

**SF-83-1 SUPPORTING STATEMENT**

**For**

**2008 Survey of Doctorate Recipients**

## TABLE OF CONTENTS

<b>A. Justification.....</b>	<b>3</b>
1. Necessity for Information Collection.....	3
2. Uses of Information .....	5
3. Consideration of Using Improved Technology.....	7
4. Efforts to Identify Duplication.....	8
5. Efforts to Minimize Burden on Small Business.....	8
6. Consequences of Less Frequent Data Collection.....	9
7. Special Circumstances .....	9
8. Federal Register Announcement and Consultations Outside the Agency .....	9
9. Payment or Gifts to Respondents.....	12
10. Assurance of Confidentiality .....	13
11. Justification for Sensitive Questions.....	13
12. Estimate of Respondent Burden.....	14
13. Cost Burden to Respondents.....	14
14. Cost Burden to the Federal Government.....	14
15. Reason for Change in Burden .....	14
16. Schedule for Information Collection and Publication.....	14
17. Display of OMB Expiration Date .....	15
18. Exception to the Certification Statement .....	15
<b>B. Collection Of Information Employing Statistical Methods .....</b>	<b>16</b>
1. Respondent Universe and Sampling Methods .....	16
2. Statistical Procedures .....	16
3. Methods to Maximize Response .....	17
4. Testing of Procedures .....	19
5. Contacts for Statistical Aspects of Data Collection .....	28

### **LIST OF ATTACHMENTS**

**Attachment 1:** National Science Foundation Act of 1950

**Attachment 2:** First Federal Register Notice for 2008 SDR

**Attachment 3:** Proposed 2008 SDR Mailing Materials

**Attachment 4:** Proposed 2008 SDR Questionnaire

**Attachment 5:** 2008 SDR/ISDR Sample Strata and Sample Allocation Table

**Attachment 6:** Changes in the 2008 SDR Questionnaire

**Attachment 7:** 2008 SDR Experiment Plan

## **2008 SURVEY OF DOCTORATE RECIPIENTS**

### **Supporting Statement**

#### **A. Justification**

This request is for a three-year reinstatement, with change, of the previously approved OMB clearance for the Survey of Doctorate Recipients (SDR). The SDR was last conducted in 2006 and the OMB clearance for the 2006 SDR expires February 28, 2009.

The SDR is one of three principal surveys that provide data for the National Science Foundation's (NSF) Scientists and Engineers Statistical Data System (SESTAT). The purpose of the SESTAT database is to provide information on the entire U.S. population of scientists and engineers with at least a bachelor's degree. SESTAT is produced by combining data from the Survey of Doctorate Recipients (SDR; representing persons in the general U.S. population who have earned a doctorate in science, engineering or health (SEH) from a U.S. institution), the National Survey of Recent College Graduates (NSRCG; representing persons with a recently earned bachelor's or master's degree in SEH from a U.S. institution) and the NSCG (representing all other individuals in the U.S. with a bachelor's degree or higher in a science or engineering (S&E) or S&E-related field, or those who had a bachelor's or higher degree in some other field, but had an S&E occupation or S&E-related occupation, including individuals who received degrees only from foreign institutions). The NSCG population is primarily drawn from eligible individuals from the decennial census.

The SESTAT integrated database derived from these surveys represents the demographic, educational, and employment characteristics of college-educated scientists and engineers in the United States. All three of these surveys are usually conducted every two years. The primary purpose of the SDR is to provide information on the doctoral scientists and engineers in the U.S. The SDR is comprised of 2 components: a longitudinal panel that tracks doctorate recipients throughout their careers until age 75, and a new cohort component that adds new doctorate recipients when they receive their doctorates. The panel portion of the SDR provides information on the stock, while the new sample in the SDR provides important data on the addition of new doctorate recipients with science, engineering and health degrees to the labor force. The SDR, as part of the SESTAT data system, is the only available source that provides detailed information to support a wide variety of policy and research analyses on science, engineering and health (SEH) labor force issues.

To provide complete representation of U.S. scientists and engineers at all degree levels, SESTAT was designed as a unified database that integrates information from all three component surveys. The system of surveys, created for the 1993 survey cycle and developed throughout the 1990s, is closely based on the recommendations of the National Research Council's Committee on National Statistics (CNSTAT) report to NSF<sup>1</sup>. That report recommended a data collection design based on three surveys, of which one (the NSCG) would be linked to the decennial Census.

#### **1. Necessity for Information Collection**

The Survey of Doctorate Recipients is sponsored by the National Science Foundation and the National Institutes of Health. The SDR is a longitudinal survey that was initiated in 1973 and was

---

<sup>1</sup>National Research Council. Committee on National Statistics. (1989). *Surveying the Nation's Scientists and Engineers: A Data System for the 1990s*. Washington: National Academy Press.

conducted on a biennial basis for NSF by the National Research Council (NRC) through 1995. From 1977 to 1995, the SDR also included a sample of doctoral recipients in the humanities with the support of the National Endowment for the Humanities (NEH). Their statutory requirement was defined in the National Foundation on Arts and Humanities Act (P.L. 209) as amended. The 1997 SDR was conducted by the National Opinion Research Center (NORC). The U.S. Census Bureau conducted the 1999 SDR and 2001 SDR. The 2003 and 2006 SDR rounds were conducted by NORC, and the 2008 SDR will be conducted by NORC.

The National Science Foundation Act of 1950, as amended by Title 42, United States Code, Section 1862 requires the National Science Foundation to:

...“provide a central clearinghouse for the collection, interpretation, and analysis of data on scientific and engineering resources and to provide a source of information for policy formulation by other agencies of the Federal Government...” (see Attachment 1 – National Science Foundation Act of 1950)

In meeting its responsibilities under the NSF Act, the Foundation relied on the National Register of Scientific and Technical Personnel from 1954 through 1970 to provide names, location, and characteristics of U.S. scientists and engineers. Acting in response to a Fiscal Year 1970 request of the House of Representatives Committee on Science and Astronautics (see U.S. Congress, House of Representatives, 91st Congress, 1st Session, Report No. 91-288), the Foundation, in cooperation with the Office of Management and Budget and eight other agencies, undertook a study of alternative methods of acquiring personnel data on individual scientists and engineers.

The President's budget for Fiscal Year 1972, as submitted to the Congress, recommended the "discontinuation of the National Register of Scientific and Technical Personnel in its present form" and that funds be appropriated "to allow for the development of alternative mechanisms for obtaining required information on scientists and engineers." The House of Representatives Committee on Science and Astronautics in its report on Authorizations for Fiscal Year 1972 states that "...it has no objection to this recommendation...." (see U.S. Congress, House of Representatives, 92nd Congress, 1st Session, Report No. 92-204).

Subsequently, the NSF established and continues to maintain the SESTAT system of surveys, the successor to the Scientific and Technical Personnel Data System of the 1980s which was the successor to the National Register. The Science and Technology Equal Opportunities Act of 1980 directs NSF to provide to Congress and the Executive Branch an “accounting and comparison by sex, race, and ethnic group and by discipline, of the participation of women and men in scientific and engineering positions.” The SESTAT database, of which the SDR is a major part, provides much of the information to meet this mandate.

The longitudinal data from the SDR provide valuable information on training, career and educational development of the Nation’s doctoral SEH population. These data enable government agencies to assess the scientific and engineering resources available in the United States to business, industry, and academia, and to provide a basis for the formulation of the Nation's science and engineering policies. Educational institutions use SDR data in establishing and modifying scientific and technical curricula, while various industries use the information to develop recruitment and remuneration policies.

In addition, the information collected from the SDR is needed by NSF to prepare two Congressionally mandated biennial reports: *Women, Minorities, and Persons with Disabilities in*

## *Science and Engineering and Science & Engineering Indicators.*

The Committee for Equal Opportunity in Science and Engineering (CEOSE), an advisory committee to the NSF and other government agencies, established under 42 U.S.C. §1885c, has been charged by the U.S. Congress with advising NSF in assuring that all individuals are empowered and enabled to participate fully in science, mathematics, engineering and technology. Every two years CEOSE prepares a congressionally mandated report that makes extensive use of the SESTAT data to highlight key areas of concerns relating to students, educators and technical professionals.

Congress also created the Commission on the Advancement of Women and Minorities in Science, Engineering and Technology Development (P.L.105-255, October 1998). In this legislation, the Commission was mandated to analyze and describe the current status of women, underrepresented minorities, and persons with disabilities in the science, engineering, and technology pipeline from early classroom education through their professional lives in industry, academe, and government. The final report, Land of Plenty: Diversity as America's Competitive Edge in Science, Engineering and Technology (September 2000), made extensive use of the SESTAT data to answer critical questions on the status of these groups in the workforce.

The importance of information on the scientific and technical workforce to inform public policy can be seen in discussions of the National Science Board's Task Force on National Workforce Policies for Science and Engineering. The taskforce relied heavily on SESTAT data to inform its deliberations about the S&E workforce and SESTAT data were an integral part of the taskforce's final report. (See <http://www.nsf.gov/nsb/documents/2003/nsb0369>.)

## **2. Uses of Information**

The time-series data produced by the SDR on the demographic, employment, and other characteristics of the Nation's SEH doctoral scientists and engineers have been used extensively in the policy and planning activities of the Foundation and the National Institutes of Health. The SDR data are used in assessing the quality and supply of the Nation's S&E personnel resources for educational institutions, private industry, and professional organizations as well as federal, state, and local governments. Other federal agencies, such as Department of Commerce, USDA, DOE, and NASA, as well as state agencies request and make use of the SDR data for a variety of purposes.

Information from the 2008 SDR will enable the NSF to continue to monitor the employment patterns of recently graduated SEH doctorates, as well as for experienced doctorates in the labor market. The SDR data on the state of SEH doctorates are used for presentations to the National Science Board. Within the Foundation, SDR data are used in the evaluation and development of programs in the Education and Human Resources (EHR) Directorate, and analysis of employment pathways by several research directorates.

The SDR provides data on the education and training, work experience, and career development of persons holding SEH doctorates from U.S. institutions. Without this information, those at the NSF, along with researchers and policymakers, would be less informed when carrying out their responsibilities. The SDR data are made available through published reports, the SESTAT on-line data system, public use files and restricted licenses. The SDR data from the past decade are incorporated in the SESTAT on-line data system for each year (1993-2006) and are available as a component of the SESTAT public-use data files, or as separate stand-alone public-use files. The

SESTAT on-line system allows Internet users to create customized data tabulations with a user-specified subject area. The SESTAT Home Page can be accessed at <http://www.nsf.gov/statistics/sestat>.

Some examples of recent SDR data uses are:

NSF Publications (all NSF publications can be accessed on the SRS website at <http://www.nsf.gov/statistics>.)

Congressionally mandated reports –

- *Science & Engineering Indicators*
- *Women, Minorities and Persons with Disabilities in Science and Engineering*

Other NSF publications –

- biennial report series: *Characteristics of Doctoral Scientists and Engineers*
- biennial report series: *Doctoral Scientists and Engineers Profiles*
- *Characteristics of Science and Engineering Doctorate Recipients: Selected Trend Tables*
- Annual report series: *Science and Engineering State Profiles*
- *Characteristics of Retiring Doctoral Scientists and Engineers Working in Educational Institutions (forthcoming)*
- *Employment Status and Career Choice of Women in the Social Sciences (forthcoming)*
- *Postdoc Participation of Science, Engineering and Health Doctorate Recipients (March 2008)*
- *Why Did They Come To The United States? A Profile of Immigrant Scientists and Engineers (June 2007)*
- *All In a Week's Work: Average Work Weeks Of Doctoral Scientists And Engineers (December 2005)*

Selected Presentations by non-NSF staff:

- *Task Assignments: Generalists vs. Specialists*, Economic Theory Workshop, University of Melbourne, September 2007.
- *Early Careers for Biomedical Scientists: Doubling (and Troubling) Outcomes*, Science and Engineering Workforce Project, February 2007.
- *Retirement Behaviors of Physicians Based on the Physicians Over 50 Survey – Preliminary Findings*, 2006 AAMC Physician Workforce Research Conference, May 2006.
- *Data That Do Not Persuade: Results of an Incentive Experiment*, 2005 Federal Committee on Statistical Methodology Research Conference, November 2005
- *Gender Equity in Higher Education: What the President of Harvard Doesn't Know or How Molehills Become Mountains of Inequity*, University of Wisconsin Women's Studies, March 2005
- *The Success of Female Scientists in the 21<sup>st</sup> Century*, Faculty Horizons Workshop, July 2005
- *A Brain is a Terrible Thing to Lose: Locating U.S.-Educated Foreign Nationals Intending to Live Abroad*, 2005 American Association for Public Opinion Research, May 2005

Selected Citations of SDR data in other sources:

- *Beyond Bias and Barriers: Fulfilling the Potential of Women in Academic Science and Engineering*, Committee on Science Engineering and Public Policy, 2007.
- *Foreign-Born Academic Scientists and Engineers: Producing More and Getting Less than their U.S.-Born Peers?*, Research in Higher Education, December 2007.

