Survey of Graduate Students and Postdoctorates in Science and Engineering: Fall 2015

Technical Notes

Survey Overview (2015 survey cycle)

Purpose. The Survey of Graduate Students and Postdoctorates in Science and Engineering (GSS) is an annual census of all academic institutions in the United States granting research-based master’s degrees or doctorates in science, engineering, or selected health (SEH) fields as of the fall of the survey year. Sponsored by the National Center for Science and Engineering Statistics (NCSES) within the National Science Foundation and by the National Institutes of Health (NIH), the GSS collects counts of graduate students, postdoctoral researchers (postdocs), and other doctorate-holding nonfaculty researchers (NFRs) at these institutions by demographics and other characteristics, such as source of financial support. Results are used to assess shifts in graduate enrollment, postdoc and NFR appointments, and trends in financial support.

Data collection authority. The information is solicited under the authority of the National Science Foundation Act of 1950, as amended, and the America COMPETES Reauthorization Act of 2010. The Office of Management and Budget (OMB) control number is 3145-0062 and expires on 30 November 2017.

Survey contractor. RTI International.

Survey sponsors. NCSES and NIH.

Key Survey Information

Frequency. Annual.

Initial survey year. 1966.

Reference period. Fall 2015.

Response unit. Organizational units (e.g., academic departments, degree-granting programs, university-affiliated research centers, and health care facilities) in academic institutions.

Sample or census. Census.

Population size. 15,202 units at 711 academic institutions.

Sample size. Not applicable.

Survey Design

Target population. The survey target population is all academic institutions in the United States and its territories (Guam and Puerto Rico) that grant master’s or doctorate degrees in SEH fields. This population includes branch campuses, affiliated research centers and health facilities, and separately organized components, such as medical or dental schools, schools of nursing, and schools of public health.

In 2015, the survey universe included 711 institutions with 824 schools and 15,202 units. There were 525 schools and 13,506 units within 412 doctorate-granting institutions and 299 schools and 1,696 units within 299 master’s-granting institutions. Data were collected at the organizational unit level. Detailed
information on the changes to the survey universe and final number of institutions, schools, and units is provided in tables A-2 through A-5.

**Sample frame.** The survey is a census in which eligible academic institutions are identified primarily through the Integrated Postsecondary Education Data Systems (IPEDS).

**Sample design.** Not applicable.

**Data Collection and Processing Methods**

*Data collection.* The survey data are collected through coordinators at eligible institutions. Coordinators are assigned by their institution and are responsible for identifying all GSS eligible units, collecting the requested data, and submitting the data to the survey contractor. While some coordinators complete the survey on their own, other coordinators enlist the aid of others (unit respondents) to gather the data. Institutions may assign multiple coordinators. For example, an institution may have one coordinator for each school within the institution or may have separate coordinators for graduate student data and for postdoc and NFR data. When a new coordinator is needed, the president’s office at the institution is asked to designate as coordinator the person most knowledgeable about the graduate student and/or postdoc data.

Once coordinators are confirmed, they are provided access to the GSS Web survey. A hard copy of the survey worksheets and GSS-eligible code lists also are mailed to each coordinator as reference. Data are collected at the organizational unit level (e.g., departments, degree-granting programs, research centers, health facilities) and include demographic and funding information for graduate students and postdocs. Data collection for the GSS is done in two parts. Part 1 requires the review and identification of organizational units (“units”) within the school. Part 2 collects aggregated counts and selected characteristics of graduate students, postdocs, and NFRs for the units. The 2015 survey cycle launched in October 2015 and concluded in May 2016. The deadline for Part 1 was 30 November 2015. The deadline for Part 2 was 29 February 2016.

