

2023–2025  
Survey of Graduate Students  
and  
Postdoctorates in Science  
and Engineering

OMB Supporting Statement  
Section B

July 2023

## TABLE OF CONTENTS

Table of Contents.....	iii
EXHIBITS.....	iii
LIST OF ATTACHMENTS.....	iii
<b>B. COLLECTION OF INFORMATION EMPLOYING STATISTICAL METHODS.....</b>	<b>1</b>
B.1 Universe and Sampling Description.....	1
B.1.1 Discussion of Institutional Frame.....	1
B.1.2 Response Rates.....	1
B.2 Information Collection.....	2
B.2.1. Collection of Data Based on Updated CIP Codes.....	3
B.3 Statistical Accuracy of the Collection.....	3
B.3.1 Methods Used to Maximize Response Rates.....	3
B.3.2 Imputation Methods for Unit and Item Nonresponse.....	4
B.3.3 Accuracy and Reliability of Data.....	6
B.4 Testing of Procedures.....	6
B.5 Individuals Consulted.....	8

## EXHIBITS

Exhibit 9. Number of GSS Institutions, Schools, Units and Enrollment, 2019-2021.....	1
Exhibit 10. GSS Institution, School, and Unit Response Rates: 2019–21.....	2
Exhibit 11. Unit response rates: 2017–21.....	5
Exhibit 12. Item imputation rates for full-time or part-time graduate students, postdoctorates, and nonfaculty researcher items: 2021.....	5
Exhibit 13. Individuals Consulted on GSS Technical and Statistical Issues.....	8

## LIST OF ATTACHMENTS

Attachment 1.	America COMPETES Reauthorization Act of 2010
Attachment 2.	NSF Act of 1950
Attachment 3.	GSS Instrument Screen Shots
Attachment 4.	Federally Funded Research and Development Centers (FFRDCs) Postdoc Survey Instrument Screen Shots
Attachment 5.	GSS InfoBrief, January 2023
Attachment 6.	Research and Media Usage
Attachment 7.	First Federal Register Notice
Attachment 8.	GSS Schedule
Attachment 9.	List of Changes During Last Clearance
Attachment 10.	GSS Code List Complete List of GSS Eligible Fields and Codes
Attachment 11.	GSS Data Collection Plan
Attachment 12.	GSS Data File Upload Templates
Attachment 13.	GSS Worksheet
Attachment 14.	Imputation Section from GSS Methodology Report

Attachment 15. Nonfaculty Researcher Methodological Study Report

## **B. COLLECTION OF INFORMATION EMPLOYING STATISTICAL METHODS**

There are no proposed changes for GSS 2023-2025 or FFRDC 2023-2025. The procedures and methodology described in other sections are in current use on the GSS and the FFRDC postdoc survey and will be continued without change. Attachment 8 documents changes implemented in the past over the 2020-22 survey cycles.

### **B.1 Universe and Sampling Description**

The GSS is an annual census of eligible institutions. The GSS universe is intended to cover all academic institutions in the United States and its territories that grant research-oriented master's degrees or doctorates, appoint postdocs, or employ non-faculty researchers (NFRs) in science, engineering, and health (SEH) fields as of the fall term. An institution is considered eligible for the GSS if it grants at least one master's or doctoral degree in at least one program listed in a GSS-eligible field (see Attachment 9 for the list of GSS fields).

The FFRDC postdoc survey is a biennial census of all federally funded research and development centers. All FFRDCs recognized by the National Science Foundation are considered part of the survey universe. Coordinators are instructed to report all employed postdocs in eligible SEH fields.

#### ***B.1.1 Discussion of Institutional Frame***

The GSS frame is updated annually, in advance of data collection. The Integrated Postsecondary Education Data System (IPEDS) and other sources are used to identify new institutions and existing institutions that now offer graduate degrees in GSS-eligible fields. See Exhibit 9 for a comparison of the number of GSS institutions, schools, units, and enrollment in 2019-2021. The frame for the FFRDC Postdoc Survey derives from the FFRDC Master List, which NSF has maintained since 1967 (available at <https://www.nsf.gov/statistics/ffrdclist/>).

