

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

OMB Supporting Statement
Section A

Updated, August 2020

TABLE OF CONTENTS

Section

n

A. JUSTIFICATION	1
A.1 Necessity for Information Collection	1
A.2 Uses of Information	4
A.2.1 Federal	4
A.2.2 NSF Uses	4
A.2.3 Use by Academic Institutions	5
A.2.4 Use by the Carnegie Foundation	5
A.2.5 Use by the Professional Societies	6
A.2.6 Use in Research	6
A.2.7 Media Uses	7
A.3 Consideration of Using Improved Technology	7
A.4 Efforts to Identify Duplication	8
A.5 Efforts to Minimize Burden on Small Business	10
A.6 Consequences of Less Frequent Data Collection	10
A.7 Special Circumstances	11
A.8 Federal Register Announcement and Consultations Outside the Agency	11
A.8.1 GSS Institution Site Visits	12
A.8.2 Other Consultations	12
A.9 Payment or Gifts to Respondents	12
A.10 Assurance of Confidentiality	13
A.11 Justification for Sensitive Questions	13
A.12 Estimate of Respondent Burden	13
A.13 Cost Burden to Respondents	18
A.14 Cost to the Federal Government	18
A.15 Program Changes or Adjustments in Burden	18
A.16 Publication Plan and Project Schedule	18
A.17 Exceptions to Displaying of OMB Expiration Date	19
A.18 Exceptions to the Certification Statement	19

EXHIBITS

Number

Exhibit 1. Burden Results for 2016-18 GSS

Exhibit 2. Burden Results for 2018 GSS

Exhibit 3. Expected Composition of the 2019 GSS Frame

Exhibit 4. Burden Estimates for the 2020 GSS

Exhibit 5. Burden Estimates for the 2021 GSS

- Exhibit 6. Burden Estimates for the 2022 GSS
- Exhibit 7. Total Burden Estimates for 2020–22 GSS
- Exhibit 8. Annual GSS Survey Federal Government Estimated Costs

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, March 2020
- Attachment 6. First Federal Register Notice
- Attachment 7. GSS Taxonomy Tool Prototype
- Attachment 8. GSS Schedule
- Attachment 9. List of Changes During Last Clearance (GSS 2017-2019)
- Attachment 10. List of Changes Requested for GSS 2020
- Attachment 11. GSS Code List—Complete List of GSS Eligible Fields and Codes
- Attachment 12. GSS Data Collection Plan
- Attachment 13. GSS Data File Upload Templates
- Attachment 14. GSS Worksheet
- Attachment 15. CIP and GSS Code Crosswalk
- Attachment 16. List of Changes to GSS Taxonomy
- Attachment 17. Imputation Section from GSS Methodology Report

A. JUSTIFICATION

This submission requests a three-year reinstatement of the previously approved OMB clearance for the Survey of Graduate Students and Postdoctorates in Science and Engineering (GSS). The survey is co-sponsored by the National Center for Science and Engineering Statistics (NCSES) within the National Science Foundation (NSF) and the National Institutes of Health (NIH). The GSS is an annual survey that was last conducted in fall 2019. The OMB clearance for the GSS will expire on October 31, 2020. With this clearance package, NSF requests approval to collect data for the 2020, 2021, and 2022 survey cycles.

A.1 Necessity for Information Collection

In 2010, the America COMPETES Reauthorization Act of 2010¹ established the previously named Science Resources Statistics division as NCSES and directed NCSES to “...collect, acquire, analyze, report, and disseminate statistical data related to the science and engineering enterprise in the United States and other nations that is relevant and useful to practitioners, researchers, policymakers, and the public...” Information obtained through the GSS is critically important to NCSES’s ability to measure science and engineering resources in the United States. Furthermore, the GSS data serve as the nation’s only source of comprehensive graduate enrollment information for specific science, engineering, and health (SEH) disciplines at the departmental level. These data are solicited under the authority of the NSF Act of 1950,² as amended, and are central to the analysis presented in a pair of congressionally mandated reports^{3,4} published by NSF, the *Science and Engineering Indicators* and the *Women, Minorities, and Persons with Disabilities in Science and Engineering*.

The GSS is the only annual national survey that collects information on the characteristics of master’s and doctoral student enrollment for specific SEH disciplines at the departmental level. It also collects information on master’s and doctoral student enrollment by degree level, race and ethnicity, citizenship, sex, source of support, and mechanism of support; information on Postdoctorates (postdocs) by citizenship, sex, source of support, mechanism of support, and origin of doctoral degree; and information on other doctorate-holding nonfaculty

¹ Section 505, Pub. L. No. 111-358. See Attachment 1

² See Attachment 2.

