SF-83 SUPPORTING STATEMENT

for

Survey of Earned Doctorates

for SED 2010 and 2011 survey cycles

TABLE OF CONTENTS

LIST OF ATTACHMENTS

SECTION A: JUSTIFICATION

This request is for OMB clearance for three years covering the 2010 and 2011 cycles of the Survey of Earned Doctorates (SED). The request represents an extension of a currently approved data collection. The SED has undergone an extensive review process to validate existing questionnaire items and test new items. The purpose of the review was to identify revisions and additions that would reduce respondent burden, improve data quality, reduce costs, and increase the efficiency of administrative processes in order to hasten the release of the survey data. The recommended changes from this review process and the reports supporting these recommendations are described in section B.4 of this report. Both the current (SED 2009) and recommended (SED 2010) questionnaire instruments are presented in Attachment 1. A list of methodological reports that have informed changes to the SED over the past 15 years appears in Attachment 9.1.

A.1. Necessity for Information Collection

The Survey of Earned Doctorates began in 1957–58 to collect data continuously on the number and characteristics of individuals receiving research doctoral degrees from accredited U.S. institutions. All individuals receiving research doctorates from accredited U.S. institutions are asked to complete the survey. A research doctorate is a doctoral degree that (1) requires the completion of an original intellectual contribution in the form of a dissertation or an equivalent project of work (e.g., musical composition), and (2) is not primarily intended as a degree for the practice of a profession. The most common research doctorate degree is the Ph.D. Doctorate recipients of professional doctorate degrees such as M.D., D.D.S., J.D., D.Pharm., and Psy.D are not included in the survey. The results of this annual survey are used to assess characteristics and trends in doctorate education and degrees. This information is vital for educational and labor force planners within the federal government and in academia.

The Survey of Earned Doctorates is sponsored by the National Science Foundation in cooperation with the National Institutes of Health (NIH), the U.S. Department of Education (USED), the U.S. Department of Agriculture (USDA), the National Endowment for the Humanities (NEH), and the National Aeronautics and Space Administration (NASA). Each sponsoring agency provides funding for the SED and obtains special tabulations from the survey each year for their unique needs, and also receives uniform data tabulations/reports that are provided to all sponsors. The representatives of each of the sponsoring agencies and the list of persons who have been consulted and/or have reviewed the SED 2010-2011 questionnaire are listed in Attachment 5. The National Science Foundation has monitoring responsibility for the project, which is currently conducted under contract by the National Opinion Research Center (NORC), affiliated with the University of Chicago. NORC was competitively awarded a new procurement in 2007 that covers the SED operations from 2007 to 2012.

The National Science Foundation Act of 1950 (as amended by Title 42, United States code Section 1862, Attachment A) requires the NSF:

... 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 ....

Statutory authority for collection of information for fields other than science and engineering comes from legislation for the other Federal sponsoring agencies. The following is a list of the applicable legislation:

1. NIH: Title I of the National Research Act of 1974 (PL 93 348);

2. ~~Department of Education: Section 406(b) of the General Education Provisions Act, as (20 U.S.C. 122le-1);~~Education Sciences Reform Act of 2002;

3. NEH: Section 956(k) of the Arts, Humanities, and Museums Amendments of 1990, as enacted in Public Law 10 1 -512;

4. USDA: Title XIV of the Agriculture and Food Act of 1977 (PL 95-113) as amended, and Title V of the Equity in Educational Land-Grant Status Act of 1994 (PL 013-382) as amended, Sec. 354;

5. NASA: Title 42 of The Public Health and Welfare and Chapter 26 of the National Space Program.

Attachment 3 provides the cited legislation for each sponsoring agency; these agencies may change depending on their continuing funding for the SED.

A.2. Uses of Information

The Survey of Earned Doctorates has been conducted annually since 1958 and is jointly sponsored by six Federal agencies in order to avoid duplication in data collection. The purpose of the SED is to compile data on all recipients of earned research doctorates awarded by U.S. universities. It is an accurate, timely source of information on one of our nation's most precious resources – highly educated individuals. There is minimal burden on the respondents and the resulting information is used extensively by many Federal agencies for program evaluation, policy formulation, and dissemination of results.

The SED is the only data source that provides comprehensive information on the education and early career commitments of persons who have recently received doctorates. The resulting information is a valuable resource for other government agencies, academic researchers and policymakers, as well.

The results of the Survey of Earned Doctorates (SED) each academic year become part of the Doctorate Records File (DRF), a virtually complete database of nearly 1.7 million U.S.-educated doctorate recipients from 1920 to 2007.

The six sponsoring agencies have made extensive use of the Survey of Earned Doctorates. Detailed tables, tabulations, and the computer files are available to representatives of the agencies that sponsor the SED for use in program planning/evaluation, policy development, and dissemination. The heads of the agencies use the data in their reports and speeches, as well as in national forum discussions of educational policy.

There is no public use data base available from the SED; however, organizations and individuals can request information from the SED database and these special tabulations are provided at cost. In addition, selected SED data items are available to the public through the WebCASPAR database on the internet. Statistical data from the SED are widely used by other Federal agencies, Congress, state agencies, universities, professional societies, and individuals doing research in science policy, graduate education, economics, and human resource planning.

Current Uses of the SED at the Federal Level

The uses made of the Survey of Earned Doctorates reflect the fact that it is the most comprehensive, accurate, and timely source of data on doctorate degree awardees in the United States. The use of SED data and reports is widespread among sponsoring Federal agencies and other Federal organizations. The data are used for policy development, in carrying out program responsibilities of the agencies, and in the administration of agency programs. The data are disseminated extensively throughout the agencies. Some of the more important recent uses, organized by user agency, are listed below. The participating Federal agencies are subject to change, pending funding availability; the current liaisons for each sponsoring agency are listed in Attachment 4.

a. The National Science Foundation

The National Science Foundation has been a sponsor of the SED since 1958. The uses made of the data on science and engineering (S&E) doctorates are many:

- The Survey of Earned Doctorates is used as the universe frame for selecting the sample of doctoral scientists to be included in the NSF's Survey of Doctorate Recipients (SDR), a longitudinal survey of doctorate recipients in science, engineering, and health fields.

- The SED serves as a measure of program effectiveness; the Graduate Research Fellowship Program uses the information on those who complete the Ph.D. to evaluate the effectiveness of the program and its design requirements. Many programs within the NSF, especially those dealing with women, minorities, and persons with disabilities use data from the SED for program planning. While these programs focus on U.S. citizens, data on foreign citizens studying here for their Ph.D. are also useful for international comparisons and for quantifying the attraction of the U.S. graduate education around the world.

