2020−2022
Survey of Graduate Students and Postdoctorates in Science and Engineering

OMB Supporting Statement

Section B

Revised, March 2020

TABLE OF CONTENTS

EXHIBITS

LIST OF ATTACHMENTS

B. COLLECTION OF INFORMATION EMPLOYING STATISTICAL METHODS

The only change proposed for GSS 2020-2022 is to implement the new Classification of Instructional Programs (CIP) 2020 taxonomy, described in section B.2.1. The procedures and methodology described in other sections are in current use on GSS 2019 and will be continued without change. Attachment 9 describes changes implemented in 2017 under the last clearance and Attachment 10 describes changes proposed for 2020-22.

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 11 for the list of GSS fields).

B.1.1 Discussion of Institutional Frame

The 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 2016-2018. Note that in the 2016 GSS cycle, to be consistent with the NCSES Higher Education Research and Development (HERD) survey, for-profit institutions were declared ineligible since they are typically not research institutions but rather focus on practitioner-oriented degrees.

Exhibit 9. Number of GSS Institutions, Schools, Units and Enrollment, 2016-2018

				Graduate enrollment by degree level		
Year	**Institutions**	**Schools**	**Units**	**Total**	**Masters**	**Doctorate**
2016	714	828	15,583	684,825	na	na
2017	703	814	18,745	649,112	378,587	270,525
2018	715	817	19,592	668,307	391,211	277,096

B.1.2 Response Rates

Exhibit 10 displays unit, school, and institutional response rates for the 2016-2018 survey cycles. Response rates from GSS 2018 were very high: 96.6 percent at the institution level, 96.5 percent at the school level, and 83.8 percent at the unit level. Response rates are projected to remain stable during the next four cycles.

Exhibit 10. GSS Institution, School, and Unit Response Rates: 2016–18

	Complete Respondents			Partial Respondents			Nonrespondents		
	2016	**2017**	**2018**	**2016**	**2017**	**2018**	**2016**	**2017**	**2018**
Institution	98.00% *n=700*	94.60% *n=665*	96.60% *n=691*	0.40% *n=3*	1% *n=7*	0.70% *n=5*	1.50% *n=11*	4.40% *n=31*	2.70% *n=19*
School	98.30% *n=814*	94.50% *n=769*	96.50% *n=788*	0.40% *n=3*	1.10% *n=9*	1.00% *n=8*	1.30% *n=11*	4.40% *n=36*	2.60% *n=21*
Unit	85.90% *n=13,617*	85.10% *n=15,946*	83.80% *n=16,410*	13.60% *n=2,157*	12.50% *n=2347*	15.20% *n=2974*	0.50% *n=79*	2.40% *n=452*	1.10% *n=208*

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 he/she is 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. Once the SC is confirmed/updated, data collection commences. 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 mails the president 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 the 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 2020 data collection plan, including a timetable and communications with GSS stakeholders, is included in Attachment 12.

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 13a, 13b and 13c) 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 14). Information can be compiled on this worksheet for each unit prior to data entry into the web instrument.

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 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 2018, approximately 80% of institutions were able to provide student data using CIP codes.

NCES has recently released the 2020 CIP code list. We plan to implement these codes in the GSS 2020 because institutions will be required to report IPEDS 2020 data using the new

codes, and the GSS data collection follows the IPED data collection. For SCs uploading data with CIP codes, the recoding of GSS codes will be completed automatically as part of the data upload process, without further input from respondents. This should mitigate reporting burden for SCs using this method.

SCs that have units without associated CIP codes will need to use a Taxonomy Tool provided in the GSS web instrument to identify the impacted units and recode them to revised GSS codes before either uploading or reporting the data for those units. The Taxonomy Tool will be adapted from the one used during the 2017 survey cycle. A prototype of the Taxonomy Tool is provided in Attachment 7. To assist SCs, they will be provided with a revised GSS Code List, a GSS/CIP Crosswalk (Attachment 15) and a list of GSS codes with changes (Attachment 16).

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 provide assistance to 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
- The inclusion of a “crosswalk” listing the fields of study for which data are requested along with CIP codes for these fields. This crosswalk is for the

convenience of the institutions using CIP codes in reporting their enrollment and degree award data to the NCES IPEDS data collection (see Attachment 15)

- 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 introduce institutions to the changes planned as part of the 2017 data collection. 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.