³ 42 U.S. Code § 1863(j)(1)

⁴ 42 U.S. Code § 1885(a), 1885(d)

researchers (NFRs) by sex and type of doctoral degree (see Attachment 3 for screenshots of the GSS web instrument). The GSS has been conducted by NCSES annually since 1972. Additional financial support for the GSS is provided by the NIH.

The GSS is a census of all organizational “units” (departments, programs, research centers, and health care facilities) in SEH fields within eligible academic institutions in the United States that grant research-based master’s or doctorate degrees. The survey collects information on graduate students enrolled in these units, as well as postdocs and NFRs working within these institutions. As a part of the GSS, NCSES also periodically surveys Federally Funded Research and Development Centers (FFRDCs) to collect information on postdocs such as race/ethnicity, sex, citizenship, source of support and area of research (see Attachment 4 for screenshots of the FFRDC Postdoc Survey web instrument).

The coronavirus pandemic is having a substantial impact on colleges, universities, and the nation’s workforce. In response, NCSES is adding short modules to some surveys to assess the pandemic’s disruption and impact. Each NCSES survey will assess the impact of the pandemic on a different subpopulation with the goal that the collective information obtained will provide insight on the impact of the pandemic on the nation’s science and engineering enterprise.

Many postsecondary institutions closed this spring and are still undecided about whether to open campuses in the fall. Some research laboratories also closed and although they may have reopened, capacity is limited due to social distancing concerns. The pandemic has also affected schools’ finances resulting in reductions in force and furloughs. These changes had only a small effect on the 2019 GSS response rates because the majority of data was submitted before the February 28 deadline, but it is possible that it could lead to major problems in the coming cycle, the 2020 GSS, and subsequent cycles. In addition, the pandemic could lead to noticeable changes in the GSS estimates of graduate student enrollment, postdoctoral appointments, and non-faculty researcher employment.

NCSES proposes adding a small number of questionnaire items to the upcoming 2020 GSS to assess the pandemic’s impact on the institutions that participate in the survey. These questionnaire items are designed to inform the following research questions.

- **Will GSS coordinators be able to participate in the 2020 GSS?**

Possible reductions in staff likely will affect the ability of our respondents to provide data. NCSES would like to learn early in the process whether this might be an issue to enable the incorporation of data collection modifications to avoid a major reduction in survey participation.

- **What changes are occurring within these institutions that may affect enrollments of master's and doctoral students?**

To address this question, proposals include items that ask about changes in school policies that may impact the counts reported for graduate students. This information will increase our understanding in two ways: 1) it will provide context for the GSS data quality checks performed after schools submit their data (a process that compares current data with past data to identify possible errors in reporting) and 2) it will provide insight on the broader data patterns related to the impact of the pandemic.

- **What changes are occurring within these institutions that may affect postdocs and doctorate-level non-faculty researchers (NFRs)?**

The GSS is the only federal survey that collects information on postdocs and NFRs, so it is important to gather information about how they may be affected by the pandemic. For example, are institutions rescinding or delaying offers to postdocs? Have the terms of postdoc appointments been changed? Answers to these questions are designed to help inform our understanding of any changes in postdoc and NFR data in the 2020 GSS.

- **Have research facilities or staffing at the institution been impacted?**

Research laboratories closed this spring at many universities and have reopened with restricted access. And many institutions have announced furloughs or other reductions in staff. Questionnaire items designed to inform this research question will provide insight on how facility changes have affected postdocs and NFRs.

NCSES, in coordination with the GSS survey contractor, plans to develop questionnaire items in the months leading up to the 2020 GSS survey launch to inform these coronavirus pandemic impact research questions. When participating schools are notified that the 2020 GSS is available, the notification will describe the additional items about the impact of the coronavirus pandemic that have been added to the beginning of the survey. The plan is to limit these questions so that burden does not exceed 7 minutes per GSS coordinator. If the pandemic continues to disrupt universities, it is possible that these same items, or a revised set of items that better inform the relevant research questions, could be included in the 2021 GSS and 2022 GSS survey cycles. Cognitive testing on potential items will be conducted in August and September of the survey year prior to the data collection start. Burden hours for the proposed cognitive testing (estimated at approximately 30 hours per survey cycle) are included in the 1,000 hours for

methodological testing included in Section A.12, Table 7. Based on the results of the cognitive testing, a request for a non-substantive change will be submitted to OMB in late September describing the questionnaire items proposed for inclusion on the survey.

A.2 Uses of Information

A.2.1 Federal

Data derived from the GSS are routinely provided to Congress and to various agencies of the Executive Branch. Recent examples of provided data include:

- Data on graduate SEH enrollment provided annually to NCES for comparison purposes and are published in the *Digest of Education Statistics*.
- Data in specially prepared GSS tabulations used by the NIH to answer specific questions to help their agencies prepare budgets and conduct program evaluation studies.