*Mode.* The Web survey was the primary mode of data submission. A paper worksheet was provided for informational purposes and to assist in preparing figures to be entered in Part 2 of the Web survey. The content and format of the paper worksheet were identical to Part 2 of the Web survey. A small number of coordinators chose not to use the Web survey but submitted their Part 2 data in an Excel spreadsheet. The survey contractor loaded this Excel data into the Web survey. Of the 701 institutions that submitted data in 2015, the modes of response were as follows:

- **Web survey:** 624 institutions reported all of their GSS data using the Web survey.
- **Upload tool on the GSS website:** 69 institutions uploaded all or some portion of the data via the data upload feature available in the GSS website or provided files for upload by GSS staff.
- **Excel spreadsheet:** 8 institutions reported all or some portion of their data via Excel spreadsheets for input by survey contractor staff into the GSS website.

*Response rates.* Response rates are calculated based on responses to the survey’s various data-collection grids (graduate student and postdoc counts, by ethnicity and race; full-time graduate student and postdoc counts, by primary source or mechanism of support; counts of postdocs, by type of doctoral degree and primary mechanism of support; counts of postdocs, by type of doctoral degree and citizenship; counts of postdocs, by origin of doctoral degree; and counts of NFRs, by type of doctoral degree and sex).

The method for calculating response rates for units has changed over time. From 2007 to the present, complete row and column totals for all grids and all details summing to the totals were complete responses; some data reported (e.g., only grand totals but data incomplete for any of the grids) were partial responses; no data reported in any grid was a nonresponse. For information about the methods

- **Unit response.** In 2015, the GSS received complete responses from 12,714 (83.6%) of the 15,202 eligible units. An additional 2,405 units (15.8%) were partial respondents. The remaining 83 units (0.5%) were nonrespondents.

- **School responses.** Of the 824 eligible schools, 810 schools (98.3%) were complete respondents (i.e., 90% or more of the school’s units provided complete or partial data), 3 schools (0.36%) were partial respondents (i.e., at least 50% but less than 90% of the school’s units provided complete or partial data), and 11 schools (1.33%) were nonrespondents (less than 50% of the schools units provided data).

- **Institutional response.** Institutional response rates were calculated using the same criteria for schools. Of the 711 eligible institutions, 697 institutions (98.03%) were complete respondents, 4 institutions (0.56%) were partial respondents, and 10 institutions (1.41%) were nonrespondents.

**Data editing.** Data quality is ensured by interactive edit checks built into the Web survey and by a comprehensive review after the coordinator submits the data. Data collection grids in the Web survey were prefilled with zeros. Respondents were asked to mark a checkbox if the unit does not have eligible data to report. Grids with a marked checkbox contributed to a complete response for the unit. Grids with unchanged prefilled zeros and an unmarked checkbox disqualified the unit from complete response status.

The edit checks built in the Web survey verify that the data entered are internally consistent and are within an expected range, often based on the previous year’s data. During follow-up, unit respondents are asked to explain the discrepancy whenever counts differ substantially from those of the previous year. The survey contractor reviews all data submitted by the academic institutions to ensure that all data fields are complete and are internally consistent. These quality checks are conducted when counts remained identical to the previous year and also when the school’s unit list, total counts, and distribution of counts had notable changes as follows:

- Changes to the unit list included unit additions and deletions, changes to the highest-degree-granted status, GSS code, or unit name.

- Units with total counts or a distribution of counts within a given data item that were substantially different from the previous survey cycle.

- Units with cell counts more than 20% above or 20% below its corresponding prior-year data.

Data fluctuations that were not sufficiently explained by the comments provided by the respondents during data collection were flagged for follow-up by e-mail or telephone call to the coordinator. Revisions were made directly in the Web survey by the coordinator, unit respondents, or survey contractor staff at the direction of the coordinator. See “Survey Quality Measures” below for a discussion of the types of measurement error detected in the data review and follow-up process.