- Several reports are published on science and engineering doctorates by the NSF for internal and external use. The first report to be released each year is available publicly in

November, seven months after survey closeout. Additional reports follow which provide more detailed data or more analysis of the results from the Survey of Earned Doctorates.

Special tabulations of data from the survey constitute a key resource in meeting policy and program information needs of the Foundation. Examples of uses within the Foundation include:

- Data on doctorates awarded to minorities and women for presentation to the National Science Board for their use in examining participation of these groups.

- Data on foreign scholars provided to an interagency committee studying foreign access to U.S. science and engineering at American colleges and universities.

- Baccalaureate institutions of science and engineering doctorate recipients supplied to the NSF's Division of Undergraduate Education for use in a study of institutions' contributions to the highly trained labor force.

- Data supplied by the NSF to outside users. At the national level, within recent years, major data users have included the White House Office of Scientific and Technology Policy, the National Academy of Sciences, and others.

- Published results in widely distributed NSF publications. Data are included in two of the Foundation's Congressionally-mandated biennial reports, *Science and Engineering Indicators,* and *Women, Minorities and Persons with Disabilities in Science and Engineering.*

- A wide range of topics related to non-U.S. doctorate recipients addressed in *Science and Engineering Indicators* report, and in selected data tables that are available electronically on the Science Resources Statistics (SRS) Web Site (www.nsf.gov/statistics/doctorates).

- Aggregated data on selected variables are available for each institution through the WebCASPAR database, available to the public (http://webcaspar.nsf.gov).

b. The National Institutes of Health

For more than thirty years, the NIH has used the results of the SED to meet a variety of planning, evaluation, and reporting needs:

- **Planning for the medical research workforce.** Since 1975, committees of the National Academies (NAS), convened at the NIH's request, have relied on the results of the SED to monitor Ph.D. production in the biomedical, behavioral, and clinical sciences and to determine the national need for investigators in these fields. Based on their findings, these NAS committees make recommendations for the future size and distribution of NIH research training programs. The next iteration in this ongoing series of congressionally-mandated studies is currently underway and expected to be completed in late 2009.

- **Evaluating NIH research training programs.** Because the SED has proven to be such a reliable and comprehensive source of information on new Ph.D.s, NIH routinely uses SED results to track the educational outcomes of NIH predoctoral trainees and fellows and to assess its research training programs. By comparing its internal records with the results of the SED, NIH regularly monitors Ph.D. completion rates for students participating in NIH-sponsored training programs, their time to degree, and their plans for postdoctoral study or employment. In evaluating its research training programs, NIH also uses the SED to identify comparison groups of non-NIH-sponsored students in the same fields of study.

- **Fulfilling reporting requirements.** By allowing comparisons and sustained tracking of selected doctoral recipients, the SED is a critical tool in Government Performance and Results Act (GPRA) reporting on the effectiveness of NIH predoctoral training grants. In 2008, NIH also used the results of the SED to report to Congress on the number and type of graduate degrees awarded with NIH support.

c. The Department of Education

The U.S. Department of Education has sponsored the Survey of Earned Doctorates since 1958. The Department's National Center for Education Statistics (NCES), Postsecondary Education Statistics Division, funds the survey and makes extensive use of a range of SED data. Reports have been published on a time-series analysis of doctorates in the field of education, as well as in other fields.

NCES has also used data on the postgraduate plans of new doctorates. Trend data are compiled each year and displayed in tables in the Center's publication *Digest of Education Statistics*. NCES has also published a report containing tables from the doctorate records file comparing education doctorates to doctorates in other fields, by selected characteristics.

Data from the SED are also used for evaluation by the Department of Education's programs, such as the Office of Student Financial Aid, the individual program offices, and by the Office of Planning, Evaluation, and Policy Development.

d. The Department of Agriculture

The United States Department of Agriculture (USDA), a sponsor since 1988, has developed a list of discipline areas in which it has particular concerns, analogous to the subsets developed by the National Science Foundation and the other sponsors, and has requested trend tabulations on doctorate recipients in these fields. Data collected in the Survey of Earned Doctorates are used to evaluate how widespread these programs and fields are in the United States. Data are also used in the evaluation and planning of 1890 Land Grant and Tribal College programs. A considerable and expanding number of tabulations from the SED are also available on their newly developed Food and Agricultural Education Information System (FAEIS) and other Websites.

e. The National Endowment for the Humanities

The authorizing legislation for the National Endowment for the Humanities (NEH) tasks the Endowment to "develop a practical system of national information and data collection on the humanities, scholars, educational and cultural groups, and their audiences." The Survey of Earned Doctorates gives university administrators, federal funding agencies, and private foundations an annual reading of a vital index of teaching and scholarship, the national output of humanities doctorates. NEH is currently participating in an effort led by the American Academy of Arts and Sciences to develop and regularly release via the web, a set of *Humanities Indicators*. SED data on doctorate production will provide a key "indicator" of the health of the humanities workforce.

f. The National Aeronautics and Space Administration

As a sponsor since 1995 of the SED, the Education Division of the National Aeronautics and Space Administration (NASA) has developed a program for the utilization of data from the SED in its planning and information dissemination activities. SED data have been useful in providing information on the progress of women and minorities in science and engineering.

g. Other Federal Agencies and Congress

Other Federal agencies have utilized the SED in several ways - through requests for special tabulations and tables, data requests from NSF, and licensing agreements. Congressional staff have called NSF for information on several topics relevant to legislation development (such as debt levels of science doctorates at graduation) and national security interests (such as nuclear engineering doctorates awarded to foreign citizens).

Academic Uses of the SED

The nation's Ph.D.-granting institutions not only contribute to the SED data set but also utilize the data for many purposes. Each year since 1997, the sponsors have provided to the dean of each graduate school profiles of their doctorate recipients' demographic characteristics, debt status, postgraduation plans, employment and other activities, compared with national and peer-institution data (see Attachment 8.1 for an example of an Institutional Profile).

Graduate and baccalaureate institutions use the data in program planning, comparison with other institutions or with national figures, and in the development of affirmative action plans. SED data on the number of Ph.D.s awarded to racial/ethnic minorities are used extensively by institutions as the only reliable source of the supply of persons with particular characteristics for academic positions. Site visits have confirmed the usefulness of the data to institutions. In the past few years, NSF has seen the increasing use of data by Graduate Deans to address issues of changes in the composition of the graduate schools and the time it takes to complete the doctorate, etc., and they have relied on the Survey of Earned Doctorates as the "ready made" data base on their graduates.

An analysis of users of the data shows that academic institutions are the primary users of the data outside of the Federal sponsors. Over the past three years, over 125 requests for data by the

graduate deans or other administrative offices have been fulfilled. The U.S. universities help administer the Survey of Earned Doctorates to their doctorate recipients but it is clear that they get something back from the survey because they can use the data for their own purposes; this is a symbiotic data collection effort.

