

# **SF-83 SUPPORTING STATEMENT**

**for**

**Survey of Earned Doctorates**

**2020 and 2021 survey cycles**

**Section B**

## TABLE OF CONTENTS

<b>SECTION B: COLLECTION OF INFORMATION EMPLOYING STATISTICAL METHODS.....</b>	<b>19</b>
B.1. Universe and Sampling Procedures.....	19
B.2. Information Collection Process.....	20
B.3. Methods to Maximize Response.....	21
B.4. Testing of Procedures.....	22
B.5. Individuals Consulted.....	23

### LIST OF ATTACHMENTS

Attachment 1:	2020 SED Questionnaire
Attachment 2:	Current Representatives from Sponsoring Agencies
Attachment 3:	Authorizing Legislation of Sponsoring Agencies
Attachment 4:	Example of SED Institutional Profile
Attachment 5:	First Federal Register Announcement
Attachment 6:	NSF Staff and Contractor Data Use Agreement for Individuals
Attachment 7:	SED Institution Communication Materials Job Aids for Institution Contacts Graduation List Template Address Roster Form Missing Information Roster
Attachment 8:	SED Nonrespondent Communication Materials
Attachment 9:	SED Methodological Research

## SECTION B: Collection of Information Employing Statistical Methods

### B.1. Universe and Sampling Procedures

The Survey of Earned Doctorate (SED) is a census of all students receiving a research doctorate between July 1 and June 30 of the following year. All institutions identified in the U.S. Department of Education’s Integrated Postsecondary Education Data System (IPEDS) as granting doctoral degrees are asked to participate *if*:

- (1) they confer “research doctorates” and
- (2) they are accredited by one of the regional accreditation organizations recognized by the U.S. Department of Education.

A “Research doctorate” is defined as a doctoral degree that requires the completion of an original intellectual contribution in the form of a dissertation or an equivalent culminating project (such as a musical composition), and that is not primarily intended as a degree for the practice of a profession. In 2016, 98.1% of awarded research doctorates were PhDs; in 2017, PhDs comprised 98.0% of awarded research doctorates.

Participating schools distribute the link to the online SED questionnaire or distribute the SED paper questionnaires to their research doctoral recipients during the final academic term prior to graduation. SED maintains the universe of research doctorate granting institutions by comparing the list of institutions from IPEDS against the institutions participating in the SED annually. If a new institution is found to be offering a research doctorate, the institution is contacted and, based on eligibility criteria, added to the SED universe.

<b>Academic Year (1 July – 30 June)</b>	<b>Number of Institutions Reporting Graduates</b>	<b>Number of Research Doctoral Graduates</b>	<b>Response Rate*</b>
2015	445	55,006	90.3%
2016	436	54,904	91.9%
2017	428	54,664	91.4%
2018	~433	~55,554	~91.0%
2019	~447	~56,038	~91.0%
2020	~450	~56,522	~91.0%

\*This response rate represents the rate at which eligible doctorate recipients complete and return SED questionnaires.

A high response rate is essential for the SED to serve its role as the frame for the Survey of Doctorate Recipients (SDR), and as the only reliable source of information on very small populations (racial/ethnic minorities, women, and persons with disabilities) in specialized fields of study at the doctoral level.

The feasibility of conducting the SED on a sample basis, and the utility of the resulting data, have been considered and found to be inadequate. Many institutions participate in the survey to receive comprehensive information about all their research doctorate recipients and to make comparisons with peer institutions. Experience indicates that the 614 Institution Contacts (ICs) (as of the 2019 round) who annually distribute the SED would have great difficulty carrying out