**Imputation.** Of the 355 data items collected in the GSS, the item nonresponse rates ranged from 1.4% to 7.1%. All missing data are subject to imputation. Different imputation techniques were used for units with and without comparable historical data. For units missing a key total (total part-time students, total postdocs, or total NFRs) with at least 1 year of qualified historical data, a carry-forward imputation method was used. Inflation factors were calculated for the key totals to account for year-to-year change. The previous year’s key totals were then multiplied by these inflation factors to calculate the imputed
values for the current year’s key totals. Finally, all other variables were imputed by distributing the imputed key totals according to the previous year’s proportions.

For units that reported totals but no details, details were imputed according to the prior distribution if qualified historical details were available. Otherwise, a nearest neighbor imputation method was used. In this method, a donor unit that was “nearest” to the unit whose data were being imputed (imputee) was identified among all responding units having similar characteristics as the imputee (such as having the same GSS code for program fields and offering a doctoral degree). When graduate student details were imputed, the nearest neighbor selected had full-time and part-time graduate enrollments that were most similar to the imputee’s enrollments. The imputed values were calculated by adjusting the donor’s values to account for the difference in full-time and part-time enrollment totals between the two units.

Similarly, when postdoc or NFR details were imputed, the total number of postdocs or NFRs, respectively, was used to choose the nearest neighbor. If the postdoc or NFR total was missing, the graduate student totals were used to select the nearest neighbor to impute the postdoc or NFR variables. If either the postdoc or NFR key total (or both) was missing, other available key totals were used to select the nearest neighbor to impute the data. The same donor was then used to impute the details corresponding to the imputed key totals.

In rare circumstances, when no graduate student data were available from a new unit, IPEDS completions and enrollment data were used to estimate graduate student totals. Based on the imputed totals, the details were then imputed by the nearest neighbor method described above. Because IPEDS does not collect data on postdocs and NFRs, a nearest neighbor was selected from the GSS data to estimate these counts, if necessary, using the graduate student totals to select a donor. For non-responding units in institutions that had not been in the GSS before, postdoc and NFR values were imputed as zero rather than using IPEDS-based imputation.

Detailed information on the institutions, schools, units, fields, response rates, imputation rates, and a crosswalk between the 2010 CIP codes and the GSS codes are provided in 17 technical tables for the 2015 GSS.

**Weighting.** Not applicable.

**Variance estimation.** Not applicable.

**Survey Quality Measures**

**Sampling error.** Not applicable.

**Coverage error.** The availability of comprehensive lists of the master’s- and doctorate-granting institutions in the United States and their high levels of participation in the survey ensures that the coverage error of institutions is minimal. The universe of higher education institutions is regularly reviewed to identify potentially eligible institutions.

**Nonresponse error.** The GSS typically has high response rates. In 2015, 99.5% of units provided complete or partial data and the overall institutional response rate was 98.6%. Of the 355 data items collected in the GSS, the item nonresponse rates ranged from 1.4% to 7.1%.

**Measurement error.** The GSS is subject to measurement error that arises when variables of interest cannot be precisely measured. Review of the data, cognitive interviews, usability tests, pilot tests, site visits and other methodological activities with the institutions have pointed to a number of possible sources of measurement error, listed below.

- **Double counting.** Anecdotal evidence indicates some misreporting may occur when an institution has more than one coordinator or offers joint programs. To reduce double counting, facilitate
communication, and allow sharing of reported data, a screen in the Web survey provides names and contact information for all school coordinators at the institution. Interactive and post-submission checks are also used to confirm that similarly named units within institutions are distinct eligible units.

- **Inclusion of practitioner degrees.** Graduate students working toward practitioner degrees, particularly in health fields with explicit exclusions may sometimes be over reported. Starting with the 2007 survey cycle, survey materials indicated that students pursuing master’s, DDS, MD, and certain other degrees in specified fields should be excluded from the counts. During the imputation process, new units that were suspected of having reported graduate students in excluded degree-field programs based on the GSS code were set to having zero graduate students to be conservative, in the absence of other information. In the 2011 survey cycle, checks were built into the Web survey to remind respondents to exclude students pursuing practitioner-based degrees.