In addition, researchers can apply to use selected microdata from the SED under the SRS Restricted Use Data Licensing Agreement, if publicly available data do not address the specific needs of their study. The SRS Licensing Agreement is executed between an institution and the SRS. The SRS license is a legal document that requires that stringent security procedures be taken to ensure that the confidential data will be protected against unlawful disclosure.

A.3. Consideration of Using Improved Technology

Planning for the development of an electronic questionnaire administration of the SED was begun in late 1999 and has been refined, implemented, and expanded since that time. The purposes of instituting an electronic, Web-based option were to ease the burden on students by offering alternative completion modes, to help assure continued high response rates, and provide an option for institutions using electronic graduation packages.

The population for this survey, graduating doctoral students, is virtually all computer literate and familiar with the internet. Offering a paperless survey version which can be submitted electronically is not only appealing to many students but is also very practical for respondents who are often relocating at the close of their studies and are not near the graduate offices for submission of their completed questionnaires. The objective of ensuring high response rates is facilitated by requesting students' e-mail addresses and following up electronically with them for missing critical items or survey questionnaires.

The phase-in of the SED electronic procedures has been gradual and consists of three parts: a Web survey instrument to parallel the paper version; follow-up for missing questionnaires and items via e-mail; and a web institution interface which is password protected for each school (see Attachment 8.2). The Web version of the instrument was carefully developed and tested to assure that mode effects between it and the paper version were minimal and non-biasing. Students accessed the Web version by entering a PIN/password which they registered for, was distributed by the school, or that they received in a letter or e-mail. Prompting for missing surveys or critical items was accomplished by the same mechanism. Because web survey design has evolved since 1999, an expert review of the web-based data collection instrument is planned for 2009, with suggested changes implemented in the 2011 SED survey cycle. See section B.4 for more information on this methodological study.

The Web Institution Interface (WII) is available to all participating institutions and it allows Institutional Contacts to enter their own password-protected site to monitor completion rates for their graduates, to link to various SED reports, and to print questionnaires from PDF document files. The WII also allows Institutional Contacts to compare their list of graduates, and their completion status, with the SED contractor-maintained database in order to track the response rate for their institution.

In general, the web-based data collection system is being adopted slowly by both graduates and institutions. In the 2007 SED, 3,112 of the 44,030 (7 percent) individual completions were done via the Web. This number included both students who graduated from schools that use the web survey as their primary or secondary method of completion, and from electronic prompting of non-respondents, which was a small but important component of the completion rate. NSF expects that this rate will continue to increase as more Graduate Schools move to web-based data collection for their graduates. NSF also anticipates that the SED web instrument will be easier to use for the respondents as a result of improvements implemented after the expert review of the web-based SED data collection instrument.

The institution (usually the Graduate Dean's office) is the main SED interface with the doctorate recipient. The Institutional Contact at the institution helps distributes the survey, tracks it, collects it and submits the completed questionnaires to the SED survey contractor. The SED work is, however, but one of many duties performed by the Institutional Contacts, who have varying degrees of computer literacy and familiarity with the SED. The SED also must fit in with the graduate schools' procedures for processing and awarding doctoral degrees. The work of maintaining the overall cooperation of the schools while phasing in web-based systems remains a continuing challenge, as the response rate usually drops when the school goes to a Web collection. The SED solicits feedback from schools using an electronic survey for ideas for process improvements and continues to promote the web survey to the institutions. However, the final decision on how to best conduct the SED is left to the administrators at each participating institution based on their individual preferences and capabilities.

The survey contractor has also been investigating the possibility of scanning returned hard copy questionnaires to reduce both the time needed to process returns and the costs associated with manual Computer-Assisted Data Entry (CADE). As with CADE, scanned questionnaires would also be edited during data entry through verification software. The feasibility of scanning will continue to be tested, and possibly implemented during the 2011 SED survey round. Implementing scanning will have no impact on the time to complete the survey or the questions themselves. It will just require printing a barcode on the cover of the questionnaire, and printing individual identifiers (in the form of a 12-digit number) on the bottom corner of every other page.

A.4. Efforts to Identify Duplication

The National Science Foundation has reviewed other governmental surveys through direct contacts with other agencies. At the present time, no survey gathers identical or similar information. In addition, the National Science Foundation actively maintains contacts with professional societies and groups, such as the Council of Graduate Schools, within both the higher education and data collection communities, so that information about any surveys similar to the Survey of Earned Doctorates would be immediately known. Both the American Psychological Association and the Hispanic Association of Colleges and Universities have sent letters to their members endorsing the SED.

SED survey content is coordinated with NSF's Survey of Doctorate Recipients (SDR) and with the SESTAT data collections on scientists and engineers. The Survey of Doctorate Recipients is designed to provide demographic and career history information about a sample of individuals with doctoral degrees in science, engineering and health fields. The results of the SDR are vital for educational planners within the federal government and in academia. The SDR results are also used by employers in all sectors (education, industry, and the government) to understand trends in employment and salaries in S&E fields for doctorate holders and to evaluate the effectiveness of equal opportunity efforts. NSF finds the results important for internal planning, because most NSF grants and fellowships go to individuals with doctoral degrees. The Doctorate Records File from the SED is used to identify SDR respondents in science, engineering and health fields. Contact information obtained in the SED is necessary for contacting new Ph.D.'s who are added to the SDR sample every two years. The coordination of content and procedures is, therefore, critical to the success of both the SED and SDR surveys.

The data base system known as the Scientists and Engineers Statistical Data System (SESTAT) combines data from the SDR, the National Survey of College Graduates, and the National Survey of Recent College Graduates. The SESTAT system is designed to provide a comprehensive picture of the number and characteristics of individuals with training and/or employment in science and engineering in the United States. The SED survey content is coordinated with the SESTAT surveys to avoid unnecessary duplication of items and to assure relevant uniform approaches on similar items such as race and ethnicity.

Differences between the Survey of Earned Doctorates and the Integrated Postsecondary Education Data System (IPEDS), which collects some information on doctoral degrees, are outlined below. The IPEDS Completions survey, conducted by the National Center for Education Statistics (NCES) of the Department of Education, collects aggregate data from institutions on numbers of degrees at each level by discipline and gender. IPEDS provides data on aggregates of institutional doctorate recipients (including race/ethnicity and gender) and not data by individuals.

The Survey of Earned Doctorates obtains information from the individual research doctorate recipient on over two dozen variables - information not collected through the IPEDS survey. As mentioned earlier, NCES uses the Survey of Earned Doctorates to present data that are not available from IPEDS.

