and imputed data. The imputed data, which are used to account for missing information, are based on an industrial performer's past expenditures and on data reported by others in the same industry. In 1998, 16 percent of the estimate for total federally funded R&D was imputed. This imputation varied by industry—from zero to 99 percent. Some of the higher percentages included 69 percent from the industrial chemical industry and 99 percent from the paper and allied products industry.

We did not assess the accuracy of the reported data. Therefore, we do not know the exact extent to which the collection and reporting issues contribute to the gap between obligations and expenditures.

Because the gap is the result of comparing two dissimilar types of financial data, it does not necessarily reflect poor-quality data, nor does it reflect whether performers are receiving or spending all the federal R&D funds obligated to them. Thus, even if the data collection and reporting issues were addressed, a gap would still exist. Expenditure data do not provide a check on the accuracy of obligation data. Expenditure data are useful for identifying trends in the conduct of R&D in any given

year by various performers, whereas obligation data are useful for identifying trends in federal R&D funding.

Agency Comments

We provided NSF and the U.S. Census Bureau with a draft of this letter for their review and comment. We met with NSF officials, including the Program Director, Research and Development Statistics, and the Economist from the Division of Science Resources Studies. We also met with Census officials, including the Assistant Chief and Branch Chief of the Manufacturing and Construction Division. These officials generally agreed with the message of the report. They provided some technical clarifications that we incorporated as appropriate.

Scope and Methodology

We conducted our review from November 2000 to April 2001 in accordance with generally accepted government auditing standards. We interviewed officials from NSF and the organizations that conduct NSF's surveys, including the U.S. Census Bureau. In addition, we reviewed various reports and supporting surveys addressing R&D funding. We did not assess the reliability of NSF's data. However, these data are widely used for R&D trend purposes, and NSF relies on them for its legislative reporting responsibilities.

- - - - -

As arranged with your offices, unless you publicly announce its contents earlier, we plan no further distribution of this letter until 7 days after the date of this letter. At that time, we will send copies to Dr. Lynda T. Carlson, Director, Division of Science Resources Studies, National Science Foundation, and William G. Barron, Jr., Acting Director, U.S. Census Bureau. We will also make copies available to others on request. The report will also be available on GAO's home page at http://www.gao.gov.

If you have any questions about this letter or need additional information, please call me on (202) 512-3841 or Robin Nazzaro on (202) 512-6246. Key contributors to this report were Diane Raynes, Michael J. Wargo, and Sandy Joseph.

Jim Wells
Director
Natural Resources
and Environment

Enclosure

GAO
Accountability * Integrity * Reliability

R&D Funding

Reported Gap Between Data From Federal Agencies and Their R&D Performers Results From Noncomparable Data

Prepared for the Senate Committee
on Commerce, Science,
and Transportation

4/4/01

GAO
Accountability * Integrity * Reliability

Introduction

- In a January 2000 report, the National Science Foundation (NSF) identified a $5-billion gap between data reported by federal agencies and by various R&D performers in 1998. Performers include industry, universities, and nonprofit organizations.
- Federal agencies reported about $72 billion for R&D support to the various performers, while the performers reported spending about $67 billion.
- The gap, which was first identified in the 1980s, has created perceptions that data quality problems exist or that performers are not receiving or spending all the federal R&D funds obligated to them.

2

GAO
Accountability * Integrity * Reliability

Objective and Results in Brief

Objective: Determine why a gap exists between federal agency-reported obligation data and performer-reported expenditure data.

Results in Brief: A gap exists for several reasons:

- Primarily, the gap results from annually comparing two separate and distinct types of financial data--federal obligations and performer expenditures--that are not comparable.
- Other potential reasons that contribute to the gap include data collection and reporting issues such as the use of estimated data. However, even if these issues were addressed, a gap would still exist.

3

GAO
Accountability * Integrity * Reliability

Background

- In serving as a clearinghouse for collecting, interpreting, and analyzing data on scientific and engineering resources, NSF funds the collection of data through several surveys, including
  - Federal Funds for R&D (conducted by Quantum Research Corporation);
  - Industrial R&D (conducted by the Bureau of the Census);
  - Federal Science and Engineering Support to Universities, Colleges, and Nonprofit Institutions (conducted by Quantum Research Corporation);
  - R&D Expenditures at Universities and Colleges (conducted by Quantum Research Corporation).
- In its biennial reports on science and engineering indicators, NSF notes the funding gap.

4

Noncomparable Data

While a number of reasons may contribute to the R&D funding gap, noncomparable data is the primary reason. The gap represents the difference between two dissimilar types of financial data-- obligations and expenditures--that are not comparable.

- Obligations are estimates of payments to be made by federal agencies without regard to when payments may (or may not) ultimately be made.
- Expenditures represent actual performer cost or expense data without regard to when the federal obligation was made. These expenses can occur years after the obligation.

5

Data Collection and Reporting

Data collection and reporting issues may also contribute to the funding gap. For example:

- Federal agencies report by federal fiscal year, while performers report by calendar year or a different fiscal year.
- Federal agencies and performers do not always agree on what type of technical activities fall under the category of R&D.
- Due to missing information, about 16 percent of the federally funded industrial R&D expenditures for 1998 were imputed. These data are based on an industrial performer's past performance and on data reported from others in the same industry.

6

GAO
Accountability * Integrity * Reliability

Concluding Observations

- The gap is the result of comparing two dissimilar types of financial data. Furthermore, the gap does not necessarily reflect poor quality data, nor does it reflect whether performers are receiving or spending all the federal R&D funds obligated to them.

- Expenditure data do not provide a check on the accuracy of obligation data. Obligation data are useful for identifying trends in federal R&D funding, whereas expenditure data are useful for identifying trends in the conduct of R&D in any given year by various performers.

7

GAO
Accountability * Integrity * Reliability

Scope and Methodology

- We reviewed NSF's report, *Science and Engineering Indicators 2000*, as well as its *National Patterns of R&D Resources: 1998* and specific performer reports and their supporting surveys. We also reviewed reports on workshops focusing on the R&D funding gap, including a Congressional Research Service report, *Challenges in Collecting and Reporting Federal Research and Development Data*, January 2000, and Quantum Research Corporation agency-specific reports.

- We interviewed officials from NSF and the Bureau of the Census.

- We did not assess the reliability of NSF's data. However, these data are widely used for R&D trend purposes, and NSF relies on them for its legislative reporting responsibility.

8

(360026)