**Exhibit 9. Number of GSS Institutions, Schools, Units and Enrollment, 2019-2021**

<b>Year</b>	<b>Institutions</b>	<b>Schools</b>	<b>Units</b>	<b>Graduate enrollment by degree level</b>		
				<b>Total</b>	<b>Masters</b>	<b>Doctorate</b>
<b>2019</b>	714	809	20,249	690,117	408,228	281,889
<b>2020</b>	712	806	21,156	697,813	414,478	283,335

<b>2021</b>	699	787	21,365	760,156	466,613	293,543
-------------	-----	-----	--------	---------	---------	---------

Source(s):  
National Center for Science and Engineering Statistics, Survey of Graduate Students and Postdoctorates in Science and Engineering.

### **B.1.2 Response Rates**

Exhibit 10 displays unit, school, and institutional response rates for the 2019-2021 GSS survey cycles. Total response rates (includes both complete and partial response) from GSS 2021 were high: 96.1 percent at the institution level, 96.1 percent at the school level, and 98.2 percent at the unit level. Response rates are projected to remain stable during the next three cycles (2023-2025). Complete responses are those where the coordinator provided data for all requested items while partial responses are those in which the coordinator left some data missing that was later imputed (see section B.3.2 for more information on response types). All but 2 of the 43 FFRDCs responded to the 2021 survey, including all of the FFRDCs that employed postdocs in 2019. Thus, the overall FFRDC response rate was 95.3%, and the response rate for FFRDCs that employed postdocs was 100%. The 2019 response rate was 95.2% overall.

**Exhibit 10. GSS Institution, School, and Unit Response Rates: 2019–21**

	Total Response			Complete Response			Partial Response		
	2019	2020	2021	2019	2020	2021	2019	2020	2021
Institution	94.9% n=678	94.9% n=676	96.1% n=669	93.8% n=670	93.1% n=663	95.7% n=666	1.1% n=8	1.8% n=13	0.4% n=3
School	95.1% n=769	94.7% n=763	96.1% n=752	93.6% n=757	92.6% n=746	95.6% n=748	1.5% n=12	2.1% n=17	0.5% n=4
Unit	97.4% n=19,718	96.9% n=20,486	98.2% n=20,990	84.1% n=17,035	84.0% n=17,764	85.1% n=18,186	13.3% n=2,683	12.9% n=2,722	13.1% n=2,804

Source(s):  
National Center for Science and Engineering Statistics, Survey of Graduate Students and Postdoctorates in Science and Engineering.

### **B.2 Information Collection**

Each institution has one or more school coordinators (SCs) that manage data collection activities. Some institutions have separate coordinators for the graduate enrollment section and the postdoc section, and some have separate coordinators for the graduate and medical schools. Each GSS survey cycle begins with a pre-data collection e-mail to the previous survey cycle’s SC to determine if they are still the appropriate contact for the upcoming cycle. The e-mail is

typically sent in early September with a telephone follow-up if confirmation is not received before data collection is slated to begin. Upon the conclusion of the confirmation period data collection commences for all institutions. Data collection begins in October with an e-mail and FedEx package providing the SC with Web access information and information about the GSS-eligible degree programs.

For new institutions, NCSES e-mails the president of the institution a survey invitation letter that asks the president to name a SC for the survey and to verify the institutions' eligibility for the GSS. Institutions that do not respond to the letter are followed up via phone call and e-mail. Hard-copy GSS worksheets are provided to any new institutions to allow them to see the types of information requested in the survey.