NCSES and NIH extensively use the information on the number and characteristics of students currently enrolled in graduate SEH programs and of persons engaged in postdoctoral programs to assess the future stock of trained SEH personnel. A variety of more general information needs are met through the annual release of data in electronic format. NCSES publishes a short InfoBrief and a set of statistical tables, *Survey of Graduate Students and Postdoctorates in Science and Engineering Data Tables*, available on the NCSES website.

Data from the GSS are also available as public use files, and on the Web through the new Interactive Data Tool (<https://ncesdata.nsf.gov/ids/gss>). The Interactive Data Tool contains institutional and summary data from all of NCSES's academic sector surveys for all institutions offering graduate-level instruction and/or maintaining research and development (R&D) activity in SEH fields.

Each year, major findings from the GSS are published in an InfoBrief. The most recent InfoBrief, *Graduate Enrollment in Science, Engineering and Health Rose by 3% in 2018*, is available on the NCSES website (<https://www.nsf.gov/statistics/2020/nsf20312/>), and is also included in this document as Attachment 5.

A.2.2 NSF Uses

Special tabulations from the GSS data constitute a key resource in meeting policy and program information needs of the Foundation. Major examples of GSS data uses are in the two

congressionally mandated biennial reports produced by NCSES, *Science and Engineering Indicators* and *Women, Minorities, and Persons with Disabilities in Science and Engineering*.

The GSS is one of four NCSES surveys whose microdata are combined into an integrated database to produce the *Academic Institution Profiles* published on the NCSES website (<https://ncesdata.nsf.gov/profiles/>). The other three surveys are (1) the Survey of Earned Doctorates (SED); (2) the Higher Education Research and Development (HERD) Survey; and (3) the Survey of Federal Science and Engineering Support to Universities, Colleges, and Nonprofit Institutions. As explained in the next section, these data are further integrated with institutional data from other NCSES surveys and with surveys conducted by the National Center for Education Statistics (NCES). Together these data provide policy makers with information on the role of higher education in the context of the national R&D effort.

Primary uses of the GSS data also include reviewing changing enrollment levels to assess the effects of NSF initiatives; tracking student support patterns; and analyzing participation in SEH fields by targeted groups for all disciplines or for selected disciplines and for selected groups of institutions. Program officers check departmental and institutional records, including data from the GSS and NCES's Integrated Postsecondary Education Data System (IPEDS), to determine department eligibility for NSF programs targeted to special populations or instructional programs.

A.2.3 Use by Academic Institutions

The surveyed institutions themselves are major users of the GSS data. Institutions use the NCSES's GSS data reports or the Interactive Data Tool to study selected groups of peer institutions for planning and comparative purposes. They combine the NCSES data with information from state and local governments on institutions in their geographic areas. Institutions also use the comparative data to review the strength of their own programs on the basis of factors such as support of students by various federal agencies and progress in reaching special target populations.

A.2.4 Use by the Carnegie Foundation

Data from the GSS are used by the Indiana University Bloomington Center for Postsecondary Research in developing the Carnegie Classification of Institutions of Higher

Education. The center uses the GSS data on postdocs and nonfaculty research staff with doctorates as components of the “research activity” measure constructed for doctorate-granting universities (for more detail see: <http://carnegieclassifications.iu.edu/methodology/basic.php>).

A.2.5 Use by the Professional Societies

Data users include the American Association of Colleges of Nursing, American Association of Universities, American Chemical Society, American Council of Education, American Geological Society, American Institute of Physics, American Physical Society, American Society for Engineering Education, Association of American Medical Colleges, Association of International Educators, Commission on Professionals in Science and Technology, Computing Research Association, Council of Graduate Schools, Federation of American Societies for Experimental Biology, and the National Postdoctoral Association. Generally, associations use GSS data to monitor trends in enrollment by field of study, and many are also interested in tracking the numbers of postdoc and NFRs.

A.2.6 Use in Research

Researchers studying a diverse range of policy issues relating to the SEH labor pipeline have used GSS data. From 2017-2019, at least 8 papers using GSS data were published. Topics included visa policies for international students wanting to work in the United States, work-life balance for postdoctoral researchers, discipline specific pipeline issues, and cross-country comparisons of women in research.⁵

⁵Anderson, Stuart (2017). International Students and STEM OPT. National Foundation for American Policy. Arlington, VA.

<http://nfap.com/wp-content/uploads/2017/10/International-Students-and-STEM-OPT.NFAP-Policy-Brief.October-2017.pdf>

Lee, J., J.C. Williams, S. Li (2017). Parents in the Pipeline: Retaining Postdoctoral Researchers with Families. San Francisco, CA, The Center for Work Life Law.