There are four data items collected on both the SED and IPEDS that may appear, on the surface, to be duplicative: field of study (FOD) and the demographic variables of citizenship, gender, and race/ethnicity. However, important purposes are served by including these variables in both databases:

- In the SED, FOD, citizenship, race/ethnicity, and gender are frequently used in analyses that link these variables with other key variables (such as the length of time spent pursuing the degree and the amount of debt accumulated during the graduate education). These other linked variables cannot be collected from the institutions that provide information to IPEDS. The FOD and demographic variables are also used to identify individuals in "rare subgroups" for oversampling in the SDR (described above). Without

these questions, the SDR would need to be greatly expanded to meet the needs for the congressionally mandated report, *Women, Minorities and Persons with Disabilities in Science and Engineering,* for education and labor market data.

- It is also not a feasible option to exclude collection of the information about doctoral degree recipients from IPEDS, because inclusion of field of degree, citizenship, gender, and race/ethnicity permits comparative analyses of trends in degree production at different degree levels. SED data cannot be substituted for the IPEDS in such comparisons, because of the inevitable differences between an institutional survey and a demographic survey. For example, individuals' racial/ethnic self-identification on these variables may differ from those maintained by the institutions.

- Including FOD, citizenship, race/ethnicity and gender questions on both surveys also provides important validity checks for both surveys at the aggregate level.

A.5. Efforts to Minimize Burden on Small Business

Not applicable. The SED does not collect information from small businesses.

A.6. Consequences of Less Frequent Data Collection

The SED is an important source for monitoring changes in academic fields of study and participation in disciplines by demographic groups of interest (including U.S. and non-U.S. citizens on both permanent and temporary visas). These data provide an annual barometer of the market conditions encountered by new doctoral degree recipients and are therefore an integral component in policy implementation and program design.

Less frequent data collection would also result in a more complicated administration of the survey in the Graduate Deans' offices. The survey collects data from each person receiving a doctorate at the time they complete the requirements for their degree. Staff at the Graduate Deans' offices insert the Survey of Earned Doctorates questionnaire into the package of materials for doctorate recipients. Any less frequent collection of the SED would yield far lower response rates because the Graduate Deans' offices would be uncertain about the timing and distribution of questionnaires to prospective doctoral graduates, a process which now occurs continuously throughout the survey year. Discussion with the Council of Graduate Schools and several universities confirms the extreme difficulty graduate schools would have in operating the survey on a stop-and-start basis. Stability of both the survey questionnaire and of the survey collection process is imperative for the usefulness of the data to the Federal agencies and for the ease of collection by the universities. A continuation of the current survey methodology serves the best interests of all involved.

If the SED were conducted less frequently, there would also be significant repercussions to the success of the Survey of Doctorate Recipients (SDR). The Doctorate Records File is the sample frame used to identify SDR respondents. Locating information obtained in the SED survey is

necessary for contacting the new Ph.D.s who are added to the SDR sample. The coordination of timing, content and procedures of these two studies is, therefore, critical to the success of both the SED and SDR surveys.

A.7. Special Circumstances

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

A.8. Federal Register Announcement and Consultations Outside the Agency

The Federal Register announcement for the SED appeared on December 5, 2008 (see Attachment 6). Public comments have been received by NSF from one person in response to the announcement, as of the close-out date of February 9, 2009. The comment came from B. Sachau of Floram Park, NJ, via e-mail on December 9, 2008. Ms. Sachau objected to the information collection. Ms. Sachau had no specific suggestions for altering the data collection plans other than to discontinue them entirely. NSF responded to Ms. Sachau on December 17, 2008, describing the program, and addressing the survey cycle frequency and the cost issues raised by Ms. Sachau. NSF believes that because the comment does not pertain to the collection of information on the required questionnaires for which NSF is seeking OMB approval, NSF is proceeding with the clearance request.

Consultations Outside the Agency

In the many years of operation of the Survey of Earned Doctorates, the six Federal sponsors and the survey contractors have consistently invited others to comment on the SED. The comments come from the Federal sponsors themselves, an expert panel convened by SRS, the Council of Graduate Schools, and other governmental and academic institutions. Comments and suggestions regarding the Survey of Earned Doctorates and the manner in which it is conducted have been received from individual respondents, university faculty advisors, Graduate Deans' offices, and professional researchers by telephone, mail, and in person contacts. University representatives have been sought out for consultation at venues such as professional conferences and meetings. These consultations have helped to determine if there are problems in the conduct of the survey or in the interpretation of certain items. These problems are discussed with graduate deans for their conceptual validity and applicability to all fields of study, and the need for such information is weighed against respondent burden.

The Federal sponsors meet together at least twice a year to discuss the SED design, operation and dissemination activities, and to plan future activities. The SED sponsors also review recent trends in the number of doctorate recipients receiving degrees in emerging fields of study – that is, fields of study not currently coded within the SED taxonomy – and in fields of study for which there are few graduates. This review is the basis for the decision made every three years on taxonomy changes for the SED.

SRS has convened a Human Resources Experts Panel (HREP) in order to help 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 accomplishes 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 meets twice a year and three meetings have been held since the panel convened in 2007. The SED was the main topic at the October 2008 meeting which discussed users' needs for race/ethnicity and gender data and possible alternative options for providing the data to meet those needs while still protecting respondent confidentiality.

Formal site visits have been conducted by National Science Foundation and survey contractor staff for the purpose of consulting with graduate deans and campus administrators. The institutions visited include those with poor response rates, primarily to resolve the survey collection problems at those institutions. However, the site visits also allow for the discussion of the uses of the Survey of Earned Doctorates by the Federal sponsors and by the universities themselves.

Other Consultations

The SED has also been informed by numerous other contacts between NSF and the user community. For example, routine information requests provide insight into the interests of the general public. In addition, there has also been consultation with members of the respondent population for the survey.

In 2008, NSF conducted a series of eight outreach meetings across the country with representatives of minority-serving doctoral degree-granting institutions, leading institutional producers of doctoral degrees to minority recipients, and STEM (Science, Technology, Engineering, and Mathematics) professional organizations. The meetings were designed to allow participants an opportunity to learn how and why SRS protects the confidentiality of data supplied by SED respondents, discuss their specific needs for and uses of SED data, and provide input to SRS on data-reporting strategies that could better meet their data needs. A list of institutions and organizations that participated in these outreach meetings is recorded in Attachment 9.2. In addition, a web-based survey was developed to better understand the preferences, concerns, needs for information, and insights of different communities of SED users. The web survey was offered to users of SED data products. The findings from these outreach efforts will help inform the redesign of statistical tables that report on the race/ethnicity, citizenship, and gender of doctorate recipients.

A.9. Payment or Gifts to Respondents

No incentives in the form of payment or gifts to the doctoral graduates are used in the SED.