The SC serves as the point of contact at the institution for all internal and external communications about the GSS. The SC may choose to delegate some reporting activity to unit respondents (URs) at their institutions or they may report the GSS data themselves. If using URs, the SC's responsibilities include notifying the URs of their assignments and ensuring that the UR submits the completed data by the established due date. The 2022 data collection plan, including a timetable and communications with GSS stakeholders, is included in Attachment 11.

The SCs are asked to prepare data files that can be uploaded directly into the GSS web survey instrument for the units that enroll graduate students, and/or employ postdocs or NFRs. The SCs are provided with survey variable and file specifications for each type of GSS data requested—graduate students, postdocs, and NFRs—as well as file templates (in the form of Microsoft Excel spreadsheets; see (Attachments 12a, 12b and 12c) to organize their data.

The SCs who are not ready or are unable to provide data through data upload method will be allowed to provide their data through the manual entry of requested data into a series of grids on the GSS web instrument. A hard copy of the GSS worksheet that corresponds to the GSS web instrument will be provided with the survey materials to the SCs, if requested (see Attachment 12). Information can be compiled on this worksheet for each unit prior to data entry into the web instrument.

The process of completing the FFRDC postdoc survey is similar to that of the GSS with exception to the availability of EDI and the use of URs because each FFRDC constitutes a single organizational unit.

### ***B.2.1. Collection of Data Based on Updated CIP Codes***

In 2017, the GSS began collecting disciplinary field data from institutions based on the Classification of Instructional Program (CIP) codes rather than NCSES's GSS codes. The collection of data by CIP codes has contributed to a reduction in response burden because these codes are commonly used at institutions. CIP is the academic field taxonomy used by the National Center for Education Statistics (NCES) for the IPEDS, a mandatory reporting requirement for institutions receiving Title IV funding. Schools still have the option of using either CIP codes or GSS codes for reporting postdoc and NFR data only. In 2021, approximately 83% of institutions were able to provide student data using CIP codes. FFRDCs report data using GSS codes as CIP codes are not applicable outside of academia.

## **B.3 Statistical Accuracy of the Collection**

### ***B.3.1 Methods Used to Maximize Response Rates***

Because the GSS is designed to produce estimates for all U.S. postsecondary institutions that offer graduate degree programs in SEH fields, care is made to maximize response rates and thus reduce the likelihood of biased estimates. The survey staff work closely with the SCs to build strong working relationships with all participating institutions and try to ensure that all contacts are positive.

Survey techniques proven successful in past surveys will again be used to maximize the GSS response rate. These techniques include:

- Early pre-data collection confirmation of the SC and their contact information
- Targeted e-mails and telephone follow-up based on response status
- Knowledgeable survey staff working at the GSS Help Desk to respond to questions and concerns and help the SCs and unit respondents via telephone and emails.
- Multiple modes of data collection allowed (two data upload options, web instrument)
- The inclusion of cover letters explaining how the provided data are used
- Enlistment of others at the institution, as appropriate, to gain cooperation

In addition to the methods listed above, a series of workshops and presentations were offered at professional conferences that GSS SCs frequently attend. To date, presentations and workshops have been offered at the Association for Institutional Research Annual Forum (AIR),