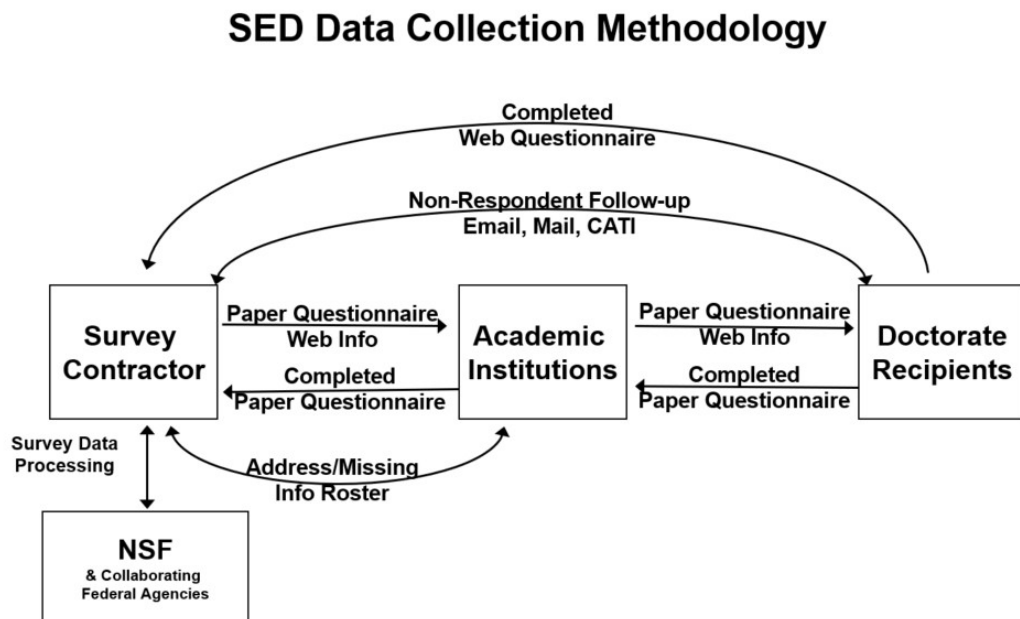
a sampling scheme. The current process is easy for ICs given that schools often refer all of their eligible doctoral students to an online graduation checklist where the SED is but one step in the graduation process. In addition, conducting the SED on a sample basis would produce poor estimates of small populations (e.g., racial/ethnic minorities) earning degrees in particular fields of study. Such data are important to a wide range of SED data users.

A second sampling option – a mailing to doctorate recipients *after* graduation – would likely result in much lower response rates. Obtaining accurate addresses of doctorate recipients is very difficult, particularly for foreign citizens, who represent an ever-growing proportion of the doctorate recipient universe each year. Although universities could help, they would incur the additional burden of determining current addresses for doctorates who have graduated. This is an ineffective process because the addresses of new doctorates are typically 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 research doctorate-granting institutions confer approximately 50 percent of all eligible degrees in a given year. Application of these sampling techniques would reduce both the utility and overall accuracy of the collected data, while increasing burden on schools that administer the survey.

Given that the SED is a census, weighting is not conducted. Missing information about non-responding individuals is obtained from public records and commencement lists where possible. Unit and item nonresponse are indicated by including categories of “unknown” for all variables in tabulated results.

## B.2. Information Collection Process



The SED (either web or paper) is distributed by ICs and completed by students in the academic term prior to the students' graduation. If paper questionnaires are completed, the IC returns them to NCSES's survey contractor. Because doctoral degrees are conferred throughout the year, questionnaire distribution and completion is a continuous process.

The institution or school, and commonly an employee in the graduate dean's office, is the main SED interface with the doctorate recipient. History shows that the interface is highly effective. The distribution of the survey by the university itself, the clear nature of the items, and the cooperation of the graduate deans, all combine to keep survey response rates above 90 percent.

When the completed paper questionnaires are received by the survey contractor, they are edited for completeness and consistency and then entered directly into a computer-assisted data entry (CADE) program. Surveys received via the web mode do not need to be data entered and are edited mainly through a series of pre-programmed skip patterns and range checks, which allow potential errors to be corrected immediately by the respondent.

The survey contractor sends an Address Roster to ICs at institutions with nonresponding students to obtain contact information these nonrespondents. The survey contractor also uses web-based locating sites to identify contact information for nonrespondents. Emails are sent to nonrespondents requesting their participation and providing a PIN and password for web access. (see Attachment 8 for a sample communication). Hard copy mailings with printed questionnaires and a link to the Web survey are sent to nonrespondents who do not respond to the emails or for whom we have no email address.

Finally, nonrespondents are given the opportunity to complete a shortened version of the survey over the telephone via a computer-assisted telephone interviewing (CATI) system. If by survey close-out an individual has not responded, public information from commencement programs or other publicly accessible sources is used to construct a skeletal record on that individual. The institution may also be asked to help provide missing data to complete skeletal records for these nonrespondents. The skeletal record contains the name, PhD institution, PhD field, degree type, month and year that the doctorate was earned, and the sex of the doctorate recipient. If a survey questionnaire is later received from a previous nonrespondent, the skeletal record is replaced by the information provided by the respondent.