Miller, C. W., B.M. Zeickl, J.R. Posselt, R.T. Silvestrini, T. Hodapp (2019). "Typical physics Ph.D. admissions criteria limit access to underrepresented groups but fail to predict doctoral completion." Science Advances 5: eaat7550.

<http://advances.sciencemag.org/content/advances/5/1/eaat7550.full.pdf>

Niu, L. (2017). "FAMILY SOCIOECONOMIC STATUS AND CHOICE OF STEM MAJOR IN COLLEGE: AN ANALYSIS OF A NATIONAL SAMPLE." College Student Journal 51(2): 218-312.

Passalacqua, N. and H. Garvin (2018). "Experiences in Applying to and Attending Biological Anthropology Programs Focused on Human Skeletal Biology." Forensic Anthropology 1(4): 201-214.

https://www.researchgate.net/profile/Heather_Garvin/publication/327275957_Experiences_in_Applying_to_and_Attending_Biological_Anthropology_Programs_Focused_on_Huma

A.2.7 Media Uses

Enrollment of graduate students in S&E fields are well reported by the press, including *Forbes*, *The Chronicle of Higher Education*, and *Science*. Recent examples of the use of GSS data are *Forbes* and *Inside Higher Ed* articles on foreign students and graduate enrollment.⁶

A.3 Consideration of Using Improved Technology

NCSES has engaged in a process of continuous improvement for the GSS, involving technical innovations to increase the utility of the data collected and reduce respondent burden. During the 2016 GSS data collection, a stratified random sample of 80 institution coordinators was selected for a Pilot survey to test the feasibility of implementing a series of redesign changes. Pilot coordinators were instructed to report master's and doctoral student data separately and upload their data file on the GSS web system using Classification of Instructional Program (CIP) codes instead of GSS field of study codes. Two data upload options were offered to the Pilot coordinators and the uploaded data were automatically loaded into the GSS web instrument. Pilot coordinators were allowed to review and edit the data before submitting to NCSES. Of 80 Pilot coordinators, 76 uploaded their data as part of the 2016 GSS pilot survey (see A.8.3 for more information).

[n Skeletal Biology/links/5baa3d5c45851574f7e447ab/Experiences-in-Applying-to-and-Attending-Biological-Anthropology-Programs-Focused-on-Human-Skeletal-Biology.pdf](https://doi.org/10.1002/ajpa.23457)

Passalacqua, N. V. (2018). "Are careers in biological anthropology sustainable?" *American Journal of Physical Anthropology*. DOI: 10.1002/ajpa.23457

<https://onlinelibrary.wiley.com/doi/pdf/10.1002/ajpa.23457>

Saxon, T. and S. Weiler (2019). "Defense spending and women in research: A cross-country comparison." *Science and Public Policy*. doi: 10.1093/scipol/scz021

<https://academic.oup.com/spp/advance-article/doi/10.1093/scipol/scz021/5485747>

Tsugawa, M. (2019). Testing an Identity-Based Motivation Conceptual Framework for Engineering Graduate Students. Materials Science & Engineering Department. Reno, NV, University of Nevada, Reno. PhD dissertation.

https://scholar.google.com/scholar_url?url=https://scholarworks.unr.edu/bitstream/handle/11714/6692/Tsugawa_unr_0139D_13032.pdf%3Fsequence%3D1%26isAllowed%3Dy%23page%3D75&hl=en&sa=X&d=1669087533915004106&scisig=AAGBfm2BDaacVNLrrFhZPTV4ND8SqKwMlw&nossl=1&oi=scholaralt

⁶ <https://www.insidehighered.com/quicktakes/2017/10/11/foreign-students-and-graduate-stem-enrollment>

<https://www.forbes.com/sites/stuartanderson/2017/10/15/are-international-students-next-on-the-menu/#3f44c4a35e9f>

NCSES leveraged these technical innovations in the 2017 GSS data collection methods, which were based on the results of the research and the 2016 GSS pilot survey. In 2017 and in subsequent GSS data collections, all schools are asked to use one of the two following data uploading options:

- Upload a file containing de-identified individual records that the Web system automatically aggregates to the unit-level format and then populates the appropriate cells in the GSS survey.
- Upload a file that contains an Excel macro program that aggregates individual-level data into unit-level data locally. This option is available for respondents who do not wish to transmit individual-level data over the Internet.

Based on analyses of respondent behavior in the 2017 and 2018 data collections, the expansion of available upload options has led to an increase in the number of respondents that supply GSS data through file uploads and has led to a reduction in the overall burden of completing the survey through the web instrument.