- *Job Satisfaction Of The Highly Educated: The Role Of Gender, Academic Tenure, And Earnings*, Scottish Journal of Political Economy, May 2006.
- *Educational Mismatch among Ph.D.s: Determinants and Consequences*, National Bureau of Economic Research Working Paper 12693, October 2006.
- *Gender Differences in Major Federal External Grant Programs*, Rand Corporation Technical Report, 2005
- *Who's Patenting in the University? Evidence from the Survey of Doctorate Recipients*, Economics of Innovation and New Technology, July 2005
- *UC Berkeley, Leads Nation In Prepping Students For Doctorates*, UC Berkeley News, January 2005

#### Data Dissemination and Access:

The NSF makes data from the SESTAT system of surveys available through published reports, the SESTAT on-line data system, public use files, and restricted licenses. Since 2005, NSF has distributed over 550 copies of SDR public-use files (2001 and 2003 survey cycles), as well as over 800 copies of the SESTAT public-use files (1993-2003 survey cycles). The SDR is also a component of the SESTAT integrated file, which describes the entire science and engineering workforce. There are currently approximately 30 licensees of the 2001 or 2003 SDR; an additional 10 licensees used these data, but their licenses have now been closed. There are pending requests for the 2006 SDR data.

### **3. Consideration of Using Improved Technology**

The 2008 SDR will collect data using three different modes of data collection: 1) paper self-administered questionnaires (mail); 2) computer-assisted telephone interviews (CATI); and 3) self-administered online surveys via the World Wide Web (Web). Until the 2003 survey cycle, SDR data were collected by first mailing paper questionnaires to sample persons and then following up the nonrespondents with CATI. The tri-mode data collection effort including mail, CATI and Web was tested in the 2003 SDR and fully implemented in the 2006 SDR. The 2008 survey cycle will be the second round of a fully implemented tri-mode data collection protocol.

During the 2003 SDR, the survey launched a beta Web version and conducted experiments on the efficacy of using the Web and CATI modes as the start mode of data collection. The experiment group sizes were relatively small because NSF wanted to ensure that using the Web and CATI as the start mode of data collection would result in high quality data, reasonable production costs and respondent satisfaction. Additionally, during the 2003 data collection effort, respondents in all modes were asked to state their mode preference for completing the survey in future rounds.

In its initial rollout in 2003, the Web survey met with very positive reactions. A number of sample members, initially asked to complete the survey in either the CATI or mail mode, completed the survey on Web. Further, of the respondents that answered the mode preference question, 49.0 percent indicated a preference for the Web survey mode.

While the Web mode is the most efficient with regard to cost, NSF wanted to verify that the data quality of this very important survey was not compromised by introducing the Web mode option. Careful quality analysis was conducted at the unit and item level. Resulting analysis showed that the Web mode showed higher response rates as well as more complete survey and contacting data than the mail mode. However, the analysis also showed that the Web mode data required more

post-processing machine editing.<sup>2</sup> Overall, the data obtained from the Web mode was considered to be higher in quality than the data obtained from the mail mode.

Thus, in 2006 SDR, the Web survey was offered as a start mode option to panel members who responded in 2003 SDR and indicated their preference for the Web survey. The mode preference question introduced in 2003 was retained in 2006. Due to the increase in sample members offered the Web mode, the 2006 SDR showed a large increase in Web completes. Further, many sample members that started in mail and CATI modes also decided to respond in the Web. Approximately 47 percent of the 2006 participants completed a Web survey and 57 percent providing a response to the future mode preference question indicated a preference for the Web. Based on the results of honoring mode preference in 2006 SDR, the 2008 SDR will also honor mode preference and it is expected that 50 percent or more of the survey response will be completed in the Web mode.<sup>3</sup>

As mentioned previously, the 2008 data collection effort will be conducted by NORC. They will utilize a proprietary case management software system to effectively manage the 2008 SDR sample in these various modes. NORC will use mrInterview, a core product of the Dimensions family of SPSS, in implementing the computer-assisted data entry (CADE), CATI with a telephone number management system (TNMS) incorporated in the case management system, and Web instruments. By using one software platform, data collection from multiple modes of data collection can easily be integrated and delivered. Optical scanning will be used to capture the digital images of the mail questionnaire after keying. The images will be stored in a database for archival purposes.

#### **4. Efforts to Identify Duplication**

Duplication does not exist. No other data collection is based on a probability sample of all components of the doctoral population in science, engineering and health fields in the United States. Data from the Current Population Survey, the American Community Survey, and Decennial Census provide occupational estimates, but no estimates by degree field. The 2008 survey is necessary to obtain trend data on education and career trends of the Nation's doctoral-level scientists, engineers and those in health related occupations as well as data that reflect the changing employment patterns.

There is no similar information available on doctoral scientists, engineers and those in health-related occupations that may be used, modified, or made comparable to the SDR.

#### **5. Efforts to Minimize Burden on Small Business**

Not applicable. The SDR collects information from individuals only.

---

<sup>2</sup>Grigorian, K. and S. Sederstrom, 2005. Qualitative Comparison of Paper and Online Self-Administered Modes. Paper presented at the American Association for Public Opinion Research annual meetings, Miami Beach, FL, May 2005.

<sup>3</sup> Grigorian, K. and Hoffer, T., 2008. "2006 Survey of Doctorate Recipients Mode Assignment Analysis Report." Report prepared for the National Science Foundation under contract number SRS-0208765, Task Order No. B-01, March 2008.



## **6. Consequences of Less Frequent Data Collection**

Because the SDR is a longitudinal survey, conducting the survey less frequently would make it more difficult to locate the persons in the sample because of the mobility of the U.S. population. This would result in both a higher attrition rate as well as less reliable estimates. Also, government, business, industry, and universities would have less recent data to use as a basis for formulating the Nation's SEH policies.

Expanding the time between interviews would also lessen the accuracy of the recall of information by the respondents. This would affect the reliability of the data collected and reduce the quality of the Congressionally mandated biennial reports prepared by the NSF.

Follow-up surveys every two years on the same sampled persons are necessary to track changes in the SEH workforce due to large movements in and out of SEH occupations over both business cycles and life cycles. To make sure of the availability of current national data, the SDR is conducted and coordinated with the National Survey of College Graduates (NSCG) and the National Survey of Recent College Graduates (NSRCG). The degradation of any single component would jeopardize the integrity and value of the entire SESTAT system of surveys and integrated database.

## **7. Special Circumstances**

Not applicable. This data collection does not require any one of the reporting requirements listed.

## **8. Federal Register Announcement and Consultations Outside the Agency**

### ***Federal Register Announcement***

The Federal Register Notice for the SDR appeared on March 7, 2008 (See Attachment 2). NSF received one public comment in response to the announcement as of the closeout date of May 6, 2008. The comment came from S. Bartholomew-Palmer, via email on May 6, 2008, and was included in the comments for a separate federal register notice on another NSF survey. Ms. Parker did not object to the information collection, but wanted to offer her services to make improvements to the survey. NSF believes that because the comment does not pertain to the collection of information on the required forms for which NSF is seeking OMB approval, NSF is proceeding with the clearance request.

### ***Consultations Outside the Agency***

The Division of Science Resources Statistics (SRS) within the NSF has responsibility for the SESTAT surveys. In the early 1990s, SRS initiated and implemented a major redesign of this system of surveys, and continued to adhere closely to the redesigned approaches in conduct of the surveys throughout the decade.

As the SESTAT survey system entered the first decade of the 21<sup>st</sup> century, SRS set a goal to further improve the efficiency and relevancy of the SESTAT system in meeting the data needs of policy makers, academic and research communities and industry. In order to accomplish this goal, SRS carefully planned and engaged in a series of formal and informal evaluations and assessments of each of the three surveys as well as the system as a whole between May 1999 and December 2002.

These evaluations covered several areas: sampling frame, population coverage, sample design, survey content, data system design and data dissemination.

After the redesign efforts, SRS began a more systematic set of activities to encourage greater dissemination of the SESTAT surveys, and to encourage greater use of the data by outside researchers.

### ***Meetings and Workshops on Redesign***

Both internal and external consultation took place through a series of meetings and workshops on various issues related to the SESTAT redesign and survey methodology.

For the 2003 survey round:

- SRS hosted a workshop on possible sources of alternate sampling frames for each of the three SESTAT surveys, including the SDR, to ensure that current frames were still the most efficient and cost effective source for the populations of interest. For this workshop, representatives from the other government agencies and survey firms were invited to provide input on other potential national frames suitable for the SESTAT surveys. This workshop confirmed that there were no alternative and better frames than those used in the 1990s and still available early in the next decade.
- SRS convened an expert content panel to provide overall guidance on the review of the SESTAT questionnaire content and the relevancy of the information collected to meet policy, research and user needs. The content panel was comprised of experts knowledgeable about scientific workforce and education issues, and represented individuals from the private-for-profit industry sector, academia, and non-profit organizations. The content panel met three times (February 2000, May 2000 and June 2002); each of the meetings included invitees from other federal agencies who either collect general workforce data, or use the SESTAT data. A report was issued from each meeting. Feedback from this activity confirmed the importance of the current content of the SESTAT surveys -- the content panel members did not recommend deleting any content -- and provided guidance on new content for the upcoming decade.
- SRS commissioned the Committee on National Statistics (CNSTAT) of the National Research Council (NRC) to examine proposed sample design options for the SESTAT surveys, with a focus on the NSCG. The CNSTAT committee held a two-day workshop on this topic, and issued a report with recommendations to NSF on the 2003 NSCG sample design. The recommendations generally were already reflected in the design plan for the surveys.<sup>4</sup>

---

<sup>4</sup> National Research Council, Committee on National Statistics. 2003. *Improving the Design of the Scientists and Engineers Statistical Data System (SESTAT)*. Washington: The National Academies Press. A copy of this report can be found at the following web address: <http://books.nap.edu/books/0309087112/html/index.html>.

For the 2006 and 2008 survey rounds:

- SRS hosted the SESAT Data collection Contractors Debriefing. All aspects of the SESTAT data collection were discussed to make improvements on the survey procedures.

For the 2008 survey round:

- SRS held a SESTAT Methodology Research Conference. This conference was held to share the results of the methodological research and experiments conducted in the three 2006 SESTAT component surveys.
- SRS held the SESTAT Research Methodology Planning Meeting. This meeting was conducted to discuss possible research experiment ideas for the future SESTAT surveys.

### ***Public Consultations on Redesign***

In addition to these meetings and workshops, SRS conducted a series of consultations with the public.

For the 2003 survey round:

- SRS conducted a series of executive interviews with high-level officials in industry, government and academe to identify the important issues likely to affect the S&E workforce in the future. The interviews were conducted with eight federal officials; one state government official; nine senior staff from private, for-profit firms; six senior staff from non-profit associations or foundations; and one university administrator.
- SRS conducted telephone interviews with eleven academic and nonacademic researchers who had previously used the SESTAT data to determine how well the survey data met their needs, and to obtain feedback on other useful content areas to include on the surveys in the future.
- SRS conducted a series of five focus groups with different sets of users on all aspects of the SESTAT survey data. Two of the focus groups were held online, an innovative approach that was cost saving, as well as providing an immediate transcript. Other focus groups were conducted in a meeting setting. The focus group participants consisted of: 1) Professional association staff who routinely used the SESTAT data as a national comparison of their internal data; 2) Researchers who used the data in the past; 3) Current and past NSF staff who had been data users; 4) Current and past SESTAT contractor staff who were familiar with the data collection and areas where improvements could be made; and 5) Staff from the NRC who also routinely used scientific workforce data.
- SRS commissioned five independent researchers who were familiar with the historical development of NSF's workforce data, and had extensively used the SESTAT microdata, to conduct thorough content reviews and evaluations of the survey data.

