

# **SF-83 SUPPORTING STATEMENT**

**for**

**Survey of Earned Doctorates**

**2024 and 2025 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.....                                                  | 26        |

## LIST OF ATTACHMENTS

|                |                                                                       |
|----------------|-----------------------------------------------------------------------|
| Attachment 1:  | 2024 SED Questionnaire - Draft                                        |
| 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:  | 2024-25 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: | SED SOGI Experiment Plan                                              |

## 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 2021, 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<br>(1 July – 30 June) | Number of Institutions<br>Reporting Graduates | Number of Research<br>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                                | 448                                           | 55,703                                   | 92.1%          |
| 2020                                | 449                                           | 55,283                                   | 92.1%          |
| 2021                                | 448                                           | 52,250                                   | 91.5%          |
| 2022**                              | 457                                           | 57,601                                   | 91.6%          |

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

\*\*The response rate for 2022 SED is as of January 10, 2023.

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. Experience indicates that many of the Institution Contacts (ICs) who work with the SED staff to distribute the SED Web survey to the doctorate graduates at their institution would have 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, collecting data from a sample of doctorate recipients would reduce data quality and utility. Many institutions participate in the survey to receive comprehensive information about all their research doctorate recipients and to make comparisons with peer institutions. 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 point of contact with the doctorate recipient. History shows that this contact is highly effective. The distribution of the survey by the university itself, the high-quality questionnaire, and the cooperation of the graduate deans, all combine to keep survey response rates above 90 percent.

The survey contacts consist of the following:

- The survey contractor asks the ICs at institutions to provide the graduation lists. (See Attachment 7 for the institution contact materials.)

- The survey contractor sends an Address Roster to ICs with nonresponding students to obtain their contact information.
- The survey contractor also uses web-based locating resources to identify the latest 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 the nonrespondent contact materials).
- 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 an individual has not responded by survey close-out, 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 the 2022 cycle, 98.5% of surveys were completed via the web and 1.5% were completed via CATI. Paper questionnaire was last used in the 2019 data collection and only available as a reference copy since 2020.

### ***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 is 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 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, the data collection methods for the educational history and the field of study questions were changed in the 2021 SED.

In April 2020, NCSES obtained an approval from OMB to add a question module designed to measure the impact of Coronavirus Disease of 2019 (COVID-19) pandemic on the graduate experience and career plans of doctoral students. The COVID-19 question module was revised based on the early results from the 2021 SED data collection, and in December 2020, OMB approved the NCSES to conduct cognitive interviews to further test the revised questions with approximately 60 individuals under the NCSES generic clearance. These study results led to the final set of COVID impact question wording and response options implemented in the 2022 and 2023 SED surveys. These questions will likely remain in the SED through 2027 (the median time to doctorate is seven years) to assess the impact of the pandemic on all stages of doctoral study. (See Attachment 9 for more information on the recent methodological research.)

As noted in Section A.8., since 2021 SED survey cycle and beyond, NCSES has 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, and the need to maintain SED's traditionally high unit and item response rates, NCSES obtained an approval from OMB to conduct cognitive testing of SOGI question with the SED target population, completing 61 cognitive interviews between May and July 2022. The SOGI questions tested were based on those used in earlier NCSES SOGI question testing efforts and the study focused on assessing question comprehension, potential confidentiality concerns, and sensitivities. NCSES incorporated probes to assess participants' likelihood of responding to the SOGI questions under different data sharing conditions.

NCSES plans to conduct an experiment within the 2024 SED survey to assess various SOGI questions to identify the question wording and response options that work best for the SED population and assess possible mechanisms for mitigating respondents' data confidentiality and privacy concerns. Due to the sensitivity of these data and because the SED data can be requested by the participating doctoral institutions, the collection and dissemination of these data provide a unique challenge for protecting the privacy and confidentiality of SED respondents. The

intended goal of this experiment is to identify a set of SOGI questions for inclusion in the 2025 SED. Attachment 10 describes this experiment in detail.

The SOGI question experiment in the 2024 survey cycle is the only methodological testing planned at this time. If during the course of this clearance additional methodological studies are needed, a separate request will be submitted to OMB under the NCSES generic clearance.

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

At NCSES, Kelly Kang, Project Officer for the SED (703-292-7796); Amber Levanon Seligson, Program Director for Science, Technology, and Innovation Survey Management; and John Finamore, 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 Peter Einaudi, Project Director (919-541-8765); Y. Patrick Hsieh, Lead Methodologist (312-777-5234); Ruth Heuer, Data Processing Task Leader (919-541-6457); Jon Gordon, Analysis Task Leader (770-407-4952); and Saki Kinney, Statistical Task Leader (919-316-3129).