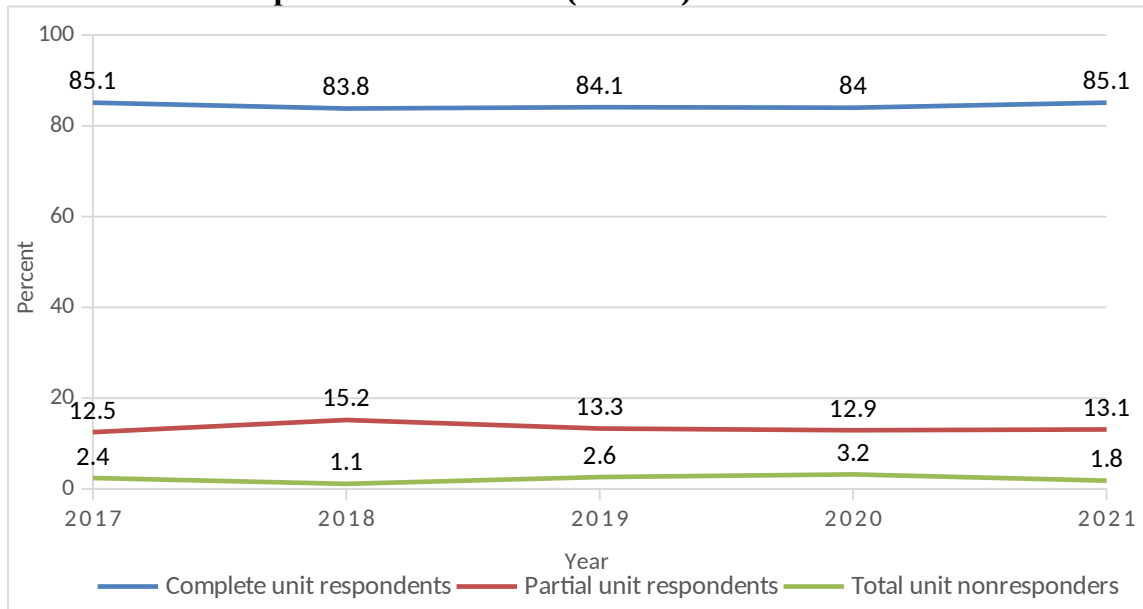
regional AIR conferences, and the Council on Graduate Schools annual meeting. Additionally, the GSS survey staff will conduct a series of webinars to provide specific instructions and support for SCs navigating the changes in the GSS data collection.

Many of these response techniques are deployed for the FFRDC postdoc survey as well, including helpdesk support, multiple phone and email contracts, and the use of site visits and virtual meetings to encourage participation.

### **B.3.2 Imputation Methods for Unit and Item Nonresponse**

*GSS Unit and Item nonresponse.* Units that provide a complete response to every item in the questionnaire are defined as *complete (unit) respondents* (needing no imputation), units that provide some but not all items are *partial (unit) respondents* (needing partial imputation), and units that provide no response are *total (unit) nonrespondents* (requiring complete imputation). Exhibit 11 shows the unit response rates by GSS survey cycle for 2017–21. The overall unit nonresponse rate for all units in the 2021 GSS is 14.9%, representing 3,179 units needing imputation. The 2021 percentage of total unit nonrespondents is 1.8%, and the percentage of partial unit respondents is 13.1%. The percentage of complete unit respondents is higher by 1.0% when compared to 2019. As seen in Exhibit 11, the unit response rates from 2017 to 2021 have remained steady for each response type.

**Exhibit 11. Unit response rates: 2017–21 (Percent)**





Source(s):  
National Center for Science and Engineering Statistics, Survey of Graduate Students and Postdoctorates in Science and Engineering.

The *item imputation rate* is a measure of the amount of missing data for each key total and grid detail variable collected on the GSS. For the 2021 GSS, all 543 data items required imputation. Exhibit 12 presents a summary of the item imputation rates for various classes of items: PT and FT master’s students, PT and FT doctoral students, PDs, and NFRs. Overall, item imputation rates ranged from 1.7% to 6.4%, with a mean item imputation rate of 3.9%, where 186 items had imputation rates between 1% and 3%, 256 items had rates between 3% and 5%, and 101 items had rates between 5% and 7%.