A.10. Assurance of Confidentiality

The SED is collected in conformance with the strict confidentiality requirements found in the NSF Act of 1950, as amended. The SED is also collected in conformance with the Privacy Act of 1974, including the section of the Privacy Act requiring notification of the respondent concerning the uses to be made of the data and the voluntary nature of his/her responses. The confidentiality pledge to the SED respondents is:

> This information is solicited under the authority of the National Science Foundation Act of 1950, as amended. All information you provide is protected under the NSF Act and the Privacy Act of 1974, and will be used only for research or statistical purposes by your doctoral institution, the survey sponsors, their contractors, and collaborating researchers for the purpose of analyzing data, preparing scientific reports and articles, and selecting samples for a limited number of carefully defined follow-up studies. The last four digits of your Social Security Number are also solicited under the NSF Act of 1950, as amended; provision of it is voluntary. It will be kept confidential. It is used for quality control, to assure that we identify the correct persons, especially when data are used for statistical purposes in Federal program evaluation. Any information publicly released (such as statistical summaries) will be in a form that does not personally identify you or other respondents. Your response is voluntary and failure to provide some or all of the requested information will not in any way adversely affect you.

~~The SED will be collected in conformance with the Privacy Act of 1974, including the section of the Privacy Act requiring notification of the respondent concerning the uses to be made of the data and the voluntary nature of his/her responses. The confidentiality pledge to the SED respondents is: "ALL INFORMATION YOU PROVIDE WILL BE TREATED AS CONFIDENTIAL and used only for research or statistical purposes by your doctoral institution, the survey sponsors, their contractors, and collaborating researchers for the purpose of analyzing data, preparing scientific reports and articles, and selecting samples for a limited number of carefully defined follow-up studies." Further, the SED is collected in conformance with the strict confidentiality requirements found in the NSF Act of 1950, as amended.~~

Specific procedures for protecting both hard copy and electronic data are used by the survey contractor, and all project staff are required to sign confidentiality agreements before they first have access to any SED data, and on a yearly basis thereafter (see Attachment 7). Data files with personal identifiers are provided to two Federal Sponsors and their contractors: NSF and NIH. NIH receives the personal identifiers only after they have executed an SRS Restricted data use agreement (license) and all contractors have signed data use agreements. As indicated explicitly in the confidentiality statement, the graduate dean of the respondent's institution may request data for respondents from that institution only with a written agreement to use such data for

statistical and program evaluation purposes only. No one outside of these groups can obtain data files with direct identifiers such as phone numbers and addresses. Qualifying researchers can obtain microdata on selected variables (but no direct identifiers) only by executing an SRS Restricted Use Data Licensing Agreement with NSF through their employer.

A.11. Justification for Sensitive Questions

The SED recognizes the growing sensitivity of requesting respondents' Social Security Numbers to an increasing segment of the population. The SED is allowed to collect respondent Social Security Numbers under the NSF Act of 1950 (42 U.S.C. 1861 et seq.), as amended, and in accordance with the Privacy Act of 1974. However, the SED only collects the last four digits of the Social Security Number to be used to ascertain the correct identity of the survey respondent in survey operation and evaluation purposes.

A.12. Estimate of Respondent Burden

The SED is a census of all individuals receiving a research doctorate in the United States in an academic year. In 2010, approximately 49,000 individuals are expected to receive research doctorates from U.S. institutions. Based on the results of cognitive interviews and tests by NORC staff using the proposed SED 2010 questionnaire, the average respondent is expected to complete the questionnaire in 20 minutes, a one minute increase over the estimate cited for earlier versions of the SED. Thus, the respondent burden for completing the survey questionnaire is estimated to be 16,333 hours.

Additional time is needed to complete the Missing Information Letter (MIL), which is sent to any survey respondent who did not provide a critical item on their original response. The MIL attempts to gather missing data on a maximum of eight items (year of Master's, year of Bachelor's, postgraduation location (state or country), birth date, citizenship status, race, ethnicity, and gender), though most MILs address fewer than eight. Based on the past results of phone retrieval efforts to obtain these critical data items, the average respondent is expected to spend two minutes completing the MIL. The SED receives an average of 2,000 completed MILs each survey round, for a burden estimate of 67 hours.

Therefore, the entire information burden for individual respondents is estimated to be 16,400 hours.

The estimated cost to respondents for this data collection is $426,400, based on the estimated 16,400 hours of response burden at a time-cost of $26.00 per hour. The $26.00 per hour time-cost estimate is derived from preliminary results from the 2006 SDR survey which indicate that the median income for individuals with science and engineering doctoral degrees who are 35 years of age or younger was $56,000 in 2006. Assuming a 40-hour work week and 52 weeks of work per year, this $56,000 per year figure works out to $26.92 of income per hour. To account for the typically lower salaries earned by doctorate recipients in non-science and engineering fields, the time-cost estimate was adjusted downward to $26.00 per hour.

In addition to the actual survey, the SED also requires the collection of administrative data from participating institutions. The Institutional Contact (IC) at the institution helps distribute the survey, track it, collect it and submit the completed questionnaires to the SED survey contractor. This involves the following materials sent to the Institutional Contacts (See Attachments 8.3 through 8.7 for examples of the materials):

- Transmittal Form, to be included in the package of completed surveys, asks for the total number of graduates for the graduation date in question, the total number of surveys enclosed, and contact information for each non-respondent;
- Interim Result Form, sent two-three times a year, reports the number of graduates currently accounted for in each graduation date;
- Address Roster Form asks for a physical mailing or e-mail addresses for non-respondents that were not already provided on the Transmittal Form;
- Missing Information Roster sent twice a year, four and seven months after the end of the data collection period, asks for the critical items for any non-respondents and the missing critical items for respondents;
- Dean/Contact change form sent twice a year, asks for any change to the contact information of the dean or Institutional Contact.

Based on focus groups conducted with Institutional Contacts, it is estimated that the SED demands no more than 1% of the contact's time over the course of a year, which computes to just under 21 hours per year (40 hours per week x 52 weeks per year x .01). At an estimated average wage rate of $24/hour for Institutional Contacts, and with 530 participating programs, the total estimated time-cost to Institutional Contacts of administering the SED is $267,120 per year. The $24/hour wage estimate is derived from the Bureau of Labor Statistics "Occupational Employment and Wages, May 2007" for a combination of Office and Administrative Support Occupations (60% of Institutional Contacts) and Education Administrators, Postsecondary (40% of Institutional Contacts).

The SED also anticipates conducting methodological research over the clearance cycle that would involve both respondents and Institutional Contacts (see Section B.4 for description of possible research). These tasks would most likely involve focus groups or cognitive interviews of as many as 60 respondents for up to two hours each, for an estimated maximum burden time of 120 hours total. Additionally, the SED expects to conduct a short web survey of all the Institutional Contacts. The survey would be designed to be completed in ten minutes or less. This would mean an expected burden of 88 hours (530 x 10 minutes).