A.4 Efforts to Identify Duplication

NCSES staff consults regularly with other federal agencies and private organizations to prevent duplication of data collection activities and to stay abreast of changes in other surveys. Such consultations take place with the NCES, the Council of Graduate Schools (CGS), and others. Specific surveys conducted by these groups will be discussed below. In addition, NCSES staff participate in a variety of NCES-related activities, including serving on the 2020 CIP Working Group and Technical Review Panels. The routine data uses of the federal agencies described in Section A.2.1 have largely determined the content of the GSS questionnaire.

Only the GSS collects the following information at the level of detailed SEH fields of study:

- For full-time master's and doctorate students, aggregate counts by
 - sources of major financial support (federal agencies, institutions, self-support, etc.)
 - mechanisms of major financial support (fellowships, teaching assistantships, etc.)
 - gender
 - citizenship
 - enrollment status (full-time or part-time; first time)
 - race/ethnicity background of U.S. citizens
- For part-time master's and doctorate students, aggregate counts by

- gender
- citizenship
- race/ethnicity background of U.S. citizens
- For postdocs, aggregate counts by
 - sources of major financial support
 - mechanism of major financial support
 - gender
 - citizenship
 - type of doctoral degree
 - doctoral degree origin
- For NFRs, aggregate counts by
 - gender
 - type of doctoral degree

Because the data are collected from all eligible institutions with graduate SEH departments, data are available at the detailed field of study by institutional characteristics, such as highest degree granted, geographical location, type of control (public or private), or any other special grouping (medical schools, historically black colleges and universities, land-grant institutions, etc.) as well as by rankings on various characteristics (foreign enrollment, minority enrollment, field-specific enrollment, etc.)

Some graduate enrollment data are collected by other organizations, either federal or private, but none of the other data collection efforts contain the detailed field distribution that is required for analyses and provides the necessary data for NCSSES and NIH. IPEDS, for example, collects race and ethnicity data every 2 years for only nine select fields (of which four are within the NCSSES definition of science and engineering, but are at more general level than is collected for the GSS). The IPEDS annual fall enrollment data collected by race and ethnicity category are not reported by the field, and hence, they do not provide a viable substitute for the race and ethnicity data collected in the GSS. No data are collected on source of support or postdocs and NFRs.

The CGS conducts an annual survey of graduate enrollment in cooperation with the Graduate Records Examinations (GRE) Board, surveying 775 institutions in 2018 that were members of the CGS or one of the four regional graduate school associations—the Conference of Southern Graduate Schools, the Midwestern Association of Graduate Schools, the Northeastern Association of Graduate Schools, and the Western Association of Graduate

Schools. The survey had a response rate of 76%, with 589 schools responding. The survey collects data by 51 fine fields of study using the GRE discipline codes as its taxonomy, type of institutional control, and highest level of degree offered, but has no data on source of financial support. It also collects information on postbaccalaureate and post-master's certificates and applications to graduate schools. Only the GSS maintains detailed data grouped into ninety-one fine fields of study on all SEH degree fields at all eligible institutions and institution-provided data on source of financial support.

A number of surveys are conducted by other professional societies or by groups of institutions and are limited to a single field or group of related fields or to institutions that are members of the organization. These surveys may collect far more detailed data on the fields of interest to the organization conducting the survey, and may even collect data on topics not covered by the GSS (e.g., on undergraduate enrollment), but they do not provide compatible data on all SEH fields, nor do they often address the issue of sources and mechanisms of financial support for graduate students.

A.5 Efforts to Minimize Burden on Small Business

The GSS does not collect information from small businesses.

A.6 Consequences of Less Frequent Data Collection

A less frequent survey cycle would have several serious consequences. First, there would be the loss of information. Because of the data uses described previously, biennial or less frequent collection means that data users would be unable to access current information. Collecting the GSS annually also increases the value of the data for monitoring trends, particularly the effects of dramatic changes in the larger context. Minor shifts in enrollment trends are monitored as early indicators of likely future changes in the supply of SEH professionals.

Other examples of trend monitoring are changes in the foreign graduate student enrollment and postdoc employment counts that correspond to the events such as September 11, 2001, the 2007-2009 Great Recession, and immigration policy changes. Less than annual data collection may not capture such changes or reveal the inflection point of a changing trend. Following the September 11, 2001, the release of the GSS fall enrollment data was eagerly

anticipated to examine trends in SEH graduate enrollment by foreign visa holders. Foreign student enrollment did not drop immediately (i.e., in 2001), and the trends varied by several years for first-time enrollment and total enrollment. Those nuances would have been lost if the data had not been collected every year.