### ***Consultations for Outreach and Dissemination***

In order to maintain the currency of the SESTAT surveys and to obtain ongoing input from the public and researchers, SRS has engaged in the following activities.

For the 2008 survey round:

- SRS has convened a Human Resources Experts Panel (HREP) in order to help the Division of Science Resources Statistics (SRS) improve data collection on the science and engineering (S&E) labor force through review and renewal of the S&E personnel surveys and to promote use of the data for research and policy analysis purposes. HREP will accomplish its mission by: 1) Suggesting methods to publicize and promote the data; 2) Providing advice on efforts to improve the timeliness and accuracy of S&E labor force data; 3) Providing a mechanism for obtaining ongoing input from both researchers and policy analysts interested in S&E personnel data; 4) Providing perspectives on the data needs of decision makers; 5) Identifying issues and trends that are important for maintaining the relevance of the data; 6) Identifying ways in which S&E personnel data could be more useful and relevant for analyses; and 7) Proposing ways to enhance the content of the SRS human resources surveys. The panel is made up of 15 members who represent the sciences, academia, business/industry, government, researchers and policy makers. The panel will meet twice a year for 3 years. Two meetings have been held since the panel convened in 2007.
- In addition to researchers and the public who use the public-use SESTAT, SDR, NSRCG or NSCG files, there are also individuals who use the restricted-use files under a license. SRS has now funded two workshops over the past 18 months where a selection of current and potential future licensees meet at NSF to present their research findings to NSF as well as to the broader research community.
- The SESTAT surveys, particularly the SDR, contain a wealth of information on highly-trained individuals in the U.S. labor force. Over the past several years, there has been a great deal of interest in leveraging the survey data that are collected with other information on productivity by some of the same individuals (for example, patenting records or publishing records). In order to pursue the feasibility of this research, SRS funded a workshop at NSF that brought in experts on database matching.
- As part of broader SRS activity on innovation, SRS participated in the 2007 Symposium on Entrepreneurship and Innovation Data in November of 2007. The purpose of the workshop was to present researchers with information on datasets that could advance research in this area.

### **9. Payment or Gifts to Respondents**

Because the SDR is a panel survey that includes both respondents and nonrespondents from one cycle in the next cycle, NSF is concerned about the impact of past refusals in subsequent rounds. In order to test a new approach to gaining early cooperation from these cases as well as to minimize expensive follow-up costs, a \$30 prepaid monetary incentive experiment is planned for 2006

refusals who are in the 2008 sample. The SDR also includes a new cohort of graduates in every survey round. Past research has shown that an early rather than late incentive is effective for these cases. In order to further refine the results of that experiment, a \$30 prepaid monetary incentive will be offered to a sample of new cohort cases at the first mailing; another group will receive a \$30 prepaid incentive at the second mailing. In addition to these incentive experiments, NSF plans to offer a \$30 prepaid monetary incentive to a sample of nonrespondents near the end of data collection to minimize potential nonresponse bias. See sections B.3 and B.4 for details on the incentives.

## **10. Assurance of Confidentiality**

The NSF is fully committed to protecting the anonymity of all survey respondents. SDR data will be collected in conformance with the Privacy Act of 1974, and NSF's authorizing legislation and the Confidential Information Protection and Statistical Efficiency Act of 2002. Cover letters and survey questionnaires to each selected respondent advise them that the information they provide is confidential (see Attachment 3 – Proposed 2008 SDR Mailing Materials and Attachment 4 – Proposed 2008 SDR Questionnaire). The same notice of confidentiality will be used in the introduction to the CATI interview and displayed prior the start of the survey in the Web instrument.

Standard data collection procedures at NORC incorporate numerous safeguards for the data. While collecting SDR data, NORC separates information that could identify a particular sample member from data about that person. Each sample member is assigned a unique identifier, and this identifier is used to store identifying information (such as name, address, etc.) in a separate database from the survey response data.

SDR hard copy questionnaires and other contact materials are housed in a secured storage room at NORC's production facility. Hard copy materials are accessed from the file room only by authorized staff and only when necessary for data collection activities. NORC's electronic systems are on a local area network (LAN). All NORC systems used to store electronic survey data are secure by design and protected by passwords only available to authorized study staff.

NORC will take special steps to ensure that data collected via the Web questionnaire are secure. First, access to the Web instrument is only allowed with a valid Personal Identification Number (PIN) and password correctly entered in combination. Second, data will be transmitted by the Secure Sockets Layer (SSL) protocol that uses powerful encryption during transmission through the Internet. If a respondent keeps a Web survey open without any activity, the Web server at NORC closes it after a short period of inactivity, thus preserving the data up to the break-off point and securely closing the connection. The Web system architecture process has been designed in a way that places authentication information and response data on physically separate servers. This strategy provides an extra layer of security to protect response data. Both development and production servers are backed up nightly, as NORC's disaster recovery plan requires.

All data and analysis are reported in aggregate form only and measures are taken so that the identity of individuals or organizations is not disclosed.

## **11. Justification for Sensitive Questions**

No questions of a sensitive nature are asked in this data collection.

## **12. Estimate of Respondent Burden**

The NSF estimates contact with about 42,600 sample persons by mail, telephone or Web. Based on experience administering the SDR interviews, the questionnaire takes an average of 25 minutes to complete. With three modes of data collection, the targeted overall response rate is 85 percent.

Based on an estimated number of 36,210 respondents, the total burden is calculated at 15,100 hours for the main data collection. Additionally, about 100 burden hours are estimated for future testing of methods to refine data collection to reduce burden and improve utility for the 2010 survey. The total cost to respondents for the 15,200 hour burden is estimated to be \$627,760. This estimate is based on an estimated median annual salary of \$85,900 per SDR respondent. Assuming a 40-hour workweek and a 52-week salary, this annual salary translates to an hourly salary of \$41.30. Salary estimates were obtained using data from the 2006 SDR.

## **13. Cost Burden to Respondents**

Not applicable. This survey does not require respondents to purchase equipment, software or contract out services.

## **14. Cost Burden to the Federal Government**

The total estimated cost to the Government for the 2008 SDR is \$6.0 million for data collection costs, NSF staff costs to provide oversight and coordination with the other two SESTAT surveys, and costs associated with the integration of the SDR data into the SESTAT data system. The cost estimate for the data collection \$5.14 million, which is based on sample size; length of questionnaire; CATI and Web data collection technology; administrative, overhead, design, printing, mail and telephone data collection costs, incentive payments, critical items data retrieval; data keying and editing; data quality control; imputation for missing item responses; weighting and estimating sampling error; file preparation and delivery; preparation of documentation and final reports; analysis, and tabulations. NSF staff costs are estimated at \$417,000, (\$111,104 annual salary of 1.5 FTE for 2.5 years of the 2008 SDR survey cycle). The SESTAT integration costs will be approximately \$400,000 for the 2008 SDR survey cycle.

## **15. Reason for Change in Burden**

The 2008 SDR will include two graduating classes from the academic years 2006 and 2007, as well as a panel component. The 2006 SDR fielded graduates from three academic years 2003, 2004, and 2005. The total sample size in the 2008 will be 40,000, while it was 42,000 in 2006, in order to accommodate three graduating classes instead of two. The change in burden hours from the 2006 SDR reflects the change in sample size and additional burden hours required for future tests of methods to improve survey procedures and utility.

## **16. Schedule for Information Collection and Publication**

The 2008 SDR data will be disseminated in a number of ways. First, NSF will issue an Info Brief, showing highlights of selected data from the 2008 SDR. Second, NSF will publish a report containing a set of about 70 detailed statistical tables from the survey in both hard copy and on the Web. These tables will be descriptive in nature and will provide extensive information on the education and employment of SEH doctorate recipients by major field of study, occupation, and demographic characteristics. Third, NSF will issue a descriptive report profiling the demographic

and employment characteristics of doctorate level scientists and engineers. Also various Info Briefs will be written to address current doctorate labor force issues. Most importantly, the SDR data will be included in the SESTAT data system, which resides on the Internet. The SESTAT system can be used to produce tabulations from the component surveys, providing a rich resource to those within and outside the government. The NSF will use SDR data in the development of key NSF reports, including the Congressionally mandated reports *Women, Minorities and Persons with Disabilities in Science and Engineering* and *Science and Engineering Indicators*.

The time schedule for information collection and publication is currently estimated as follows:

<b>2008 SDR Milestone Task</b>	<b>Start Date</b>
Data Collection (mail, CATI, Web)	October 2008-June 2009
Coding and data editing	October 2008-November 2009
Final edited/weighted/imputed data file to NSF	December 2009
SDR InfoBrief	Spring 2010
SDR Detailed Statistical Tables	Summer 2010
SDR Public Use File	Summer/Fall 2010

#### **17. Display of OMB Expiration Date**

The OMB Expiration Date will be displayed on the 2008 SDR questionnaire.

#### **18. Exception to the Certification Statement**

Not Applicable.

## **B. Collection Of Information Employing Statistical Methods**

### **1. Respondent Universe and Sampling Methods**

The population for the 2008 SDR will be selected from approximately 795,000 individuals from the Doctorate Records File (DRF) which is a census of research doctorates awarded from U.S. institutions since 1920. The DRF is compiled through the annual Survey of Earned Doctorates (SED).

The sample design for 2008 will be consistent with the sample redesign developed and implemented in 2003 SDR and retained in 2006 SDR. To be eligible for the 2008 SDR target population, respondents have to:

- 1) receive a doctoral degree in science, engineering or health from U.S. institutions between 1958 and the academic year 2007;
- 2) indicate on the SED their plan to stay in U.S. after receiving their doctorate degree;
- 3) be under age 76;
- 4) be living in the U.S. as of October 1, 2008 (new survey reference date).

For 2008, a sample will be selected from the new 2006-2007 doctoral cohort groups and added to the longitudinal sample (which covered graduates through 2005) that is conveyed from cycle to cycle. To offset this new cohort addition and to limit the overall sample size, a maintenance cut will be performed on the longitudinal sample.

There are two types of SEH doctorate recipients who have been excluded from the eligible SDR sample frame: 1) non-U.S. citizens who reported plans in the SED to leave the U.S. after earning their science, engineering or health doctorate were considered permanently ineligible for SDR sample selection; 2) non-U.S. citizens who were selected into the SDR but who had been found to reside outside of the U.S. for two or more survey cycles were also considered permanently ineligible for the SDR. In both cases, it is possible that the individuals assumed to reside outside the U.S. were actually living in the U.S. on the survey reference date and should thus have been classified as eligible. These additional cases will be asked to participate in the 2008 SDR as a separate subsample group, named the International Survey of Doctorate Recipients (ISDR), thus making the overall SDR sample more inclusive and representative. The sample size for the ISDR is approximately 2,600. These cases will be fielded with the 40,000 cases included in the 2008 SDR sample and subject to the same sampling procedures, data collection protocol and data processing treatment.

The targeted overall weighted response rate on the 2008 SDR is 85 percent. The plan for maximizing the response rate is presented in Section 3.

### **2. Statistical Procedures**

As mentioned in the previous section, the 2008 sample design will be consistent with the 2003 and 2006 SDR sample designs. Stratification variables for the sample include: demographic group, field of doctorate, and sex. The demographic group is a composite variable recording disability status, race/ethnicity, and citizenship at birth (U.S. or foreign).

The 2008 SDR sample will be selected using sampling strata based on a multi-way cross of the stratification variables. (See Attachment 5 – 2008 SDR/ISDR Sample Strata and Sample



Allocation Table). For 2008, a sample will be selected from the new 2006-2007 doctoral cohort groups and added to the longitudinal sample that is conveyed from year to year. To offset this new cohort addition and to limit the overall sample size, a maintenance cut will be performed on the longitudinal sample. The SDR sample size and sample design ensure NSF will maintain the ability to produce the small demographic/degree field estimates that are needed for the Congressionally mandated report on *Women, Minorities and Persons with Disabilities in Science and Engineering* (See 42. U.S.C., 1885d). The 2008 ISDR sample is, like the SDR, drawn from the DRF. However, the stratification, with ten strata defined by race/ethnicity and gender, is simpler than that for the SDR. Within each stratum, the frame will be sorted by degree field for implicit stratification prior to systematic selection.