The chart below summarizes the annual burden anticipated for all the tasks involved with conducting the SED:

Description	# of Responses	Respondent Burden	Annual Burden Hours	Annual Cost Burden
Doctorate Recipient SED Questionnaire	49,000	20 minutes	16 ,333	$424,658

Doctorate Recipient Follow-up	2,000	2 minutes	67	$1,742
Institutional Contacts	530	21 hours	11,130	$267,120
Proposed Methodological Research with Respondents	60	2 hours	120	$3,120
Proposed Methodological Research with Institutional Contacts	530	10 minutes	88	$2,112
Total SED Annual Burden			27,738	$698,752

A.13. Cost Burden to Respondents

There is no cost to the SED respondents other than the burden hour cost noted in A.12. Respondents need not purchase, operate, or maintain capital equipment, software, or storage facilities.

A.14. Cost to the Federal Government

The cost to the Federal Government for this annual data collection is approximately $3 million a year. This amount was based on the negotiated contract cost for the 2008 SED.

A.15. Program Changes or Adjustments

The respondent burden has increased from the last OMB clearance submission. Some changes to survey questions (detailed in Attachment 2) are estimated to produce a small increase in the estimate of the time required for survey completion from 19 minutes to 20. In addition, there has been an increase in the number of research doctorate awards (from 39,364 to 49,000) since the last OMB clearance submission which has increased the burden hours. This will be the first time that the burden hours on the institutional contacts has been included in the total SED burden estimate, resulting in a large increase in the total burden. In addition, burden hours for future methodological studies were projected and included in this request.

A.16. Tabulation and Publication Plans and Project Schedule

The results of the SED will be disseminated in a number of ways. To release the data, an NSF InfoBrief will be published. Then the NSF Detailed Statistical Tables report (containing a set of approximately 10 detailed statistical tables from the survey) will be released. These tables will be descriptive in nature and will provide extensive information on the education and employment plans of S&E doctoral graduates by field of study, granting institution, degree, future occupational and postdoctoral training plans, and demographic characteristics.

The six Federal agencies participating in the SED sponsor the compilation of survey results on all fields of study. The survey contractor prepares an interagency summary report for release by NSF and the other federal sponsors. The interagency Summary Report is provided free of charge

to responding institutions and is available via the Web, the address of which is noted in the SED questionnaire.

The SED data will also be used in the development of key NSF reports, including the Congressionally mandated reports *Science and Engineering Indicators* and *Women, Minorities, and Persons with Disabilities in Sciences and Engineering*. All of these publications, plus additional detailed tables, will become available on NSF's Science Resources Statistics (SRS) Web site.

There are also plans to include SDR variables taken from the SED data in the SESTAT (Science and Engineering Statistical Data System), which resides on the Web. The SESTAT system, described in section A.4, can be used to produce tabulations from the component surveys, providing a rich resource to those within and outside the government. As noted above, microdata will also be disseminated to Federal co-sponsors and collaborating researchers (with current data use licenses) so that they can conduct specialized studies. These, in turn, are expected to result in reports and other publications that further disseminate the data. Finally, it is anticipated that substantive analyses of the SED data will be presented at appropriate professional meetings, such as the annual meetings of the Association for Institutional Research, the Council of Graduate Schools, the American Educational Research Association, the American Statistical Association, the American Economics Association, the American Sociological Association, etc.

The SED project schedule is compressed. The 2010 survey covers the period from July 1, 2009 to June 30, 2010. The questionnaires will be mailed to the graduate schools upon OMB approval (anticipated in April 2009), for continuous distribution to graduate students as they complete the requirements for their doctorate. Returned survey questionnaires are edited and coded until survey close-out, which is April 2011 for the 2010 academic year (SED 2010). After the survey close-out, data variables are constructed, edited, evaluated, and reviewed for trend consistency in May/June. In July, the file is further evaluated and quality control checks are made. Data are tabulated in August and prepared for publication by October. Aggregate data are made available to the public in November via an online data release by NSF.

Project Schedule

The 2010 SED survey schedule follows. The 2011 SED survey schedule is expected to be similar except lagging by one year.

Phase	**Time**
Receive OMB clearance approval	April 10, 2009
Mailing of new forms to graduate deans	April 2009
Forms distributed to graduates	July 2009 – June 2010
Data collection close-out	April 15, 2011
Preparation of data file	July 2011
Production of tabulations	August 2011
Release of data by NSF	October 2011
Interagency Summary Report released	November 15, 2011

A.17. Display of OMB Expiration Date

The OMB Expiration Date will be displayed, as indicated.

A.18. Exception to the Certification Statement

The 2010-11 SED will comply with the certification statement on form OMB 83 1.

SECTION B: Collection of Information Employing Statistical Methods

Survey Data Collection Procedures Background

The Survey of Earned Doctorates questionnaire is distributed to new doctorate recipients by the Graduate Deans of the approximately 420 doctorate granting institutions, and approximately 530 independent programs within those institutions, in the United States. The SED questionnaires (either paper or web) are filled out at the time the individuals complete all requirements for their doctoral degrees and are returned to NSF's survey contractor by the Graduate Dean. Because doctorates complete the requirements for graduation throughout the year, the questionnaire distribution and completion process is continuous.

The institution (usually the Graduate Dean's office) is the main SED interface with the doctorate recipient and experience shows that the process is highly effective. The distribution of the questionnaire by the university itself, the clear nature of the questionnaire, and the cooperation of the Graduate Deans all combine to keep survey response rates around 92 percent.

When the completed paper survey questionnaires are received by the survey contractor, they are entered directly into the survey contractor's Computer-Assisted Data Entry (CADE) program. This system permits edits (for completeness, consistency, valid ranges, etc.) during data entry. Surveys received via the Web do not need to be keyed, but do receive edit checks. Errors which can clearly be remedied are corrected immediately; any questionnaire failing the edit for critical items will have a follow up Missing Information Letter (MIL) or e-mail generated for the respondent. The MIL attempts to gather missing data on a maximum of eight items (year of Master's, year of Bachelor's, postgraduation location (state or country), birth date, citizenship status, race, ethnicity, and gender), though most MILs address fewer than eight.

The survey contractor works with Institutional Contacts to obtain contact information for students who have not submitted their SED questionnaires. An Address Roster is sent to Institutional Contacts asking for the addresses of the nonresponders. Sometimes the IC's can provide other basic data items, as well as the addresses. The survey contractor also utilizes Web-based locating sites to locate contacting information for nonresponders. A series of letters is sent to any graduate who did not complete the survey through their graduate school, requesting their participation and containing PIN/password for web access plus a paper questionnaire.