Annual collection also helps reduce respondent burden. Most colleges and universities have automated record keeping systems, facilitating their ability to respond to the GSS on an annual cycle. These automated record systems considerably reduce the time required to assemble and report information needed for the GSS related to graduate enrollment by field, demographics, postdoctoral appointments, and sources and mechanisms of support, etc. Thus, because the database and software are retained, kept current, and easily accessed, collecting consistent data annually considerably reduces respondent burden for academic institutions with automated data systems.

Annual collection also helps to maintain contacts with the coordinators within institutions. Having this continuity helps the coordinators maintain their databases and, therefore, maintain the quality of the data.

A.7 Special Circumstances

This data collection does not require any of the reporting requirements listed.

A.8 Federal Register Announcement and Consultations Outside the Agency

The Federal Register notice was published on April 6, 2020 (85 FR 19169) (see Attachment 6). NCSES received two comments. NCSES received the first comment on 4 April 2020 from an economics professional association requesting a copy of the draft information collection request (ICR) including the survey instrument and supporting statement. NCSES informed the commenters that the ICR was currently undergoing internal review within NCSES with plans to submit it for public review in June and that the GSS would be largely unchanged from its current design.

NCSES received a second comment on 6 May 2020 from a group representing several organizations. The commenters requested that NCSES include measures of sexual orientation and gender identity on the GSS. NCSES informed the commenters that it shares their interest in improving federal data collections and providing reliable measures for important segments of the

population. Furthermore, NCSES described its process for evaluating possible questionnaire additions, including the extensive experimentation involved and the time and resources required. Finally, NCSES informed the commenters that it is conducting research to evaluate these measures and does not intend to include them in the 2020-22 GSS.

As described in the next sections, in the past three years, several consultations with the respondents have taken place to examine different aspects of the GSS data collection and to inform the changes introduced in 2017.

A.8.1 GSS Institution Site Visits

NCSES routinely conducts site visits to better understand the reporting experiences of institutions that participate in the GSS. From 2017 to 2019, NCSES conducted 5 site visits to institutions of various size and reporting capacity. The site visits focused on data availability, barriers to using Electronic Data Interchange (EDI), and assisting coordinators in making the transition from manual entry to EDI. Most coordinators expressed an interest in uploading but did not fully understand the benefits, work required, or potential to reduce burden in the subsequent years. The information gathered in these visits also informs data collection procedures to better assist other coordinators complete the survey.

A.8.2 Other Consultations

NCSES regularly consults with the Department of Education's NCES, and other federal agencies, such as NIH, professional societies, and institutions. NCSES staff members maintain frequent contact with members of the data-using community as well as with major academic data providers through attendance at professional society meetings and consultation with institutional and agency officials. GSS sessions are typically held at the Association for Institutional Researchers (AIR) Annual Forum and the CGS Annual Meeting each year to obtain respondent input.

A.9 Payment or Gifts to Respondents

There are no payments or gifts to GSS respondents.

A.10 Assurance of Confidentiality

No pledge of confidentiality is given to institutions providing data to the GSS because all data collected in the GSS are aggregate counts of students, postdocs, and NFRs. Data are published only at the departmental summary level.

A.11 Justification for Sensitive Questions

The survey does not contain any questions of a sensitive nature.

A.12 Estimate of Respondent Burden

Each survey cycle, when respondents reach the end of the GSS web instrument, they are asked to report how long it took them to complete the data collection. In the past three cycles (2016-2018 data collections), the average burden per coordinator reported each cycle was 17.8 hours. However, burden varies considerably across respondents. Factors impacting burden include the number of organizational units at the institution, the degree to which requested data can be queried from centralized institutional databases, whether the GSS survey coordinator relies on the Unit Respondents (URs) in various units, for some of the requested data, and whether the respondent uploads their data or manually enters it into the GSS web instrument.

Prior to the 2017 data collection nearly all data was manually keyed into the web instrument. A 2017 survey redesign requested master's and doctorate student data to be reported separately and introduced Electronic Data Interchange (EDI) as the primary response method. Switching to EDI was necessary as reporting master's and doctorate students separately doubles the amount of data requested and would otherwise increase the required burden. A 2016 pilot conducted among a stratified random sample of respondents confirmed that EDI would mitigate any burden increases. There was, however, a one-year increase in burden as coordinators mastered the new data collection methods, as shown in Exhibit 1. Data from the second year of the redesigned survey (GSS 2018) corroborate that EDI has led to a decline in respondent burden when compared to both the transition year and the legacy survey, even in the face of collecting more detailed data (i.e., separate reporting of master's and doctorate student data).