Estimates from the 2008 SDR/ISDR will be based on standard weighting procedures. As was the case with sample selection, the weighting adjustments will occur separately for cases for the old cohort and new cohorts. Each case will have a base weight defined as the probability of selection into the 2008 SDR/ISDR sample. This base weight will reflect the differential sampling across strata. For the old cohorts, the base weight will be equal to the final weight from the previous survey cycle. The final analysis weights will be calculated in three stages:

- 1) First, a base weight will be calculated for every case in the sample to account for its selection probability under the sample design.
- 2) Second, an adjustment for unknown eligibility will be made to the base weight by distributing the weight of the unknown eligibility cases to the known eligibility cases proportionately to the observed eligibility rate within each adjustment class.
- 3) Third, an adjustment for nonresponse will be made to the adjusted base weight to account for the eligible sample cases for which no response was obtained.

*Replicate Weights.* A set of replicate weights based on the Balanced Repeated Replication (BRR) method will also be constructed. The entire weighting process applied to the full sample will be applied separately to each of the replicates to produce a set of replicate weights for each record.

*Standard Errors.* The BRR method will be used to estimate the standard errors of the 2008 SDR estimates as in the past. The variance of a survey estimate based on any probability sample may be estimated by the method of replication. This method requires that the sample selection, the collection of data, and the estimation procedures be independently carried through (replicated) several times. The dispersion of the resulting estimates then can be used to measure the variance of the full sample.

### **3. Methods to Maximize Response**

#### ***Maximizing Response Rates***

The weighted response rate for the 2006 SDR was 79 percent. Extensive locating efforts, follow-up survey procedures and targeted data collection protocols will be used to maximize the survey response rate to maintain at least an approximately 80 percent response rate and to target an 85 percent response rate in 2008. Additionally, monetary incentives are also being planned, building on experiments conducted in the 2003 and 2006 rounds.

### ***Locating***

The contact information obtained from the 2006 SDR and from the 2006 and 2007 SED surveys for the sample members as well as the people who are likely to know the whereabouts of the sample members will be used to locate the sample members in 2008.

The U.S. Postal Service's (USPS) automated National Change of Address (NCOA) database will be used to update addresses for the sample. The NCOA incorporates all change of name/address orders submitted to the USPS nationwide, which is updated at least biweekly. Vendors also maintain up to 36-month historical records of previous address changes. It will also be used to track persons who have moved from their previous address at the time of 2008 survey. The names and addresses of mail nonrespondents will be matched to the most recent NCOA address updates with a vendor who appends telephone numbers.

The locating efforts will also utilize a specially-trained locating team who has proven themselves successful at searching for and finding nonrespondents with problem addresses or telephone numbers. Their locating strategy will include contacting employers, educational institutions and alumni associations, online publication searches, change of address searches, and Directory Assistance and administrative record searches. In addition to last known address, locators have past contacting information available as far back as 2001. Locators will also have access to contact names and addresses given by respondents in past survey rounds, where available. An automated commercial telephone number matching service and the national death registry will also be used.

### ***Data Collection Strategies***

As described above, NORC will continue to incorporate the Web mode in the data collection protocol to improve both data completeness and sample members' satisfaction.

A core set of contact materials (Prenotice Letter, Thank You/Reminder Postcard, and Cover Letters accompanying the SAQ) will be used in mailing to the SDR sample members. These contact materials will be tailored to address the particular issues or concerns of the sample groups to whom they are targeted. Tailoring will be based on cohort (2006 Panel member versus new cohort), response in the past round, citizenship, and expressed mode preference. NORC will also utilize email versions of the contacting materials for sample members with email addresses on file.

NORC will conduct extensive CATI follow-ups for those sample members who do not submit a completed questionnaire via a paper or Web form. The CATI Interviewing team will include Refusal Avoidance and Conversion specialists who have a proven ability to work with sample members to obtain their consent and participation.

### ***Incentive Plan for 2008***

In 2003 and 2006 SDR, multiple incentive experiments were conducted. In 2003, the experiments showed that a monetary incentive significantly improved response at the end of the field period compared to a non-monetary incentive or no incentive. Further, the response to a lower value prepaid incentive was better, though not significantly, than a higher value postpaid incentive. In 2006, the incentive experiment results show that the prepaid incentive had the greatest effect on response for the new cohort with least follow-up cost when offered earlier. The second greatest effect of a prepaid incentive on response was for those that cooperated in the prior survey cycle, and that the response effect was still significantly better than no incentive. It was also found that

the incentive cost was less expensive for the past cooperative respondents when it was offered later in the field period.

These results obtained in the 2003 and 2006 SDR incentive experiments will be applied to the 2008 SDR data collection protocol as part of a gaining cooperation contacting strategy. A \$30 prepaid incentive will be offered as required to nonrespondents who are non-locating problems towards the end of the field period to gain cooperation and boost the response rate. The cases selected to receive the incentive will be determined after the 2008 incentive experiments and standard data collection protocols have been implemented.

The overall survey response rates and the number of respondents at three months prior to the end of the field period will be analyzed by sampling cell. The overall strategy would be to give all nonrespondents a probability of receiving an incentive. A greater probability of selection for receiving an incentive will be given to cases in those cells where the number of completed cases is low, in order to improve the survey estimates.

We will examine each sampling cell, and apply the following allocation formula to determine the number of cases in each sampling cell that will receive the incentive.

$$n_h = n * \frac{\frac{p_h^a}{r_h^b}}{\sum_{h=1}^H \frac{p_h^a}{r_h^b}}, \text{ here } a \geq 0, b \geq 0.$$

where:

- $n_h$  = is the number of incentives allocated to sampling stratum  $h$
- $n$  = number of incentives available (based on budget) for this survey
- $p_h$  = nonresponse rate for stratum  $h$  before the incentives are issued
- $r_h$  = number of respondents before the incentives are issued
- $H$  = total number of sampling strata
- $a, b$  = parameters used to determine whether  $p_h$  or  $r_h$  should be given more weight.

NSF will ensure that all nonrespondents have a 0.25 up to 1.0 probability of receiving the incentive. The final determination of the incentive distribution will be made in consultation with the SRS Chief Statistician. The incentive will be sent via first class U.S. mail, with a cover letter and a paper questionnaire. The date of this incentive mailing would likely be in mid-February 2009.

#### 4. Testing of Procedures

Because data from all three SESTAT surveys are combined into a unified data system, the surveys must be closely coordinated to provide comparable data from each survey. Most questionnaire items in the three surveys are the same.

Although there will be no new questions in the 2008 SDR questionnaire, all content items in the SESTAT questionnaires have undergone an extensive review and testing before they were included

in the final version. The changes made in the questionnaires are a result of a variety of activities that included extensive review of the entire content in each of the SESTAT survey questionnaires and additional research on specific items to provide more information before a final decision was made on placement and wording of the item in the questionnaires. Content evaluation and testing activities for the 2003 and 2006 surveys included:

- External and internal consultation with questionnaire design experts on questionnaire layout and formatting to improve user-friendliness and minimize respondent reporting errors;
- External consultation on improving the messages in the survey contact materials; and
- A two-stage pretest of the survey questionnaires consisting of mail and telephone.

All of these activities contributed to the development of the questions in the 2008 SDR questionnaire.

### ***Survey Questionnaire Review and Research***

The SESTAT survey questionnaire items are divided into two types of questions: core and module. Core questions are defined as those considered to be the base for all three SESTAT surveys. These items are essential for sampling, respondent verification, basic labor force information, and/or robust analyses of the science and engineering workforce in the SESTAT integrated data system. They are asked of all respondents each time they are surveyed, as appropriate, to establish the baseline data and to update the respondents' labor force status and changes in employment and other demographic characteristics. Module items are defined as special topics that are asked less frequently on a rotational basis of the entire target population or some subset thereof. Module items tend to provide the data needed to satisfy specific policy, research or data user needs.

After identifying the core and module items that would be included in the SESTAT surveys, SRS reviewed and identified content items needing improvement, and engaged in research to craft new questions. SRS conducted separate studies on six core items, and one study on a module for the 2003 survey questionnaires. The core item research covered the following topics on the SESTAT questionnaires: employer's main business, academic positions, academic institutions, work activities, marital status, and degrees earned abroad. Based on the external consultations (See Section A.8), a study was conducted to develop a module to capture more information on postdoctoral employment histories in the SDR, which was included on the 2006 SDR.

The core item research resulted in some wording changes to those questions on the SESTAT questionnaires, and a revision of how the occupation code frame is presented. The module research led to the addition of a series of questions on postdoctoral employment for up to three postdoctoral positions in the 2006 SDR questionnaire. The 2008 SDR questionnaire will not include new questions not previously fielded before.

For 2008, the SDR questionnaire content will be revised from 2006 as follows:

- Survey reference date changed from April 1, 2006 to October 1, 2008.
- Removed a 2006 module on collaborative activities (it has not yet been decided if this will be rotated back in at a future time).
- Rotated out a module on postdoctoral history, which was asked in 1995 and 2006.
- Rotated in a module on second job (status, job description, job category, relatedness of second job to highest degree), which was asked in 1993 to 2001.

- Rotated in a module on the respondent's and spouse's areas of technical expertise, which was asked in 1993 to 2003.

A complete list of questions proposed to be added, dropped, or modified in the 2008 SDR questionnaire is included in Attachment 6.

The 2008 SDR questionnaire retains all content changes that were tested implemented in the 2006 SESTAT questionnaires. In 2005, SRS conducted an extensive pretest under a generic clearance (OMB No. 3145-0174) that consisted of two phases: (1) two rounds of in-depth cognitive interviews, and (2) a small-scale field test of the mail questionnaires.

#### Pretest Phase I – Cognitive interviews

Mathematica Policy Research, Inc. (MPR) and the U.S. Census Bureau (Survey Research Division) were contracted to conduct in-depth cognitive interviews on the 2006 SDR and the other two SESTAT survey questionnaires. Cognitive interviews were conducted in two waves, with the waves being scheduled during the same time period at MPR and the Census Bureau. MPR tested the full-length questionnaires for the three surveys, while the Census Bureau was asked to focus on the employment section of the NSCG (which is the same as is used in the SDR). In addition to the questionnaires, the cognitive interviews were also used to test improvements to the cover letters for the 2006 survey administration.

The first round of cognitive interviews was conducted between February 2 and February 25, 2005. During this period MPR and Census Bureau each interviewed 30 respondents. The second round of cognitive interviews was conducted between March 25 and May 2, 2005. MPR interviewed 40 respondents (28 in-person and 12 via telephone) and the Census Bureau interviewed 30 respondents. Based on the results of the cognitive interviews, MPR and NSF worked together to develop a series of experiments to test in the mail portion of the pretest.

#### Pretest Phase II – Mail Field Test

The field test consisted of two mailings of SDR and the other two SESTAT surveys with a reminder postcard in between; no further nonresponse follow-up was conducted due to time constraints. The NSCG mail pretest included a sample of 1,500 selected from a commercial list of 5,000 names of bachelor's degree holders with address, sex, age, and occupation information, and between the ages of 21 and 75. To mimic the proportion of science and engineering cases from the 1995 NSCG, MPR selected 15 percent of the cases from computer occupations, 20 percent from engineering occupations, and 65 percent from other occupations for a total of 1,500 sample members. Each sample member was randomly assigned to one of four control or experimental groups.

Pretest questionnaires were mailed on June 24, 2005 using first class mail. Although mailing a reminder was not part of the original pretest plan, a postcard reminder was sent to all non-respondents because of the low response (12 percent) to the first mailing. The postcard was mailed on July 20, 2005, and provided an additional boost of about 2 percentage points to the response rate for a 14 percent cumulative overall response rate from all three SESTAT surveys to the first mailing. A second mailing was sent on August 3, 2005 with a cover letter urging participation with a "respond by" date in a Priority Mail envelope. Mail returns were accepted until August 26, 2005. Final response rate to the NSCG mail pretest was about 25%. Final response rate for respondents from all three surveys was 27 percent.