Finally, any graduate who does not complete the SED through their graduate school and does not return a survey through the non-respondent mailing effort is given the opportunity to complete a slightly shortened version of the survey over the telephone. If by survey close-out an individual has not responded, public information from the commencement programs is used to construct a skeletal record on that individual. The skeletal record contains the name, PhD institution, PhD field, degree type, calendar year that the doctorate was earned, month that the doctorate was earned and usually the sex of the doctorate earner. If a survey questionnaire is later received from a previous non-respondent, the skeletal record is replaced by the information provided by the respondent.

B.1. Universe and Sampling Procedures

The SED is a census of all students receiving a research doctorate between July 1 and June 30 of the following year. Because it is a census, no sampling is involved. All institutions identified in IPEDS as granting doctoral degrees are asked to participate IF they confer "research doctorates" and if they are accredited by one of the regional accreditation organizations recognized by the Department of Education. If so, the schools are asked to distribute survey questionnaires, or cooperate in the electronic distribution the self-registration link, to their research doctoral recipients at the time of graduation. The SED maintains the universe of research doctorate-granting institutions each year by comparing the list of research granting institutions from IPEDS against the schools participating in the SED. If a new institution is found to be offering a research doctorate, they are contacted and added to the SED universe.

A high rate of response is essential for the SED to fulfill its role as a key part of the universe frame for longitudinal sample surveys, such as the Survey of Doctorate Recipients, and as the only reliable source of information on very small groups (racial/ethnic minorities, women, and persons with disabilities) in specialized fields of study at the Ph.D. level.

The feasibility of conducting the Survey of Earned Doctorates on a sample basis, and the utility of the resulting data, have been considered and found to be unacceptable. First, it is highly unlikely that the 530 graduate offices that distribute the SED questionnaire voluntarily could be expected to effectively carry out a sampling scheme such as handing out the questionnaire to every fifth doctoral candidate. In addition, one of the reasons many institutions participate in the survey is to receive complete information about all of their doctorate recipients and to be able to make comparisons with peer institutions.

A second sampling option – a mailing to doctorate recipients AFTER graduation – would likely result in a much lower response rate because of difficulties in obtaining accurate addresses of doctorate recipients, particularly the foreign citizens who represent an ever growing proportion of the doctorates recipient universe each year. Such a technique would impose on the universities the additional burden of providing current addresses of new graduates, a somewhat ineffective process because experience with mailing surveys to new doctorates shows many addresses are outdated almost immediately after graduation.

A third alternative, sending the questionnaire to doctorate recipients at a selected subset of institutions, would result in only a marginal decrease in respondent burden because the largest universities, all of which would need to be included in such a scheme, grant a disproportionate number of doctoral degrees. For example, the 50 largest institutions annually grant 51 percent of all doctoral degrees. Application of these sampling techniques would reduce both the utility of the data and the overall accuracy of the collected data. Matrix or item sampling – a widely used technique in achievement testing – would not be feasible because the characteristic information is needed for each doctorate recipient for use in selecting the sample for the follow-up SDR. It would reduce the utility of the information to request, for example, sex, or race, or field of

degree information for some doctorate recipients and not for others. These characteristics are not evenly distributed across the doctorate population, and the extensive uses made of the data base rely on the completeness and accuracy of the information on doctorate recipients.

Therefore, sampling doctorates would decrease the utility of the data while increasing burden on the Graduate Schools which administer the survey and decrease the incentives for the institution to participate.

B.2. Survey Methodology

Because there is no sampling involved in the SED, there has traditionally been no weighting necessary. Basic information about non-responding individuals is obtained, where possible, from public records at their graduating institutions, graduation lists, etc. Both unit and item nonresponse are handled by including categories of "unknown" for all variables in tabulated results. The statistical experts associated with this survey are Colm O'Muircheartaigh, Vice President of Statistics and Methodology at NORC (312-759-4017) and Rachel Harter, Senior Statistician on the project at NORC (312-759-4058). At NSF, Mark Fiegener, Project Officer for this survey (703-292-4622) and Stephen Cohen, SRS Chief Statistician, (SRS) (703-292-7769), will provide statistical oversight.

B.3. Methods to Maximize Response

The SED has enjoyed a high response rate during its existence, with an average of 92% completions over the past 30 years. It owes this high rate, in part, to the use of the data by the Graduate Deans who go to extraordinary lengths to encourage participation on the part of their graduates. Each Graduate Dean receives a profile of their graduates, compared with other institutions in their Carnegie class, soon after the data are released each year. It is also due to extensive university outreach efforts on the part of the survey contractor, NORC, and National Science Foundation staff and to the importance the universities themselves place on the data.

Throughout the data collection period, schools are constantly monitored for completion rates. Data on doctorates awarded on each commencement date are compared to data from the previous round in order to flag fluctuations in expected returns. Schools with late returns or reduced completion rates are individually contacted. Site visits, primarily to institutions with low response rates, by NSF staff and survey contractor staff are also critical to maintaining a high response rate to this survey. NORC's electronic monitoring systems are particularly important to these efforts, as each institution's graduation dates or SED submission dates can vary from monthly to annual.

In addition to the broad efforts to maintain high completion rates, targeted efforts to prompt for missing surveys and critical items are also key. The survey contractor works with Institutional Contacts and also utilizes Web-based locating sites to contact students by mail and e-mail for missing surveys or items. A Missing Information Roster is sent to Institutional Contacts who can sometimes provide basic items, in addition to addresses. A series of letters is sent to any

graduate who did not complete the survey through their graduate school, requesting their participation and containing PIN/passwords for web access plus paper questionnaires are sent to non-responding students. Additionally, a Missing Information Letter (MIL) is sent to any respondent who did not provide one of the "critical items" on the survey. This letter asks for the missing item and provides both a return envelope and a special e-mail address where the response can be sent. Finally, any non-respondent who does not complete the SED through their graduate school and does not return a survey through the non-respondent mailing effort is given the opportunity to complete a slightly shortened version of the survey over the phone. Data received via the different modes are merged and checked to avoid duplicate requests going out to the various sources. The results of these varied efforts significantly increase the number of completions as well as reduce the number of missing critical items, thereby improving the quality of the SED data.

The response rates of institutions as well as the response rates to questionnaire items are evaluated annually. For example, the evaluation of the response rate for 2007 indicated that over half of the non-response was due to 21 institutions. Institutions with poor response rates were targeted for special letters or site visits by NSF or survey contractor staff and, to a large extent, these efforts have been successful in raising the response rates at institutions.