Exhibit 1. GSS 2016-2018 Burden by Institutional Reporting Size

	2016 Average Burden (hours)	n	2017 Average Burden (hours)	n	2018 Average Burden (hours)	n
Large Reporter	30.7	251	37.8	232	29.0	211
Small Reporter	7.7	411	10.8	363	8.1	355
<i>Total Burden</i>	<i>16.4</i>	<i>662</i>	<i>21.3</i>	<i>595</i>	<i>15.8</i>	<i>566</i>

While coordinators are asked about the burden required to complete the survey, only 63% provided a response in 2018. Several techniques were used to estimate the burden for remaining coordinators. First, outliers defined as 3 standard deviations above the mean were removed. Next, missing data was replaced with the mean of the analysis group. Finally, unit respondent burden was calculated and included with the coordinator’s burden. This calculation is necessary because when a school utilizes URs, the coordinators burden is minimal while the response burden falls to individual unit respondents. In 2018, the average burden for unit respondents was 2.3 hours. This figure was applied to all units at schools utilizing URs and was then added to the coordinator’s burden. Estimates of the GSS 2018 burden are presented in Exhibit 2. These figures best represent the total effort required to complete the survey and are the basis for the 2020, 2021, and 2022 burden estimates.

Exhibit 2. GSS 2018 Burden by Institutional Reporting Size and Data Provision Method

Institution Type	Coordinators without Unit Respondents	Avg. Burden (hours)	Coordinators with Unit Respondents	Avg. Burden (hours)	Total Coord inator s	Total Burden (hours)
More than 15 units, EDI	299	29.4	19	167.0	318	11,989
More than 15 units, Manual data entry	32	20.5	10	107.6	42	1,730
15 or fewer units, EDI	359	8.2	4	17.4	363	3,013
15 or fewer units, Manual data entry	157	7.9	21	16.5	178	1,602
Estimated total	847	16.1	54	86.4	901	18,334

In 2018, the average response burden for coordinators using unit respondents was 86.4 hours compared to only 16.1 for coordinators not using unit respondents. This difference highlights why coordinators should prioritize reporting all data using EDI. Even schools that

uploaded partial data and used unit respondents to report missing data saw a significant increase when compared to schools that uploaded all data and did not use unit respondents. As described in section 8.a, NCSES periodically conducts site visits to these types of institutions to assist in the transition to EDI reporting.

To estimate burden for the next three data collection cycles, the GSS frame is split by response method (EDI or manual entry) and the number of organizational units reported by the institution (more than 15 units are large reporters and 15 or fewer units are small reporters). This stratification better aligns with the current state of the GSS and will provide a more accurate estimation of the burden required to complete the survey. Based on the 2018 GSS, 35.3 percent of schools were large uploaders, 4.7 percent were large manual entry, 40.3 percent were small uploaders, and 19.8 were small manual entry.

The expected frame for the 2019 GSS includes 720 institutions comprising 822 schools with 906 total respondents. Assuming a steady state in terms of the use of EDI, the same number of coordinators per category are retained in each year. Five small manually entering schools were added in 2020, 2021, and 2022 due to anticipated organic growth in the census frame, resulting in 318 large uploaders, 42 large manual entry, 363 small uploaders, and 183 small manual entry reporters expected in the 2019 GSS (see Exhibit 3). Given the historically high levels of participation, a 100 percent school response rate is used in these estimates.

Exhibit 3. Expected Composition of the 2019 GSS Frame

Institution Type	# of Schools	Percent
More than 15 units, Uploading	318	35.1%
More than 15 units, Manually Enter	42	4.6%
15 or fewer units, Uploading	363	40.1%
15 or fewer units, Manually Enter	183	20.2%
Totals	906	100.0%

As mentioned, burden estimates for the 2020-2022 GSS project burden are reported by coordinators by response method and institution size. As some schools utilize more than one coordinator and use different response methods, the number of coordinators is greater than the number of schools and institutions. These figures also include burden from other individuals on campus that assist with the data collection, such as unit respondents. Estimates for the 2020 GSS are provided in Exhibit 4. In 2020, the National Center for Education Statistics (NCES) is expected to update their academic taxonomy to CIP 2020. The transition burden from CIP 2010

to CIP 2020 should be minimal for coordinators that rely on EDI, as the transition to the new coding schema will be handled programmatically. Coordinators that manually enter data may need to use the Taxonomy Tool (Attachment 7) to recode select GSS codes impacted by this transition. While this change imposes a considerable burden on the GSS survey contractor in terms of programming and data editing, it is expected to impose a minimal burden on GSS respondents.

Exhibit 4. Burden Estimates for the 2020 GSS

Institution Type	Respondents (# of schools)	Average Burden (hours)	Total Burden (hours)
More than 15 units, Uploading	318	37.7	11,989
More than 15 units, Manually Enter	42	41.2	1,730
15 or fewer units, Uploading	363	8.3	3,013
15 or fewer units, Manually Enter	188	9.0	1,692
Estimated total	911		18,424

Exhibit 5 presents burden estimates for the 2021 GSS data collections, which will include a biennial Survey of Postdocs at FFRDCs. Response burden for the FFRDCs is estimated based on the 2017 data collection in which FFRDCs required an average of 1.7 hours per center to complete the information request. The 2022 GSS estimates are provided in Exhibit 6.