The primary goal of the field pretest was to test the various recommended questionnaire changes from the cognitive interviews. Specific test conditions were incorporated to obtain research data that might further improve the questionnaires. These are described below:

- 1) Testing the placement of the sample person’s name and address label on the questionnaire (front versus back cover).
- 2) Testing the Field of Study and Job Category Code Lists in a new format.
- 3) Testing a different approach to “anchoring” the reference date in the employment questions.
- 4) Testing a new wording and format of the principal employer type question.

In addition, the experimental versions of the questionnaires had small wording and formatting changes for some questions of interest such as work activity categories, employer name and location, supervising, etc. The control versions of the questionnaire retained the same wording for most questions of interest and Field of Study/Job Category Code Lists used in 2003. Testing the label placement by the presence versus absence of the content changes created a two-by-two design, shown in the table below.

Mail Pretest Design			
		Content, Anchor, and Code List	
		Old Content (Control)	New Content (Experimental)
Address Label	Back	Questionnaire Version 1	Questionnaire Version 3
	Front	Questionnaire Version 2	Questionnaire Version 4

The mail pretest also included testing of a new 2006 module on the method and means of collaboration; using “Yes/No” response options in a few remaining questions with the “Mark All That Apply” response options used in 2003; moving the part-time employment questions to a different section and revising the work-related training reasons to fine tune the measurement of the concepts for these two items.

Based on the mail pretest results, decisions were made to keep the sample person’s name and address labels on the front cover of the questionnaire; use the revised wording and format of the employer sector question; use the new Field of Study/Job Category Code Lists; no longer use the ‘Mark All That Apply’ response option; not use the reference week “anchoring” question but use consistent question wording in all references to the principal job.

***Survey Contact Materials***

Survey contacting materials will be tailored to best fit sample members need for information about the SDR and gain their cooperation. Materials requesting sample member participation via the Web survey will include access to the survey online. As was done in 2003 and 2006 SDR, NSF and NORC will develop 2008 SDR letterhead stationery that includes project and NSF website information, and NORC’s project toll-free telephone line, USPS and email addresses. Additionally, the stationery will contain a watermark that shows the survey’s logo to help brand the communication for sample members for ease of recognition.

***Questionnaire Layout***

SRS has previously engaged the services of Dr. Don Dillman to further improve the visual presentation of the 2003 and 2006 SESTAT questionnaires. An SRS staff member with expertise in visual design theory was also involved in this process. The suggested revisions to the questionnaires included the standardization and consistent use of formatting, placement of instructions, and placement of privacy act notices. Also revised from previous versions were the items that include a format that requires the respondent to review a long list of items before reporting a response to make the selection process easier for the respondents.

### ***Web-Based Survey Instrument***

Because of technological improvements and the wide proliferation of Internet users, offering a Web option to SDR respondents has become both feasible and desirable. The Web mode has the potential to become a valuable asset to the survey with regard to decreased cost and enhanced respondent satisfaction. In the 2003 SDR, this new mode was carefully introduced to avoid having a negative impact on the response rate or the high data quality that the SDR project has realized over the years.

The 2008 SDR will maintain the same functionality and software design as used in the 2003 and 2006 survey rounds. However, due to questionnaire changes, it will be necessary to recode some portions of the instrument. This development will take place during Summer 2008, and full testing of the reinstated questions as well as the entire instrument will be completed during August and September 2008.

### ***2006 SDR Survey Methodology Tests***

#### **Contacting Experiments Analysis**

This report details the three contacting experiments that were implemented during the 2006 SDR. The three experiments included in the report are 1) the Brochure Experiment, 2) the Cover Letter Experiment, and 3) the Endorsement Letter Experiment. Each of these is briefly detailed below.

#### **A. The Brochure Experiment**

The Brochure Experiment was developed to help determine the most effective means of gaining cooperation of new cohort sample members. In the 2006 SDR, the new cohort sample consisted of the three most recent SED cohorts, 2003, 2004 and 2005. Because the 2005 SED cohort was not available until after the start of the 2006 SDR field period, the new cohort sample was selected and fielded in two stages. New cohort cases sampled from 2003 and 2004 SED were available at the start of the data collection effort, and these cases are referred to as the first stage new cohort sample. This contacting experiment was conducted on the first stage new cohort sample so that the results could be used to help inform the best way to contact the second stage new cohort sample. In the past, the SDR sent a Frequently Asked Questions (FAQ) brochure to all new cohort sample members with an advance letter at the start of data collection. The FAQ brochure is a tri-fold brochure that addresses sample members' concerns about survey participation. Thus, the FAQ brochure with an advance letter served as the control treatment for this experiment. The two treatments were 1) including a Flyer brochure, which was a shorter, less detailed brochure with the same advance letter instead of the FAQ, and 2) sending the advance letter without any type of brochure.

*Summary results:* Excluding the FAQ from the new cohort mailing, treatment 2, had a significantly positive effect on the response rate.

#### B. The Cover Letter Experiment

The purpose of the Cover Letter Experiment was to test whether sample members who previously refused to participate in the survey would respond better to different versions of the questionnaire cover letter. More specifically, we wanted to see whether this group of past refusals would respond better to cover letters with an “authoritative” appeal or an “altruistic” appeal. The SDR traditionally utilizes an altruistic appeal in its letters and NORC was interested to see whether a firmer tone would be more effective in persuading past-refusals to participate in the survey.

*Summary results:* The response from the group receiving the “authoritative” letter was slightly worse than the “altruistic” letter, but the difference was not significant.

#### C. The Endorsement Letter Experiment

The Endorsement Letter Experiment sought to increase response to a questionnaire mailing sent to all nonrespondents to the initial 2006 contact (whether by mail, CATI or web). The endorsement letters were included in a questionnaire mailing along with a cover letter. The endorsement letters were from 10 different professional organizations encouraging sample members in their particular field to participate in the 2006 SDR (e.g., an endorsement letter from the American Psychological Association was sent to psychology doctorates). The results of similar mailings in the 2003 round, when no endorsement letters were sent, were used as a control for this experimental treatment.

*Summary results:* Including an endorsement letter with the SDR mailings had a significantly negative impact on response.

#### Mode Assignment Analysis

The 2003 SDR included a starting mode experiment and the questionnaire included a mode preference question. In 2006 SDR, three different data collection modes were available at the start of data collection. The three different starting modes were 1) a paper self-administered questionnaire sent in the mail (SAQ), 2) a computer-assisted telephone interview (CATI), and 3) a self-administered online survey (Web). Using mode preference information reported during in the 2003 SDR and response information from the 2003 SDR mode experiments, the 2006 selected sample was assigned to various starting mode data collection protocols. Old cohort sample members who responded to the 2003 SDR were stratified by explicit (their stated preference) or implicit (if no stated preference, the mode by which they responded) mode preference, and the cases were assigned to start mode accordingly. Explicit responses were determined by the answer to the mode preference question on the 2003 SDR survey; for those that did not respond to the preference question or indicated no preference, implicit preference was defined as the mode they used to complete the 2003 SDR. 2003 SDR non-respondents were assigned a starting mode based on analysis conducted on the 2003 data which indicated that past refusals are more likely to cooperate if started in the SAQ mode and other non-respondents were most likely to cooperate if started in the Web mode. All new cohort members were assigned to the CATI mode; this decision was also based on analysis conducted on the 2003 SDR data. Those sample members that were living abroad and who had not completed the 2003 SDR were started in the Web mode to decrease mailing costs for sample members most likely to be ineligible for the 2003 SDR. Those without any physical or e-mail address were started in CATI.



The Mode Assignment Analysis report documents the results of the 2006 SDR starting mode assignments. The 2006 SDR results were compared to the results from the 2003 SDR at the case level for the panel, and in the aggregate for the panel nonrespondents and the new cohort. The outcomes of interest include response rates, level-of-effort indicators, response time intervals, and data quality measures. Analysis examined these outcomes by demographic variables of sex, citizenship, doctorate field, ethnicity/race and age, and also by locating status.

*Summary results:* Assessment of the 2006 SDR mode assignment for particular groups revealed the following about the 2006 data collection approach:

- Honoring explicit and implicit mode preference was an effective strategy for the 2006 SDR. While following this strategy did not affect the response rate or the number of contacts required to achieve that response, it did improve both the time to respond and the quality of the data provided.
- Assigning panel members who refused to cooperate in the prior survey cycle to the mail starting mode kept the response from this type of case consistent from 2003 to 2006. However, it required a greater level of effort to maintain the response rate for this group of cases. And while the response rate is generally low for this group, the data provided by panel members who were converted was of a higher quality in 2006.
- Locating problem cases were assigned to the Web starting mode in 2006 SDR. This strategy appears to be the most effective for yielding a positive response with a lower level of contacting effort, based on an analysis of the number of times these cases needed to be contacted.
- Ineligibles and other nonresponse panel cases from the 2003 SDR were also assigned to the Web starting mode in 2006 SDR. The 2006 SDR data collection results for this small group of cases are less clear. The response rate improved for the ineligible cases in 2006, and remained static for the other nonresponse cases. However, the level of effort required to achieve these results increased considerably. While this is understandable for the ineligible cases that are largely emigrants that have returned to the U.S., it is not clear why it should increase for other nonresponse cases. Potentially, other nonresponse cases would respond more readily in another mode.
- Finally, new cohort cases were assigned to the CATI starting mode protocol in 2006 SDR. This approach worked well for the new cohort cases that were missing sampling stratification variables from SED, but did not appear to work as well for the new cohort cases with complete stratification variables. While the time to respond was decreased for the new cohort cases overall, unweighted response rates dropped slightly and item nonresponse increased.

### Incentive Experiment Analysis

In the current environment of declining response rates, many survey researchers have begun to use incentives to increase response rate. In the 2003 SDR, a late-stage data collection experiment showed that offering a pre-paid incentive not only significantly increased response, but also yielded significantly higher quality data. In the 2006 SDR, the research team implemented a follow-up controlled experiment to determine the most efficacious time to offer a pre-paid \$25 incentive to non-respondents late in the field period after the full protocol of contacting attempts had been executed. The incentive experiment design included four different sample groups which were selected on September 12, 2006. At that time the main SDR sample had achieved an unweighted response rate of 61.7%. The incentive experiment groups were selected and identified in the following way:

- Early Control – 500 cases were sent a gaining cooperation letter and email message on September 22, 2006 and followed up with a telephone call approximately one week later.
- Early Incentive – 5,000 cases sent a \$25 pre-paid check and gaining cooperation and email message on September 22, 2006 and followed up with a telephone call approximately one week later.
- Late Control – 500 cases were selected; of these 433 remained pending on October 17, 2006. The remaining pending cases on October 23 were sent a gaining cooperation letter and email message on October 23, 2006 and followed up with a telephone call approximately one week later.
- Late Incentive – 2,600 cases were selected; of these 2,217 remained pending on October 17, 2006. The remaining pending cases were sent a \$25 pre-paid check and gaining cooperation letter and email message on October 23, 2006 and followed up with a telephone call approximately one week later.

*Summary results:* Those sample members receiving incentives, regardless of Early or Late, had higher completion, cooperation, and response rates, than those not receiving the money. While the Late groups caught up to the Early groups, their response came later in the field period. Concerning cost, the incentive experiment supported the 2003 finding that pre-paid incentives are a cost effective gaining cooperation strategy. Few sample members cashed their incentive check without completing a survey.

### Web Screen Experiments Analysis

One challenge of Web questionnaires is the presentation of long lists of response options, particularly those that cannot be fit well within the confines of a single computer screen. One such problematic question in the SDR is the work activity question. In the 2003 SDR Web questionnaire, this item was presented on a single page but respondents were obliged to scroll down the screen in order to view all response options. The 2003 data indicated that the scroll-down requirement may have affected responses, as evidenced in the relatively low frequencies for the response options in the middle of the range (which may have been skipped over too quickly in some Web interfaces) among the Web respondents compared to the paper and CATI respondents. This experiment examines the effect of two different presentations of the work activities question in the 2006 SDR Web instrument. The cases assigned to the Web starting mode data collection protocol were scientifically assigned to a treatment or control group. The treatment group saw a compact preview screen before the work activity question that summarized all fourteen work activity response options on a single screen before being presented with the work activity question

in the scroll-down format. The control group did not have a preview screen, but followed the same protocol used in 2003 and went directly to the work activity question. In addition to the comparisons between the treatment and control group cases that completed the Web version of the survey, we compare the results for the work activity question from respondents who completed the paper version of the survey to the Web respondents.