B.3.2 Imputation Methods for Unit and Item Nonresponse

Unit and Item nonresponse. One of the key goals of the redesign was to improve the quality of data collected by the GSS. Balancing the request for additional detail by degree level and revising the primary method of data reporting while maintaining historically high response rates on a voluntary survey were overarching challenges for NCSES. NCSES understood that in light of major changes to the survey methodology, some respondents might choose not to participate in the 2017 GSS. As Exhibit 11 demonstrates, this concern was appropriate. School nonresponse increased from 1.3% in 2016 to 4.4% in 2017, and unit nonresponse increased from 0.5% to 2.4%. However, results from the 2018 GSS data collection show that the increase in school nonresponse rates may be temporary. The school-level nonresponse rate for the 2018 survey cycle was 2.6%, while unit nonresponse was 1.1%. Item-level nonresponse increased slightly in 2018 but still remains lower than the levels seen prior to the redesign.

Exhibit 11. School and unit level nonresponse rates: 2014–18

			Nonresponse rate		
Year	Schools	Units	School	Unit	Item
2014	821	14,845	0.9	0.3	3.7
2015	824	15,202	1.3	0.5	3.7
2016	828	15,853	1.3	0.5	3.3
2017	814	18,745	4.4	2.4	2.3
2018	816	19,592	2.6	1.1	2.8

SOURCE: National Center for Science and Engineering Statistics, National Science Foundation, Survey of Graduate Students and Postdoctorates in Science and Engineering.

There is evidence that the improvement in nonresponse rates may be due to the shift to electronic data interchange (EDI) as the primary form of data reporting. Exhibit 12 presents 2017 item nonresponse by data entry type. As can be seen, item nonresponse for data provided through EDI is lower than that provided through manual data entry for nearly every data type. The overall nonresponse rate for uploaded data items was 2.1%, while the nonresponse rate for manually entered data was 3.1%. Differences were most dramatic for student demographic counts (0.7% for uploaders, 4.1% for manual data entry) and for NFR counts (0.4% for uploaders, 4.7% for manual data entry). These results are promising and will continue to be monitored in future data collections.

Exhibit 12. Item nonresponse, by data entry type: 2017

Data type	Number of items	Uploaded			Manual		
		Master's	Doctoral	All	Master's	Doctoral	All
Graduate enrollment	188	2.8	1.7	3.6	4.2	2.5	5.5
Demographics	90	0.5	0.3	0.7	3.4	1.4	4.1
Primary support	98	4.9	3.0	6.3	4.9	3.5	6.9
Postdoctoral appointments	152	-	-	1.9	-	-	2.4
Demographics	30	-	-	1.9	-	-	2.0
Primary support	87	-	-	2.1	-	-	2.6
Type	35	-	-	1.5	-	-	2.3
Nonfaculty researchers	15	-	-	0.4	-	-	4.7
Overall item nonresponse rate							
All survey items	-	-	-	2.1	-	-	3.1
355 items asked 2010–16	-	-	-	2.7	-	-	4.2

NOTE: Graduate enrollment, demographics, and primary support for "All" uploaded data and "All" manually entered data were derived from addition of master's and doctoral enrollment data.

SOURCE: National Center for Science and Engineering Statistics, National Science Foundation, Survey of Graduate Students and Postdoctorates in Science and Engineering, 2017.

Imputation Methods. The 2018 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 2017 GSS Methodology Report.[1] The imputation section from that report is included in Attachment 17.

[1] RTI. (2019). *2017 Survey of Graduate Students and Postdoctorates in Science and Engineering: Survey Methodology Report* Deliverable under contract NSFDACS49100418F1261 Prepared by RTI International, April 2019.

B.3.3 Accuracy and Reliability of Data

Because this is a census with a high response rate, and statistical imputation is conducted for nonresponse, the accuracy of these data are very high. The use of CIP codes for reporting helps ensure reliability as these codes are used for IPEDS which is a mandatory data collection. And GSS has an extensive review process to check the consistency of each coordinator's data within and across years.

B.4 Testing of Procedures

No methodological research is planned for GSS 2020. NCSES will submit plans informing OMB of its intention to use methodological research burden hours before any applicable studies are undertaken.

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 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
Mr. John Finamore Program Director	National Science Foundation, NCSES, Alexandria, VA	703-292-2258
Dr. Wan-Ying Chang Mathematical Statistician	National Science Foundation, NCSES, Alexandria, VA	703 292-2310
Ms. Rebecca L. Morrison Survey Methodologist	National Science Foundation, NCSES, Alexandria, VA	703 292-7794
Dr. Patricia Green GSS Project Director	RTI International Chicago, IL	312-456-5260
Dr. Jonathan Gordon Redesign Task Leader	RTI International Atlanta, GA	770-407-4952
Mr. Peter Einaudi Data Analysis Task Leader	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 FFRDC Postdoc Survey Task Leader	RTI International Berkeley, CA	510-898-4812
Mr. Jim Rogers Data Delivery Task Leader	RTI International Research Triangle Park, NC	919-541-7291
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