

# **SF-83 SUPPORTING STATEMENT**

**for**

**Survey of Earned Doctorates**

**2022 and 2023 survey cycles**

**Section B**

## TABLE OF CONTENTS

<b>SECTION B: Collection of Information Employing Statistical Methods.....</b>	<b>22</b>
B.1. Universe and Sample Descriptions.....	22
B.2. Information Collection Procedures.....	23
B.3. Statistical Accuracy.....	24
B.4. Testing of Procedures.....	25
B.5. Contact Information.....	25

## LIST OF ATTACHMENTS

Attachment 1:	2021 SED Questionnaire Reference Copy
Attachment 2:	Current Representatives of the SED Sponsoring Agencies
Attachment 3:	Authorizing Legislation of the SED Sponsoring Agencies
Attachment 4:	SED Institutional Profile - Example
Attachment 5:	2022-23 SED First Federal Register Notice
Attachment 6:	NSF NCSES Individual Data Use Agreement for NSF Staff and Contractors
Attachment 7:	SED Institution Contact Materials
Attachment 8:	SED Nonrespondent Follow-Up Contact Materials
Attachment 9:	SED Recent Methodological Research
Attachment 10:	Proposed Revisions to COVID-19 Question Module

## SECTION B: Collection of Information Employing Statistical Methods

### B.1. Universe and Sample Descriptions

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.

The SED universe also includes a small number of institutions that grant research doctorates that are not mandated to report to IPEDS because they do not participate in Title IV Federal financial assistance programs.

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 2019, 98.4% of awarded research doctorates were PhDs.

Participating schools distribute the link to the online SED Web instrument to their research doctorate 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.

Academic Year (1 July – 30 June)	Number of Institutions Reporting Graduates	Number of Research Doctoral Graduates	Response Rate*
2015	445	55,006	90.3%
2016	436	54,904	91.9%
2017	428	54,664	91.4%
2018	431	55,195	92.1%
2019	438	55,703	92.1%
2020**	452	55,258	91.5%

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

\*\*The number of doctoral graduates and response rate for 2020 SED are not yet final.

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 approximate 600 Institution Contacts (ICs) (as of the 2020 round) who annually distribute the SED web survey to the doctorate graduates 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 doctoral student who have graduated. This is an ineffective process because the addresses of new doctorate recipients 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 Procedures***

The link to the SED Web instrument is distributed by ICs at each institution and completed by students in the academic term prior to the students’ graduation. Because doctoral degrees are conferred throughout the year, questionnaire distribution and completion are 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.

The survey contractor sends an Address Roster to ICs at institutions with nonresponding students to obtain contact information for 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 mailing of letters with 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.

Nonrespondents are also given the opportunity to complete a shortened version of the survey over the telephone via a computer-assisted telephone interviewing (CATI) system.

A few months prior to the end of data collection, the survey contract sends a Missing Information Roster (MIR) to ICs to obtain information on a small number of critical SED data items for nonrespondents from their institution. 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 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 is later completed by a previous nonrespondent, the skeletal record is replaced by the information provided by the respondent.

In 2020, 97.3% of surveys were completed via the web, and 2.7% via CATI. Starting in the 2020 survey cycle, the paper questionnaires was not used in data collection and were available only as a reference copy.

### ***B.3. Statistical Accuracy***

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. Thus, information collected by SED are reviewed and used by participating institutions, and if anomalies are noted, the contractor can review the data. The high response rates and the use of the data by institutions result in accurate, reliable data.

In addition to the importance the universities themselves place on the SED 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 SED survey contractor works with ICs and also uses web-based locating sites to contact graduates 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. Additionally, nonrespondents are given the opportunity to complete a slightly shortened version of the survey over the phone.

Finally, a MIR provided by ICs include critical item information (sex, race/ethnicity, citizenship, etc.) for some nonrespondents. 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 review of the SED data collection methods and testing of the SED questions. Attachment 9 summarizes methodological research studies that have occurred since the last SED survey clearance in 2019.

In December 2019 and January 2020, a methodological study was conducted under the NCSES Generic Clearance of Statistics Improvement Projects (OMB No. 3145-0174) to improve the quality of education data collected in the SED while also reducing respondent burden. Based on the study results, methods for collecting the educational history and the field of study questions were changed in the 2021 SED.

In April 2020, NCSES requested and obtained an approval from OMB to add a question module designed to measure the immediate challenges encountered by the doctoral students from the impact of Coronavirus Disease of 2019 (COVID-19) pandemic on their graduate experience and career plans. The COVID-19 question module was revised based on the early results from the 2021 SED data collection, and in December 2020, NCSES submitted a request for OMB approval to conduct cognitive interviews to further test the revised questions with approximately 60 individuals under the NCSES generic clearance. The revised COVID-19 questions are listed in Attachment 10. The final set of questions for inclusion in the 2022 SED and the cognitive testing results will be submitted to OMB in March 2021.

As noted in Section A.8., during the 2021 SED survey cycle and beyond, NCSES will be engaged in several important research efforts to develop and evaluate questions on sexual orientation and gender identity (SOGI) for their potential use in the SED and other relevant NCSES surveys. Given the unique aspects of the SED and its respondents, NCSES anticipates additional testing may be required. For example, because SED has a greater proportion of temporary visa holders than other federal or NCSES surveys, additional cross-cultural testing may be needed. Also, because the SED respondent data, including personally identifiable information, is routinely provided to their doctorate institutions, data privacy and informed consent statements may need to be examined and revised before implementation, to ensure SED's traditionally high unit and item response rates are maintained. NCSES will submit requests for these additional studies as needed.

No other methodological testing is proposed at this time. If during the course of this clearance additional methodological studies are needed, a separate request will be made under the NCSES generic clearance.

#### ***B.5. Contact Information***

At NCSES, Kelly Kang, Project Officer for the SED (703-292-7796), John Finamore, Program Director for Human Resources Statistics and acting Chief Statistician (703-292-2258) 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, Data Processing Task Leader (919-

541-6457), Jon Gordon, Analysis Task Leader (770-407-4952), Saki Kinney, Statistical Task Leader (919-316-3129), and David Wilson, Statistician (919-541-6990).