*Summary results:* The most important lesson the 2006 SDR web preview screen experiment results offer is that SDR respondents can effectively navigate the standard Web questionnaire without mode effect in long list questions. And a preview screen listing all options on a single screen before a long list question does not seem to have a positive effect on response frequency regarding the overall number of responses.

### ***Survey Methodology Tests to be Undertaken***

As described in Section A, NSF proposes two experiments to examine the impact of the timing of incentives in order to assess the impact on response rates and on costs. In the first experiment (“Past Refusal Incentive Experiment”), the 2006 SDR refusals in the 2008 SDR will be split into two groups; the first will receive a \$30 prepaid incentive at the first questionnaire mailing, and the second group would receive no incentive. This experiment is designed to determine whether offering an incentive to past refusals early in the survey cycle will improve response rates, as well as save expensive follow-up costs. The second incentive experiment (“New Cohort Incentive Experiment”) will build on past research that shows that offering an early rather than late incentive to these new panel members is effective in increasing response rate. The 2008 experiment will build on the prior experiment to try and determine an effective early time point to offer the incentive. The new cohort cases in the 2008 SDR will be divided into two groups: the first will be offered a \$30 prepaid monetary incentive at the first mailing, and the offer of the incentive will be repeated in the second mailing; the second group will receive no incentive at the first mailing, but will be offered a \$30 prepaid monetary incentive at the second mailing. The purpose of this experiment is to determine when the early incentive is most effective for improving response rate, and for reducing follow-up costs.

In addition to the two incentive experiments described above, NSF also plans to conduct a third experiment on the data collection strategy involving panel members. The 2006 SDR included a question that asked respondents to indicate their preferred response mode (e.g. mail, telephone, web or no preference). In general, the data collection strategy for the SDR is to honor the preference the respondents provided to the mode preference question, and if there is no preference, to assign them the mode in which they actually completed the survey. For those 2006 cases that indicated no preference, but completed the 2006 survey in the CATI or mail modes, we would like to assess whether first offering these cases the opportunity to complete in web will yield comparable or better response rates compared to assigning them to CATI or mail. These cases will be divided into two groups: the first will be assigned the web start mode, and the second will be assigned the mode they completed in 2006 (either CATI or mail).

Details on the SDR experiment plans are in Attachment #7. NSF plans to conduct additional methodological tests in the current and future rounds of the survey to reduce burden and increase utility of the survey under the burden hours in this survey clearance for the next survey cycle. Proposals for these additional tests are still under consideration. These will be submitted for OMB approval prior to implementation.

## **5. Contacts for Statistical Aspects of Data Collection**

SRS Chief Statistician, Stephen Cohen, has overall responsibility of statistical aspects of the survey. Consultation on statistical aspects of sample design was provided by Brenda Cox, (703-875-2983, Senior Staff, Battelle) and Rachel Harter, (312-759-4025, Statistics and Methods Vice President, NORC). At NSF the contacts for statistical aspects of data collection are Nirmala Kannankutty (703-292-7797, SDR Project Manager) and Stephen Cohen (703-292-7769, SRS Chief Statistician).

**Attachment 1**

National Science Foundation Act of 1950

**SECTION I**  
**NATIONAL SCIENCE FOUNDATION ACT OF 1950**

FUNCTIONS (42 U.S.C. §1862)

§ 1862. Functions

(a) Initiation and support of studies and programs; scholarships; current register of scientific and engineering personnel

The Foundation is authorized and directed—

(1) to initiate and support basic scientific research and programs to strengthen scientific research potential and science education programs at all levels in the mathematical, physical, medical, biological, social, and other sciences, and to initiate and support research fundamental to the engineering process and programs to strengthen engineering research potential and engineering education programs at all levels in the various fields of engineering, by making contracts or other arrangements (including grants, loans, and other forms of assistance) to support such scientific, engineering, and educational activities and to appraise the impact of research upon industrial development and upon the general welfare;

(2) to award, as provided in section 1869 of this title, scholarships and graduate fellowships for study and research in the sciences or in engineering;

(3) to foster the interchange of scientific and engineering information among scientists and engineers in the United States and foreign countries;

(4) to foster and support the development and use of computer and other scientific and engineering methods and technologies, primarily for research and education in the sciences and engineering;

(5) to evaluate the status and needs of the various sciences and fields of engineering as evidenced by programs, projects, and studies undertaken by agencies of the Federal Government, by individuals, and by public and private research groups, employing by grant or contract such consulting services as it may deem necessary for the purpose of such evaluations; and to take into consideration the results of such evaluations in correlating the research and educational programs undertaken or supported by the Foundation with programs, projects, and studies undertaken by agencies of the Federal Government, by individuals, and by public and private research groups;

(6) to provide a central clearinghouse for the collection, interpretation, and analysis of data on scientific and engineering resources and to provide a source of information for policy formulation by other agencies of the Federal Government;

(7) to initiate and maintain a program for the determination of the total amount of money for scientific and engineering research, including money allocated for the construction of the facilities wherein such

research is conducted, received by each educational institution and appropriate nonprofit organization in the United States, by grant, contract, or other arrangement from agencies of the Federal Government, and to report annually thereon to the President and the Congress; and

(8) to take a leading role in fostering and supporting research and education activities to improve the security of networked information systems.

#### **BIENNIAL REPORT (42 U.S.C. §1885d)**

##### § 1885d. Biennial reports

(a) By January 30, 1982, and biennially thereafter, the Director shall simultaneously transmit a report to the Congress, the Attorney General, the Director of the Office of Science and Technology Policy, the Chairman of the Equal Employment Opportunity Commission, the Director of the Office of Personnel Management, the Secretary of Labor, the Secretary of Education, and the Secretary of Health and Human Services.

(b) The report required by subsection (a) of this section shall contain—

(1) an accounting and comparison, by sex, race, and ethnic group and by discipline, of the participation of women and men in scientific and engineering positions, including—

- (A) the number of individuals in permanent and temporary and in full-time and part-time scientific and engineering positions by appropriate level or similar category;
- (B) the average salary of individuals in such scientific and engineering positions;
- (C) the number and type of promotional opportunities realized by individuals in such scientific and engineering positions;
- (D) the number of individuals serving as principal investigators in federally conducted or federally supported research and development; and
- (E) the unemployment rate of individuals seeking scientific and engineering positions;

(2) an assessment, including quantitative and other data, of the proportion of women and minorities studying scientific and engineering fields, including mathematics and computer skills, at all educational levels; and

(3) such other data, analyses, and evaluations as the Director, acting on the advice of the Committee on Equal Opportunities in Science and Engineering, determines appropriate to carry out the Foundation's functions as well as the policies and programs of sections 1885 to 1885d of this title.

**Attachment 2**

**First Federal Register Notice for 2008 SDR**



et seq.) and Secretary of Labor's Order No. 5-2007 (72 FR 31159).

Signed at Washington, DC, on February 29, 2008.

Edwin G. Foulka, Jr.

Assistant Secretary of Labor for Occupational Safety and Health.

[FR Doc. E8-4478 Filed 3-6-08; 8:45 am]

BILLING CODE 410-26-P

---

## NATIONAL SCIENCE FOUNDATION

### Agency Information Collection Activities: Proposed Collection; Comment Request

**AGENCY:** National Science Foundation.

**ACTION:** Notice and request for comments.

**SUMMARY:** Under the paperwork Reduction Act of 1995, Public Law 104-13 (44 U.S.C. 3501 et seq.), and as part of its continuing effort to reduce paperwork and respondent burden, the National Science Foundation (NSF) is inviting the general public or other Federal agencies to comment on this proposed continuing information collection. The National Science Foundation (NSF) will publish periodic summaries of the proposed projects.

**Comments:** Comments are invited on: (a) Whether the proposed collection of information is necessary for the proper performance of the functions of the Foundation, including whether the information will have practical utility; (b) the accuracy of the Foundation's estimate of the burden of the proposed collection of information; (c) ways to enhance the quality, utility, and clarity of the information to be collected; and (d) ways to minimize the burden of the collection of information on those who are to respond, including through the use of automated collection techniques or other forms of information technology.

**DATES:** Written comments on this notice must be received by May 6, 2008, to be assured consideration. Comments received after that date will be considered to the extent practicable. Send comments to address below.

**FOR FURTHER INFORMATION CONTACT:** Ms. Suzanne H. Plimpton, Reports Clearance Officer, National Science Foundation, 4201 Wilson Boulevard, Suite 295, Arlington, Virginia 22230; telephone (703) 292-7556; or send e-mail to [splimpto@nsf.gov](mailto:splimpto@nsf.gov). Individuals who use a telecommunications device for the deaf (TDD) may call the Federal Information Relay Service (FIRS) at 1-800-877-8339 between 8 a.m. and 8

p.m., eastern time, Monday through Friday.

#### SUPPLEMENTARY INFORMATION:

**Title of Collection:** 2008 Survey of Doctorate Recipients.

**OMB Approval Number:** 3145-0020.

**Expiration Date of Approval:** February 28, 2009.

**Type of Request:** Intent to seek approval to reinstate an information collection for three years.

#### 1. Abstract

The Survey of Doctorate Recipients (SDR) has been conducted biennially since 1973. The 2008 SDR will consist of a sample of individuals under the age 76 who have earned a research doctoral degree in a science, engineering or health field from an U.S. institution. The purpose of this longitudinal panel study is to provide national estimates on the doctoral science and engineering workforce and changes in employment, education and demographic characteristics. The study is one of three components of the Scientists and Engineers Statistical Data System (SESTAT), which produces national estimates of the size and characteristics of the nation's science and engineering population.

The National Science Foundation Act of 1950, as subsequently amended, includes a statutory charge to " \* \* \* provide a central clearinghouse for the collection, interpretation, and analysis of data on scientific and engineering resources, and to provide a source of information for policy formulation by other agencies of the Federal Government." The SDR is designed to comply with these mandates by providing information on the supply and utilization of the nation's doctoral level scientists and engineers. Collected data will be used to produce estimates of the characteristics of these individuals. They will also provide necessary input into the SESTAT data system, which produces national estimates of the size and characteristics of the country's science and engineering population.

The Foundation uses this information to prepare congressionally mandated reports such as *Women, Minorities and Persons with Disabilities in Science and Engineering and Science and Engineering Indicators*. The NSF publishes statistics from the survey in many reports, but primarily in the biennial series, *Characteristics of Doctoral Scientists and Engineers in the United States*. A public release file of collected data, designed to protect respondent confidentiality, also will be made available to researchers on CD-

ROM and on the World Wide Web. A private contractor is currently being selected to conduct this study for NSF. Data will be obtained by mail questionnaire, computer-assisted telephone interviews and web survey beginning October 2008. The survey will be collected in conformance with the Confidential Information and Statistical Efficient Act of 2002. The individual's response to the survey is voluntary. NSF will insure that all information collected will be kept strictly confidential and will be used only for statistical purposes.

#### 2. Expected Respondents

A statistical sample of approximately 40,000 individuals with U.S. earned doctorates in science, engineering and health will be contacted in 2008. The total response rate in 2006 was 79%. NSF is also considering sampling 2,000 additional U.S. doctorates that received their degrees in the 2001-2007 academic years, who are non U.S. citizens, and indicated they planned on leaving the United States after they received their doctorate.

#### 3. Estimate of Burden

The amount of time to complete the questionnaire may vary depending on an individual's circumstances; however, on average it will take approximately 25 minutes to complete the survey. We estimate that the total annual burden will be 16,700 hours during the collection. If the additional 2,000 respondents who had plans to leave the United States are included in the sample, that will increase the burden an additional 850 hours to a total of 17,550 hours.

Dated: March 4, 2008.

Suzanne H. Plimpton,  
Reports Clearance Officer, National Science  
Foundation.