B.4. Testing of Procedures

The SED has undergone extensive review and testing of the questionnaire and the methods employed in conducting the survey, and there has been extensive outreach about the uses of the data by the SED's stakeholders. The changes made to the SED 2010 survey version are a result of many activities which have helped inform changes to instruments and procedures over time. The following major activities have conducted since the previous OMB clearance submission (see Attachment 9.1 for a list of the methodological studies conducted over the past 15 years). The NSF project officer will be pleased to provide any of the documents referred to in this section or those referred to throughout the supporting statement.

Questionnaire Review and Research

Cognitive interviews were conducted to explore the effects on respondents of item format changes from the 2007 and 2008 SED questionnaire instruments, and to gauge respondents' understanding of the term "interdisciplinary". Twenty respondents were brought in and asked to complete the 2009 version of the questionnaire, then probed for their reaction to specific questions. They were also asked to compare the 2009 version to a mock-up version of the possible 2010 questionnaire, and asked to rate their preference on any changes that were made. The recommendations based on the results of these cognitive interviews are incorporated within the SED 2010 questionnaire, and will expand the questionnaire from a 10-page to a 12-page survey. (See Attachment 9.3.)

Data Collection Related Tests

The accuracy of the data from the Survey of Earned Doctorates has been one of its strongest assets. An ongoing evaluation of the accuracy of coding, editing, and data entry processes is conducted. It consistently indicates that the error rate is very low (less than one percent). During data collection, the frequency distribution of variables is monitored on a continuous basis, so that emerging problems, such as high item non-response rates, can be identified early in the data collection phase and appropriate corrective measures implemented, if necessary. Additional quality control checks on the merger of paper and electronic questionnaires as well as the merger of missing information into the master data base are also ongoing. The survey questionnaires are constantly compared with the universities' graduation lists and commencement programs to make sure that only those persons with earned research doctorates are included.

Survey Methodology Tests and Research

Several tasks were completed since the last OMB package, and include several that informed the recommendations for the next cycle. These tasks ranged from continuous assessments of everyday processes to overarching reviews of the institutions and degrees included in the survey to confirm the completeness and accuracy of the SED universe.

The following tasks are done regularly throughout each survey round:

- Review of systems, programming, and quality control data preparation processes with a goal of earlier release of the data;
- Merging data on a flow basis to identify and correct data inconsistencies and reduce the amount of time between the close of data collection and the release of the data.

These tasks are done annually, prior to the beginning of data collection or to the start of data preparation:

- Comparison of the IPEDS database of doctorate-granting institutions to the SED universe to identify institutions newly offering doctorate programs that are not currently in the SED;
- Review of the IPEDS database to determine if any institutions currently participating in the SED are offering eligible degrees that are not currently being included;
- Discussion of possible improvements in the coding and editing processes to ensure faster data entry resulting in more timely follow-up with non-respondents;
- Consultation with data processing managers on issues of paper and electronic data handling and mergers;
- In-depth analysis of confidentiality issues, particularly of data products that will be publicly available;
- Coordination of items common to the SDR and SESTAT instruments (see section A.4).

The following tasks are completed annually at the end of each data collection period. The results are compiled and reviewed before each new OMB clearance cycle to inform possible changes:

- Extensive reviews of unit and item-by-item frequencies ;
- Item analysis for floor and ceiling effects;
- Review of all respondent comments for concerns over confidentiality or item improvements;

- Detailed review of emerging and declining fields of study and alignment with the CIP (Classification of Instructional Programs);
- Review of "other, please specify" information in consideration of expanding or changing answer options;
- Coordination of items common to the SDR and SESTAT instruments (see section A.4).

Finally, the following tasks were conducted during the last OMB clearance cycle, and will be conducted periodically in the future:

- conduct of cognitive interviews, as noted above, with doctoral students from various disciplines;
- specific analysis of the items changed during the prior cycle (in the case of this package, changes made to the 2007 questionnaire) ;
- review of the non-PhD doctorate degrees included in the SED to confirm that they are research degrees and thus eligible for the survey;
- extensive literature reviews on targeted topics, such as disclosure avoidance and other confidentiality issues, as well as the history and contemporary purpose of the Doctorate of Education (Ed.D).

Research of SED Data Needs and Uses

SRS conducted a series of eight outreach meetings in fall of 2008 to learn about the specific data needs and uses of institutions, associations, and organizations that make extensive use of the SED's race/ethnicity and gender data. The meetings provided important input to SRS's efforts to redesign the statistical tables that report this information. SRS also conducted a web survey during this same period to gather similar information from the SED data user community.

Proposed Tests and Research

Over the course of the proposed OMB cycle (April 2009 – April 2012), the SED anticipates conducting several methodological tasks that would involve both SED respondents and the Institutional Contacts (ICs) at participating institutions. The burden hours for these tasks are included in Section A.12. Proposals for these additional tests are still under consideration. These will be submitted for OMB approval prior to implementation.

The SED anticipates conducting focus groups and/or cognitive interviews with potential or already existing SED respondents over the next 3 years. One set of interviews would involve the web survey. Dr. Don Dillman ~~will be~~ has begun ~~undertaking~~ an expert review of the SED web survey ~~in the upcoming year, with~~and provided some recommendations for changes to the design of individual questions. Some of these recommended changes have been incorporated in the 2010 SED questionnaire (see Attachment 1) and are included in the list of changes in Attachment 2. Dr. Dillman will continue his expert review and may have further question changes as well as recommendations about ~~and possibly~~ the web survey methodology and its administration as a whole. Interviews may be conducted with respondents to gauge their reaction to these changes, their reaction to the "Field of Study" lists on the paper survey versus the web survey, or other possible mode effects.

Additionally, another set of cognitive interviews may be conducted prior to the next OMB submission (for the 2012-2013 survey rounds) to test any changes to the questionnaire that would be recommended in the next review. The recommended changes would be based on analysis of the data from previous rounds to identify problem questions or emerging trends not being captured by the current instrument.

Finally, the SED anticipates conducting a web survey of all the ICs from participating institutions. This short survey would collect information on the specific practices involving the conduct of the survey at the institutions in the effort to identify new technologies, practices, or trends that impact the SED. The goal of this survey would be to identify areas where the SED can better support the ICs and adjust practices to meet the changing needs of the graduate schools.

The draft SED questionnaire was reviewed by Federal sponsors in November of 2008, and the final questionnaire was reviewed and then approved by the sponsors in January of 2009. (See Attachment 5 for the list of persons who were consulted or who reviewed the questionnaire.) See Attachment 2 for a list detailing changes made to the SED 2010 questionnaire from the 2009 version and the rationales for those changes.

B.5. Individuals Consulted

NORC at the University of Chicago is the organization contracted to collect and analyze the SED data for the 2010-2011 survey rounds. Staff from NORC who have consulted on the aspects of the design are listed in Attachment 5.

Additional individuals both inside and outside of NSF who have consulted on the statistical and methodological aspects of the design are also listed in Attachment 5.