**Exhibit 12. Item imputation rates for full-time or part-time graduate students, postdoctorates, and nonfaculty researcher items: 2021 (Number and percent)**

Item imputation rate	Total	Part-time master’s students		Full-time master’s students		Part-time doctoral students		Full-time doctoral students		Postdoctorates		Nonfaculty researchers	
		N	%	N	%	N	%	N	%	N	%	N	%
Total items	543	30	100.0	158	100.0	30	100.0	158	100.0	152	100.0	15	100.0
≥ 1% and < 3%	186	30	100.0	63	39.9	30	100.0	63	39.9	0	0.0	0	0.0
≥ 3% and < 5%	256	0	0.0	0	0.0	0	0.0	95	60.1	150	98.7	11	73.3
≥ 5% and < 7%	101	0	0.0	95	60.1	0	0.0	0	0.0	2	1.3	4	26.7

Source(s):  
National Center for Science and Engineering Statistics, Survey of Graduate Students and Postdoctorates in Science and Engineering.

*GSS Imputation Methods.* The 2021 GSS collected 543 data items related to enrollment and financial support for master’s and doctoral full-time and part-time students, postdocs, and NFRs. All missing data within partial responding and total nonresponding units were imputed. To address the unit and item nonresponse, three methods of imputation were used including a carry-forward imputation method, a nearest neighbor method, and adjusted enrollment method based on IPEDS data. Details for these methods are documented in the 2020 GSS Methodology Report.<sup>1</sup> The imputation section from that report is included in Attachment 13.

<sup>1</sup> RTI. (2022). *2020 Survey of Graduate Students and Postdoctorates in Science and Engineering: Survey Methodology Report* Deliverable under contract NSFAC49100418F1261 Prepared by RTI International, April 2022.

*FFRDC Postdoc Survey Imputation Methods.* The 2021 FFRDC Postdoc Survey collected five distinct data items related to postdocs employed at FFRDCs. The item response rate was 100% among FFRDCs known to employ postdocs. One FFRDC did require partial imputation of demographic and field-level information. The lone FFRDC that required imputation was addressed through logical imputation. This approach replaced the percentages provided by the coordinator with estimated numbers of postdocs for demographic and field-level information.

### ***B.3.3 Accuracy and Reliability of Data***

Because both surveys are censuses with high response rates, and statistical imputation is conducted for nonresponse, the accuracy of these data is high. And both the GSS and the FFRDC postdoc survey have extensive review processes to check the consistency of each coordinator's data within and across years. For the GSS, the use of CIP codes for reporting helps ensure reliability as these codes are used for IPEDS which is a mandatory data collection.

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

NCSES, in coordination with its survey contractor plans to conduct several methodological activities with the goal of improving data quality and broadening the scope of demographic information associated with the students enrolled in eligible academic institutions and the postdoctoral and nonfaculty researchers employed by same. Changes under consideration include modifying the definition of doctorate holding NFRs and expanding collection of NFR demographic data, expanding the collection of gender identity data on students, postdocs, and NFR beyond male and female, and determining the feasibility of collected race and ethnicity data in alignment with the revisions to race and ethnicity categories proposed by the Federal Interagency Technical Working Group on Race and Ethnicity Standards under OMB's 1997 Statistical Policy Directive No. 15: Standards for Maintaining, Collecting, and Presenting Federal Data on Race and Ethnicity (SPD 15). To conduct appropriate methodological testing for these changes to GSS and potentially the FFRDC postdoc survey, NCSES is requesting 2,000 hours for the three-cycle survey clearance period.

GSS has been collecting separate data on postdoctoral researchers and NFR since 1979. Separating counts of these two groups of researchers was originally implemented to improve accuracy in the counting of postdocs. As a separate category of research staff generally working

on a non-term limited basis (as opposed to postdocs who are often limited in the number of years they can be employed under this job title), the NFR was offered as a residual category to keep research staff from being erroneously reported as postdocs. Considerably less information is collected about NFRs than postdocs or graduate students: in the current GSS, there are 188 data items collected on graduate enrollments, 152 data items on postdocs, and only 15 data items regarding NFRs. NFRs also account for a modest proportion of overall counts in the GSS. For example, the 2021 survey reported information on 760,503 graduate students, 63,296 postdocs, and 30,725 NFRs. Despite this relatively lower profile within the GSS, the NFR data are used regularly to help determine research intensity of U.S. academic institutions. Since 2005, the Carnegie Classification of Institutions of Higher Education has used the NFR data to help distinguish R1, R2, and other doctoral or professional institutions. Over past three survey cycles, NCSES conducted a series of investigations to better appreciate the challenges that institutions face when dealing with the GSS definition of NFR. The investigations were necessary because institutional interpretations of the GSS definition of NFR vary, complicating comparisons of NFR data between institutions. Attachment 15 presents the results of these investigations.