[FR Doc. E8-4483 Filed 3-6-08; 8:45 am]

BILLING CODE 755-01-P

---

## NATIONAL SCIENCE FOUNDATION

### Agency Information Collection Activities: Proposed Collection; Comment Request

**AGENCY:** National Science Foundation.

**ACTION:** Notice and request for comments.

**SUMMARY:** Under the paperwork Reduction Act of 1995, Public Law 104-13 (44 U.S.C. 3501 et seq.), and as part of its continuing effort to reduce paperwork and respondent burden, the National Science Foundation (NSF) is inviting the general public or other

**Attachment 3**

**Proposed 2008 SDR Mailing Materials**

**Draft 2008 SDR Contact Letters** (on customized NSF/SDR letterhead)

Prenotice #1: Past Respondents

Prenotice #2: Past Refusals

Prenotice #3: New Sample Members

Notice Email #1: Web start mode

Cover Letter #1: Past Respondents

Cover Letter #2: Past Refusals

Thank you/Reminder Postcard: All Sample Members

Follow-up Cover Letter: All Sample Members

*Prenotice Letter #1: Past SDR Respondents*

---

September 24, 2008

Dr. [FIRST NAME] [MI] [LAST NAME]  
[STREET ADDRESS]  
[CITY], [STATE] [ZIP CODE]

Dear Dr. [LAST NAME],

In a few days, we will be requesting your participation in the 2008 Survey of Doctorate Recipients (SDR), sponsored by the National Science Foundation and the National Institutes of Health. In the past, you contributed to this unique study of science and engineering doctorate holders. **Thank you for your past participation.** Your answers provided government, business, and academic institutions with crucial information concerning the availability of highly educated personnel in a variety of fields.

**The SDR has been conducted biennially since 1973 and is the only source of data on the careers of science and engineering doctorate holders from U.S. academic institutions.** The value of the information obtained over the years with the help of respondents like you is immeasurable. **Your cooperation in this research helps to ensure the validity and accuracy of the survey results.** Results from earlier studies are available from the NSF website listed below.

Because of their experience in directing studies of this type, we have asked the University of Chicago's National Opinion Research Center (NORC) to conduct this survey for us. The letter accompanying the questionnaire will explain more about this survey and our reasons for contacting you.

If you do not receive a questionnaire within two weeks or have any questions regarding this survey, please contact NORC via the toll-free number or e-mail address listed below. Staff are available from 9 a.m. to 9 p.m. (Central Time) to assist you.

I would greatly appreciate your continued participation in this significant research effort.

Sincerely,

Arden L. Bement, Jr.  
Director

*Prenotice Letter #2: Past Refusals*

---

September 24, 2008

Dr. [FIRST NAME] [MI] [LAST NAME]  
[STREET ADDRESS]  
[CITY], [STATE] [ZIP CODE]

I am requesting your participation in the 2008 Survey of Doctorate Recipients (SDR), sponsored by the National Science Foundation and the National Institutes of Health. In the past, we tried to contact you to contribute to this unique study, but did not receive a questionnaire from you.

The SDR is the only source of data on the careers of science and engineering doctorate holders from U.S. academic institutions. This survey provides government, business, and academic institutions with crucial information concerning the availability of highly educated personnel in a variety of fields. As such, your response is needed whatever your current employment status or occupation. Because you were scientifically selected for this study, we cannot substitute any other person for you. Your participation in this research will help to ensure the validity and accuracy of the survey results. Results from earlier studies are available from the NSF website listed below.

We will be sending you the questionnaire in the mail and will ask you to participate in the 2008 SDR at that time. Because of their experience in directing studies of this type, we have asked the University of Chicago's National Opinion Research Center (NORC) to conduct this survey for us. The letter accompanying the questionnaire will explain more about this survey. All information you provide will be kept strictly confidential and safeguarded in accordance with the Privacy Act of 1974.

If you do not receive a questionnaire within two weeks or have any questions regarding this survey, please contact NORC via the toll-free number or e-mail address listed below. Staff are available from 9 a.m. to 9 p.m. (Central Time) to assist you.

I would greatly appreciate your cooperation in this important research effort.

Sincerely,

Arden L. Bement, Jr.  
Director

*Prenotice Letter #3: New Sample Members*

---

September 24, 2008

Dr. [FIRST NAME] [MI] [LAST NAME]  
[STREET ADDRESS]  
[CITY], [STATE] [ZIP CODE]

Dear Dr. [LAST NAME],

I am requesting your participation in the 2008 Survey of Doctorate Recipients (SDR), sponsored by the National Science Foundation and the National Institutes of Health. This survey of people who have earned science and engineering doctorates in the United States has been conducted biennially since 1973. The survey, and the information it obtains, provides government, business, and academic institutions with crucial information concerning the availability of highly educated personnel in a variety of fields. The Survey of Doctorate Recipients is the only source of data on this important population.

You were scientifically selected from a database that contains the name and degree information for all individuals earning a research doctorate in the U.S. As such, we cannot substitute any other person for you. Your participation in this study will help ensure the validity and accuracy of the survey results. Results from earlier studies are available from the NSF website listed below.

We will be contacting you by telephone in a few days to ask for your participation in the 2008 SDR. Because of their experience in directing studies of this type, we have asked the University of Chicago's National Opinion Research Center (NORC) to conduct this survey for us. The interview should take roughly 25 minutes. All information you provide will be kept strictly confidential and safeguarded in accordance with the Privacy Act of 1974.

When the interviewer calls, they can answer any questions you have about participation in the survey. If you do not receive a call within two weeks or if you have any questions regarding this survey, please contact NORC via the toll-free number or e-mail address listed below. Staff are available from 9 a.m. to 9 p.m. (Central Time) to assist you.

I would greatly appreciate your cooperation in this important research effort. We look forward to talking with you.

Sincerely,

Arden L. Bement, Jr.  
Director

**Notice Email #1: Web Start Mode**

---

October 1, 2008

Dear Dr. [LNAME],

I am requesting your participation in the 2008 Survey of Doctorate Recipients (SDR) on behalf of its sponsors, the National Science Foundation and the National Institutes of Health. In the past, you contributed to this unique study of science and engineering doctorates holders. Thank you for your past participation. Your answers provided government, business, and academic institutions with crucial information concerning the availability of highly educated personnel in a variety of fields. Results from earlier studies are available from the NSF website listed below.

To access the 2008 SDR survey, please go to the following URL address:

<https://survey.norc.org/sdr>

Because this format is secure and confidential, you'll need to use a unique Personal Identification Number (PIN) and Password to access the survey after you have gone to the URL address.

Your unique PIN and Password are...

PIN: [PIN]      Password: [PASSWORD]

The information you provide will be collected at the University of Chicago's National Opinion Research Center (NORC). These data will be kept strictly confidential and safeguarded in accordance with the Privacy Act of 1974. Because we were not certain that this email address was still active, we have also sent online survey access information to your mailing address.

Your response to the 2008 SDR is needed whether you are working in or out of your doctoral field, are seeking employment, are retired or are in another situation. Important measures from this study include how many doctorate holders are working, what field they are working in, and how career patterns change over time. We can only learn this from you.

If you have any questions regarding the survey or the Internet format, please contact NORC toll-free at 1-800-685-1663 or respond to this e-mail. Staff are available from 9 a.m. to 9 p.m. (Central Time) to assist you. Or you can link to the SDR Frequently Asked Questions at

[http://www.norc.uchicago.edu/sdr/sdr\\_faq.asp](http://www.norc.uchicago.edu/sdr/sdr_faq.asp)

Thank you in advance for contributing to the SDR. We look forward to receiving your online questionnaire.