Exhibit 5. Burden Estimates for the 2021 GSS

Institution Type	Respondents (# of schools)	Average Burden (hours)	Total Burden (hours)
More than 15 units, Uploading	318	37.7	11,989
More than 15 units, Manually Enter	42	41.2	1,730
15 or fewer units, Uploading	363	8.3	3,013
15 or fewer units, Manually Enter	193	9.0	1,737
FFRDCs	43	1.7	73
Estimated total	959		18,542

Exhibit 6. Burden Estimates for the 2022 GSS

Institution Type	Respondents (# of schools)	Average Burden (hours)	Total Burden (hours)
More than 15 units, Uploading	318	37.7	11,989
More than 15 units, Manually Enter	42	41.2	1,730
15 or fewer units, Uploading	363	8.3	3,013
15 or fewer units, Manually Enter	198	9.0	1,782
Estimated total	921		18,514

The annual burden estimates are presented in Exhibit 7, along with the cost burden estimate for respondents. At an estimated cost of \$36.83 per hour (based on the Bureau of Labor Statistics May 2018 average hourly wages for “Management Analysts,” within NAICS 611300 - Colleges, Universities, and Professional Schools, accessed on March 3, 2020, at <http://data.bls.gov/oes/>), the average annual cost to respondent institutions is \$682,901 (\$712 per respondent). In addition, the burden estimate includes 1,000 hours for conducting GSS site visits, methodological testing, and other survey improvements. This 1,000 hour estimate includes the burden hours for all three GSS survey cycles (2020, 2021, and 2022) associated with any cognitive testing needed for the development of new GSS questionnaire items to assess the impact of the coronavirus pandemic.

Exhibit 7. Total Burden Estimates for 2020-2022 GSS

Survey Cycle	Respondents (# of coordinators)	Total Burden (hours)	Total Annual Cost to Coordinators	Annual Cost per Coordinator
2020 GSS	911	18,424	\$678,555	\$745
2021 GSS	959	18,542	\$682,901	\$712
GSS Coordinators	916	18,469	\$680,213	\$743
FFRDC Coordinators	43	73	\$2,688	\$63
2022 GSS	921	18,514	\$681,870	\$740
Future methodological testing (all 3 years)		1,000	\$36,830	NA
Total estimated burden	2,791	56,480	\$678,555	\$745
Estimated average annual burden	930	18,827	\$682,901	\$712

A.13 Cost Burden to Respondents

This survey does not require the purchase of equipment, software, or services beyond those normally used in universities as part of customary and usual business. See exhibit 7 for annual respondent personnel costs.

A.14 Cost to the Federal Government

The average cost per cycle of conducting the GSS is \$1.75M based on the total estimated value of the current contract (\$7M) to conduct four cycles, 2018-21. The estimated total cost of the GSS to the federal government is approximately \$2.1M per cycle. Exhibit 8 presents more detailed information on this estimate.

Exhibit 8. Annual GSS Survey Federal Government Estimated Costs

GSS Resources and Activities	Total (\$)
Data collection and processing contract	1,750,000
GSS survey manager (1.0-person year)	150,000
Other NCSES staff (program manager, statistician, editor, etc.)	210,000
Publication Web posting, printing and mailing costs	1,000
Estimated total	2,111,000

For the 2019 GSS, NIH contributed \$445,546 (20%) of the annual contract costs. It is assumed that NIH will continue that level of support. NCSES funds the remainder of the annual costs to the federal government.

A.15 Program Changes or Adjustments in Burden

Burden estimates have been lowered substantially from previous clearances. The new procedures for data uploads have streamlined the process, especially for large universities.

A.16 Publication Plan and Project Schedule

The GSS project schedule (Attachment 8) for the entire project from design to final publication is similar each year. Institutions are contacted to confirm the survey coordinators in September, and the survey is launched in October, with a final closeout date in April of the following year. The most recent InfoBrief was published in March 2020 along with the detailed data tables, and a description of the survey methodology (see Attachment 5). There are no

complex analytical techniques incorporated into the GSS, except for the use of imputation for nonresponse (see Section B.2.3).

A.17 Exceptions to Displaying of OMB Expiration Date

Not applicable. The OMB control number and expiration date will be displayed on the GSS web survey login page and on GSS worksheets provided to respondents for reference purposes (worksheets are no longer used for actual data submission).

A.18 Exceptions to the Certification Statement

No exceptions to the certification statement are being sought.