NCSES plans to follow-up its NFR methodological work by conducting a record-keeping survey of GSS respondents to determine what data they currently maintain on research staff and the salient challenges posed in reporting if GSS adopts changes to NFR definitions.

NCSES plans to continue its methodological efforts to exploring expanding the collection of demographic data beyond male and female for graduate students. NCSES is currently testing revised sexual orientation and gender identity (SOGI) categories in several of its household surveys and seeks to understand data user needs and reporting capabilities of respondents to its establishment surveys as well. NCSES conducted a discussion session at the 2022 Association for Institutional Research Annual Forum to better understand how gender identity data is currently collected and maintained by GSS institutions, as well as what policies might exist that restrict institutions from reporting these data to NCSES. NCSES expects to conduct a series of virtual meetings to gain further insights into the collection of gender identity data beyond a binary question for students and postdoctoral researchers and will seek to include items related to SOGI data in the record-keeping study described above.

Finally, depending on future guidance by OMB regarding updating the race and ethnicity statistical standard in data collection, NCSES may pursue methodological work to test the impact

of revised race and ethnicity questions on the GSS and the FFRDC postdoc survey. This work may include convening technical review panels of data users, cognitive interviews with GSS/FFRDC respondents to test revised questions, and record-keeping studies with GSS/FFRDC respondents in order to determine the feasibility of any changes to race and ethnicity questions on the GSS and the FFRDC survey.

**B.5 Individuals Consulted**

The individuals consulted on GSS technical and statistical issues are listed in Exhibit 13, along with project staff at RTI International, the data collection contractor that conducts the GSS.

**Exhibit 13. Individuals Consulted on GSS Technical and Statistical Issues**

Name	Affiliation	Telephone Number
Mr. Michael Yamaner GSS Project Officer	National Science Foundation, NCSES, Alexandria, VA	703-292-7815
Dr. Amber Seligson Levanon Program Director	National Science Foundation, NCSES, Alexandria, VA	703-292-7829
Dr. Wan-Ying Chang Mathematical Statistician	National Science Foundation, NCSES, Alexandria, VA	703 292-2310
Ms. Jennifer L. Beck Survey Methodologist	National Science Foundation, NCSES, Alexandria, VA	703 292-8328
Dr. Jonathan Gordon GSS Project Director	RTI International Atlanta, GA	770-407-4952
Ms. Laura Burns-Fritch Methodology Task Leader	RTI International Research Triangle Park, NC	919-514-7014
Mr. Peter Einaudi Senior Consultant	RTI International Research Triangle Park, NC	919-541-8765
Ms. Jennifer Pauli Data Collection Task Leader	RTI International Research Triangle Park, NC	919-485-5598
Dr. Caren Arbeit Data Analysis Task Leader	RTI International Research Triangle Park, NC	510-898-4812
Mr. Chris Davies Data Quality Task Leader	RTI International Cleveland, OH	919-541-7237
Ms. Miranda Guardiola FFRDC Postdoc Survey Task Leader	RTI International Research Triangle Park, NC	919-541-5956

Mr. Mark McLean Data Delivery Task Leader	RTI International Research Triangle Park, NC	919-541-7248
Mr. Bob Steele Systems Development Task Leader	RTI International Research Triangle Park, NC	919-316-3836
Dr. Kimberly Ault Mathematical Statistical Task Leader	RTI International Research Triangle Park, NC	919-541-7455

---