Sincerely,

Thomas B. Hoffer, Ph.D.  
SDR Lead Research Scientist  
NORC

~~~~~  
~  
For assistance completing the survey:  
Call NORC at 1-800-685-1663, e-mail [SDR@norc.uchicago.edu](mailto:SDR@norc.uchicago.edu) or visit  
[www.norc.uchicago.edu/sdr](http://www.norc.uchicago.edu/sdr).

For more information about the survey:  
Visit [www.nsf.gov/statistics/srvydoctoratework](http://www.nsf.gov/statistics/srvydoctoratework) or write NSF at 4201 Wilson Blvd, Suite 965, Arlington, Virginia 22230





*Cover Letter #1: First Questionnaire Mailing to Past Respondents*

---

October 1, 2008

Dr. [FIRST NAME] [MI] [LAST NAME]  
[STREET ADDRESS]  
[CITY], [STATE] [ZIP CODE]

Dear Dr. [LAST NAME],

I am requesting your participation in the 2008 Survey of Doctorate Recipients, sponsored by the National Science Foundation and the National Institutes of Health. In the past, you contributed to this unique study of our country's science and engineering doctorate holders. Thank you for your past participation. Your answers helped academic and government institutions anticipate shortages or surpluses in personnel and make decisions about graduate student support and research and development funding. The information you provided has also proven valuable for students who want to learn about the relationship between graduate education and careers.

At this time, we are asking for your participation in the 2008 SDR. Regardless of your employment situation—whether you are working in or out of your doctoral field, are seeking employment, are retired, or are in another situation—your response is vital. Important measures from this study include how many doctorate holders are working, what fields they are working in, and how career patterns change over time. We can only learn this from you. Results from earlier studies are available from the NSF website listed below.

Please complete the enclosed questionnaire and return it in the postage-paid envelope to the University of Chicago's National Opinion Research Center (NORC) as soon as possible.

If you have any questions about the survey, please contact NORC via the toll-free number or e-mail address listed below. Staff are available from 9 a.m. to 9 p.m. (Central Time) to assist you.

Thank you for your continued participation. I look forward to receiving your questionnaire.

Sincerely,

Lynda T. Carlson, Ph.D.  
Director, Division of Science Resources Statistics  
National Science Foundation

*Cover Letter #2: First Questionnaire Mailing to Past Refusals*

---

October 1, 2008

Dr. [FIRST NAME] [MI] [LAST NAME]  
[STREET ADDRESS]  
[CITY], [STATE] [ZIP CODE]

Dear Dr. [LAST NAME],

I am requesting your participation in the 2008 Survey of Doctorate Recipients, sponsored by the National Science Foundation and the National Institutes of Health.

The SDR is the only source of data on the careers of science and engineering doctorate holders from U.S. academic institutions. Regardless of your employment situation—whether you are working in or out of your doctoral field, are seeking employment, are retired, or are in another situation—your response is vital. Important measures from this study include how many doctorate holders are working, what field they are working in, and how career patterns change over time. We can only learn this from you.

The questionnaire takes about 25 minutes to complete. All information you provide will be kept strictly confidential and safeguarded in accordance with the Privacy Act of 1974. The data collected will be aggregated and used to prepare scientific reports, articles, and statistical summaries, but any information released publicly will maintain the confidentiality of all participants. Results from earlier studies are available from the NSF website listed below.

If you have any questions about the survey, please contact The University of Chicago's National Opinion Research Center (NORC) via the toll-free number or e-mail address listed below. Staff are available from 9 a.m. to 9 p.m. (Central Time) to assist you.

Please complete the enclosed questionnaire and return it in the postage-paid envelope to NORC as soon as possible. Thank you in advance for your participation.

Sincerely,

Lynda T. Carlson, Ph.D.  
Director, Division of Science Resources Statistics  
National Science Foundation

*Thank You/Reminder Postcard: All Sample Members*

---

October 8, 2008

About a week ago we asked for your help with an important national study. The study is being conducted by the University of Chicago's National Opinion Research Center (NORC) on behalf of the National Science Foundation. Many thanks to you if you have already participated and submitted the questionnaire.

If you do not have the materials necessary to complete the survey, or if you have questions about the study, please call NORC's toll-free number at 1-800-685-1663 between 9 a.m. and 9 p.m. (Central Time), visit their website at [www.norc.uchicago.edu/sdr](http://www.norc.uchicago.edu/sdr), or send an e-mail to [SDR@norc.uchicago.edu](mailto:SDR@norc.uchicago.edu). If you received the survey information but have not yet had time to complete it, please do so as soon as possible. Your participation is important for the success of this study.

With appreciation,

Lynda T. Carlson, Ph.D.  
Director, Division of Science Resources Statistics  
National Science Foundation

*Follow-up Cover Letter #1: All Sample Members*

---

November 26, 2008

Dr. [FIRST NAME] [MI] [LAST NAME]  
[STREET ADDRESS]  
[CITY], [STATE] [ZIP CODE]

Dear Dr. [LAST NAME],

We recently sent you the 2008 Survey of Doctorate Recipients (SDR) questionnaire to complete. The SDR is being conducted on behalf of the National Science Foundation and the National Institutes of Health by the University of Chicago's National Opinion Research Center (NORC). We have not yet received your completed survey, but hope that you will respond. Your participation will have a significant impact on the overall accuracy of the results and the ultimate usefulness of this research.

Your response is needed wherever you live, and whether you are working in or out of your doctoral field, are seeking employment, are retired, or are in another situation. Since you were scientifically selected for the SDR, we cannot substitute any other person for you.

Please take a few minutes to complete this important survey. We are offering three different ways by which you may participate.

1. Paper: If you prefer to complete a paper version of the survey, please call 1-800-685-1663.
2. Telephone: If you prefer a telephone interview, please contact us at 1-800-685-1663.
3. Web: If Internet access is convenient for you, we invite you to complete this survey online.

Please go to the following URL address: <https://survey.norc.org/sdr>

To ensure security, please use a unique Personal Identification Number (PIN) and Password below to access the survey after you have gone to the URL address.

PIN: [PIN] Password: [PASSWORD]

If you have any questions regarding the survey, please contact NORC via the toll-free number or e-mail address listed below. Staff are available from 9 a.m. to 9 p.m. (Central Time) to assist you.

Thank you in advance for your cooperation in this important research effort.

Sincerely,

Thomas B. Hoffer, Ph.D.  
SDR Lead Research Scientist  
National Opinion Research Center  
University of Chicago

**Attachment 4**

**Proposed 2008 SDR Questionnaire**

**Attachment 5**

**2008 SDR/ISDR Sample Strata and Sample Allocation Table**

**Attachment 6**

**Changes in the 2008 SDR Questionnaire**

## 2008 SDR Questionnaire Changes: Questions Added

---

The following questions were rotated in from the 2003 SDR questionnaire for the 2008 survey (2003 item names are in parentheses):

**2008 A20 (2003 A23):** Did your duties on this job require the technical expertise of a bachelor's degree or higher in...

*Mark Yes or No for each item.*

Engineering, computer science, math, or the natural sciences  
The social sciences  
Some other field (e.g., health, business, or education) – *Specify* →

**2008 A23 (2003 A26):** Which two reasons in question A22 were your most important reasons for taking this postdoc?

*Enter number of appropriate reason from question A22 above.*

- (1) \_\_\_\_ Most important reason  
(2) \_\_\_\_ Second most important reason (*Enter "0" if no second reason*)

**2008 C1 (2003 C1):** Since October 2003, how many...

Number

(1) Papers have you (co)authored for presentation at regional, national or international conferences? (*Do not count presentations of the same work more than once.*)

\_\_\_\_\_  
(2) Articles, (co)authored by you, have been accepted for publication in a refereed professional journal?

\_\_\_\_\_  
(3) Books or monographs, (co)authored by you, have been published or accepted for publication?

*enter "0")*

\_\_\_\_\_  
*(If none,*

**2008 C2 (2003 C2):** Since October 2003, have you been named as an inventor on any application for a U.S. patent?

- (1) Yes  
(2) No → *Go to question C4*

**2008 C3 (2003 C3):** Since October 2003...

Number

(1) How many applications for U.S. patents have named you as an inventor?



\_\_\_\_\_

(2) How many U.S. patents have been granted to you as an inventor?

\_\_\_\_\_

(3) How many of the patents recorded as granted (in category 2 above) have resulted in commercialized products or processes or have been licensed?

\_\_\_\_\_

*enter "0"*)

*(If none,*

**2008 E3 (2003 E3):** Did your spouse's or partner's duties on this job require the technical expertise of a bachelor's degree or higher in...

*Mark Yes or No for each item.*

- (1) Engineering, computer science, math, or the natural sciences
- (2) The social sciences
- (3) Some other field (e.g., health, business, or education) – *Specify* →

The second job module is new for the 2008 questionnaire as follows:

**2008 A38:** During the week of October 1, 2008, were you working for pay (or profit) at a second job (or business), including part-time, evening, or weekend work?

- (1) Yes
- (2) No → *Go to page 7, question A43*

**2008 A39:** (*If Yes*) What was the title of the second job you held during the week of October 1, 2008?

*If you had more than two jobs that week, report the job where you worked the second most hours.*

*Example: Clinical psychologist*

**2008 A40:** What kind of work were you doing on this job—that is, what were your duties and responsibilities on your second job? Please be as specific as possible, including any area of specialization.

*Example: Diagnose mental and emotional disorders and develop treatment programs*

**2008 A41:** Using the JOB CATEGORY list on pages 13-14, choose the code that best describes the second job you held during the week of October 1, 2008.

CODE \_\_\_ \_ \_

**2008 A42:** To what extent was your work on your second job related to your first U.S. doctoral degree? Was it...

*Mark one answer.*

- (1) Closely related
- (2) Somewhat related
- (3) Not related

## 2008 SDR Questionnaire Changes: Questions Dropped

---

The following questions were included in the 2006 SDR questionnaire, but will be rotated out from the 2008 survey questionnaire:

**2006 A27:** *“In performing the principal job you held during the week of April 1, 2006, did you...”*

**2006 A28:** *“(If Yes to Item 4 above) Did your work with individuals located in other countries*

*involve...”*

**2006 A29:** *“In your work with individuals located in other countries, did you...”*

**2006 A38:** *“Since completing your first doctoral degree, how many “postdocs,” if any, have you*

*held? Please include any postdocs you held through April 1, 2006.”*

**2006 A39acd:** *“Please provide the following information for each postdoc reported in A38.”*  
*(A39b*

*has been modified and retained; see below.)*

*2006 A39a: Date postdoc started and ended (or date you left)*

*2006 A39c: Which sector best describes where you worked for this postdoc?*

*2006 A39d: For this postdoc position, did your employer provide...Health benefits? Retirement benefits?*

**2006 A40:** *“To what extent did your most recent (or current) postdoctoral appointment...”*

## **2008 SDR Questionnaire Changes: Questions Modified**

---

One question from 2006 was modified to include two additional response categories (5) and (6), and the response category (4) was changed from 'Chronic illness or permanent disability' to 'Chronic illness or disability' based on verbatim responses reported in the 2006 questionnaire.

### **2008 A37 (corresponds to 2006 A37).**

- (4) Chronic illness or disability?
- (5) Full-time job not available
- (6) Held more than one job

**Attachment 7**

**2008 SDR Experiment Plan**

## 2008 SDR Experiment Plan

### **Experiment #1: No Preference Web Assignment Experiment**

**Synopsis:** Assess whether assigning 2006 CATI and mail respondents with no preference to 2008 Web start mode yields comparable or better response rates than assigning those cases to their 2006 completion mode.

**Background:** There are 3,169 SDR and ISDR sample members who completed the 2006 survey via CATI or mail, but who did not provide a mode preference for future survey rounds. They either expressed no mode preference (n=1,916) or did not answer the question (n=1,253). The current default mode for these individuals is the mode they completed in 2006. However, Web is the most cost-effective mode of data collection and yields high quality data in a timely manner and thus moving sample members with no preference into this start mode is recommended. In general, respondents who chose “no preference” are cooperative sample members. This is an opportunity to test whether assigning these types of cases to the more cost effective Web start mode can be assumed for future rounds.

**Sample:** This consists of 2006 respondents who meet the following conditions:

- a) Completed in mail or CATI
- b) Indicated “no preference” or did not answer the mode preference question
- c) Reported an e-mail address in the 2006 survey
- d) Are eligible for 2008 SDR sampling.

There are 2,573 experiment eligible cases available in the 2008 SDR panel frame. The distribution by 2006 SDR past response mode and 2008 age group for the experiment eligible cases in Table 1 below.

Table 1. “No Preference Web Assignment” experiment eligible frame by 2006 response mode and 2008 age group

| 2006 Response Mode      | Total | Under 35 | 35-39 | 40-44 | 45-49 | 50-54 | 55-59 | 60-64 | 65-75 |
|-------------------------|-------|----------|-------|-------|-------|-------|-------|-------|-------|
| All Experiment Eligible | 2,573 | 105      | 270   | 286   | 312   | 326   | 335   | 389   | 550   |
| Mail                    | 1,387 | 17       | 70    | 117   | 143   | 182   | 186   | 259   | 393   |
| Complete mail           | 1,278 | 16       | 69    | 108   | 134   | 167   | 174   | 246   | 364   |
| Retrieval needed        | 89    | 1        | 1     | 9     | 9     | 15    | 12    | 13    | 29    |
| CATI                    | 1,206 | 88       | 200   | 169   | 169   | 144   | 149   | 130   | 157   |
| Complete CATI           | 840   | 56       | 136   | 118   | 112   | 94    | 116   | 88    | 120   |
| Critical item only      | 366   | 32       | 64    | 51    | 57    | 50    | 33    | 42    | 37    |

**Design:** The treatment group will be assigned the Web start mode; the control group will be assigned either the CATI or mail start mode depending on their 2006 end mode. In experiment group selection, the sample will be stratified by 2006 response mode and sorted by age group. Analysis will be done by comparing response rate, days to response, and quality of data for the treatment versus the control groups. Ultimately, the results of the data will indicate if these types of cases can be assigned to the Web start mode in future rounds of the survey.

## **Experiment #2: Past Refusal Incentive Experiment**

**Synopsis:** Assess the effects on response rates and follow-up costs of offering an early incentive to sample members that refused to participate in the 2006 SDR.

**Background:** Different types of 2008 SDR panel refusal cases were identified utilizing the 2006 and 2003 final survey response. Panel refusal cases are always a challenge for the SDR. The 2003 and 2006 SDR incentive experiments showed that offering an incentive to past refusal cases significantly improved their cooperation rate.

**Sample:** After excluding panel cases that are classified as hostile and Congressional refusal cases, the remaining panel refusal cases (n=3,292) in the eligible 2008 SDR panel frame can be described as follows:

1. Chronic refusals – those that refused the 2003 and 2006 SDR (n=1,391)
2. Recent refusals – those that responded in 2003, but refused 2006 SDR (n=1,695)
3. Nearly new refusals – those that were new cohort in 2006 SDR and refused (n=206)

**Design:** Based on the results from the 2003 and 2006 mode experiments, the panel refusal cases will be assigned to the 2008 SDR mail starting mode. Within the mail start mode protocol, we will test the efficacy of including an early \$30 incentive check included in the first questionnaire mailing. Thus, the treatment group will receive a \$30 check in the first questionnaire mailing and the control group will receive the standard mailing package that includes a cover letter, questionnaire form and return mailing envelope.

These three types of panel refusal cases would be included in the experiments, and treatment and control assignments would be randomized across them. In addition to examining outcome differences among these groups within the treatment and control groups, we will also examine differences by their refusal type (soft, hard, or gatekeeper refusal) and retirement status. Ideally, the results will help determine the best method to avert refusals for the different types of panel cases that have refused the prior round.

**Experiment #3:** New Cohort Incentive Experiment

**Synopsis:** Test the efficacy of offering an early incentive to new cohort cases in the first versus the second questionnaire mailing.

**Background:** The 2006 SDR incentive experiment showed that an early incentive to new cohort members was more effective than one sent later in the field period, and definitely more effective than no incentive. For the new cohort, optimal incentive timing for maximum cost efficiency and greatest response is still ambiguous.

**Sample:** This incentive experiment will involve all 2008 SDR new cohort cases in the mail starting mode (estimated  $n = 3,500$ ) that have complete sampling stratification variables reported in the Survey of Earned Doctorates. These cases would be stratified by citizenship at birth and field of degree.

**Design:** Based on the results from the 2006 mode experiments, the 2008 SDR new cohort cases with complete stratification variables will be assigned to the mail starting mode. Within the mail start mode protocol, we will test the efficacy of including an early \$30 incentive check in the first questionnaire mailing versus including a \$30 incentive check in the second questionnaire mailing.

The treatment group will receive a \$30 check in the first questionnaire mailing and the control group will receive the standard mailing package that includes a cover letter, questionnaire form and return mailing envelope. At the second questionnaire mailing, the new cohort cases that are not locating problems and have not responded will be identified. Nonresponse cases in the treatment group will receive the standard second questionnaire mailing package with a repeat of the offer of the incentive, and the nonresponse cases in the control group will receive a \$30 incentive check in the second questionnaire mailing.

To analyze the effects of this experiment, we will compare the response rates at different times in the field period and at the close, days to complete, data quality, and cost.