In 2017, 94.9% of surveys were completed via the web, 2.7% via the paper questionnaire, and 2.4% via CATI.

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

The SED has traditionally obtained a high response rate, with an average of above 90% over the past 30 years. It owes this high rate, in part, to the use of the data by the ICs and graduate deans, who go to great lengths to encourage participation on the part of their graduates. Soon after the data are released each year, each graduate dean receives a profile of their graduates, known as the Institutional Profile, compared with those of other institutions in the same Carnegie class.

In addition to the importance the universities themselves place on the data, the high response rate is also due to university outreach efforts pursued by NCSES. Throughout the data collection period, school participation is constantly monitored. Doctorates awarded each commencement date are compared to data from the previous round, and fluctuations in expected returns are

flagged. Schools with late returns or reduced completion rates are individually contacted. Staff site visits, primarily to institutions with low response rates, are also critical to maintaining the SED's consistently high response rate.

Along with the broad efforts to maintain high response, targeted efforts to prompt for missing surveys and critical items are also key. The survey contractor works with ICs and also uses web-based locating sites to contact students by email and mail. A series of contacts is sent to any graduate who did not complete the survey through their graduate school, requesting their participation and including a PIN and temporary password for web access, as well as a paper questionnaire when postal mailings are sent. Additionally, nonrespondents are given the opportunity to complete a slightly shortened version of the survey over the phone.

Finally, a Missing Information Roster (MIR) is sent to ICs who can sometimes provide critical item information (sex, race/ethnicity, citizenship, etc.) in addition to addresses. 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 SED data.

Institutional and individual response rates are evaluated annually. Institutions with poor response rates are targeted for conference calls or site visits to discuss their procedures and potential improvements for achieving a higher response rate. Such efforts typically have been successful in raising response rates.

#### ***B.4. Testing of Procedures***

In recent years, NCSES has conducted extensive review of the SED survey methods and testing of the SED questionnaire. Attachment 9 summarizes methodological research that has occurred since 2016.

#### ***Data Collection and Processing Tests***

During data collection, item nonresponse is continuously monitored to ensure that emerging problems can be identified early, and appropriate corrective measures implemented. Completed questionnaires are constantly compared with the universities' graduation lists and commencement programs to ensure that only questionnaires for persons with earned research doctorates are included. Quality control checks related to processing for paper and electronic questionnaires and missing information are also continuous. Overall, NCSES conducts a continuous evaluation of the accuracy of SED coding, editing, and data entry processes. The results consistently indicate that the error rate is very low, less than 1 percent. Measurement error in the SED is attributable to errors that occur during data processing. Data entered by respondents regarding their educational history, including institution and field of study, are cleaned and standardized by trained coders. Average error rates for all 2017 cases are 0.24% for institutions and 0.58% for fields of study. (less than one percent).

#### ***Survey Methodological Research***

Over the course of the upcoming OMB clearance cycle, NCSES anticipates conducting methodological research tasks and analyses of data user needs under the Generic Clearance of Survey Improvement Projects (OMB #3145-0174), as needed. In addition, in 2018, two expert survey methodologists, Dr. Jolene Smyth and Dr. Kristen Olson of the University of Nebraska at

Lincoln, were asked to review the education history section of the SED survey instruments for potential improvement to make it easier for the respondents to provide information about their degrees. NCSES is currently reviewing the suggested changes and will submit any questionnaire changes for OMB approval at a later time.

### ***B.5. Individuals Consulted***

At NCSES, Kelly Kang, Project Officer for the SED (703-292-7796), John Finamore, Program Director for Human Resources Statistics (703-292-2258), and Samson Adeshiyan, Chief Statistician (703-292-7769), provide oversight for the survey.

The survey experts from the SED data collection contractor, RTI International, associated with the SED are Patricia Green, Project Director (312-456-5260), Stephanie Eckman, Methodological Task Leader (202-974-7816), Ruth Heuer, Analysis and Reporting Task Leader (919-541-6457), Kaleen Healey, SED Digest (312-777-5210), Saki Kinney, Confidentiality Review (919-316-3129), and David Wilson, Statistical Task Leader (919-541-6990).