- **Difficulty in reporting source and mechanism of support.** Feedback from respondents and methodological research indicates that financial support data are often difficult for respondents to report. The information may not be stored in one centralized database; financial support may not always be channeled through the institution (e.g., self-support); and foreign sources of support may not always be known. Respondents may also have difficulty categorizing financial information by field, such as when a student is enrolled in one unit but receives support from another. Therefore, these data may be more prone to measurement error than other survey data items. Finally, institutions define mechanisms of support differently (e.g., fellowships versus traineeships) and may report individuals according to the institution’s definition rather than that provided by the GSS. Beginning with the 2010 survey, the grids include “unknown” categories.

- **Difficulty in reporting postdocs and NFRs.** Many respondents indicate in the Web survey that they are unable to provide data on their unit’s postdocs or NFRs because they do not know all of the units that employ postdocs and NFRs. Starting with the 2010 survey cycle, schools were given the option of appointing a separate postdoc coordinator who may be more knowledgeable about a school’s postdocs or NFRs to provide these data.

**Data Comparability**

*Changes in survey coverage and population.*

- **Fields of study.**

2011: GSS-eligible, degree-granting programs were updated from the 2000 Classification of Instructional Programs (CIP) taxonomy to the 2010 CIP taxonomy. A total of 58 new 2010 CIP fields were mapped to GSS codes, 14 CIP fields were moved between GSS codes, and 24 CIP fields were removed as ineligible. The impact on field-level counts was typically small and did not change the overall trend from 2010 to 2011. A crosswalk between the 2010 CIP codes and the GSS codes are provided in table A-16. The taxonomy changes yielded a net increase of 0.2%, 0.1%, and 0.2% in the total number of graduate students, postdocs, and NFRs, respectively in 2011. For more details, see the Technical Notes section of *Graduate Students and Postdoctorates in Science and Engineering: Fall 2011* at https://www.nsf.gov/statistics/nsf13331/.

*Institutions and Units.* In 2014, the survey frame was updated following a comprehensive frame evaluation study. The study identified potentially eligible but not previously surveyed academic institutions in the United States with master’s- or doctorate-granting programs in SEH. Eligible units at 151 newly eligible institutions were added, and two private for-profit institutions offering mostly practitioner-based graduate degrees were determined to be ineligible. Four additional
institutions dropped out of the data collection in 2014 because they no longer grant graduate
degrees in SEH fields; two merged with previously eligible institutions; and one began reporting
data under another institution. As a result, the total number of institutions included in the GSS
increased from 564 in 2013 to 706 in 2014. The total net increase in the number of GSS-eligible
units was 826, rising to 14,845 in 2014 from 14,019 in 2013 (see table A-1). For more
information on the survey frame update, see the special report Assessing the Impact of Frame

- **Eligibility and degree-granting status.** Institutions are classified as doctorate-granting if at least
one GSS-eligible unit confers doctoral degrees. In 2015, eight institutions became newly eligible
for GSS. Seven institutions changed GSS degree-granting status; two from doctorate-granting to
master’s-granting institutions and five from master’s-granting to doctorate-granting institutions.
Three institutions merged into two other institutions (see table A-2). As a result, the total number
of institutions included in the GSS increased from 706 in 2014 to 711 in 2015 (see table A-3).

**Changes in survey content.**

- **Sex.**

  2010: Began collecting citizenship, ethnicity, and race data on postdocs by sex, and type of
doctoral degree data on NFRs by sex.

  2008: Began collecting the number of first-time, full-time male graduate students by ethnicity and
race; full-time male graduate students by source of support; male postdocs by source of support;
and male NFRs. Previously, the number of men was inferred by subtracting the number of women
from the total.

- **Ethnicity and race.**

  2010: Began collecting ethnicity and race data for postdocs who are U.S. citizens and permanent
residents using the same categories as used for graduate students.

  2008: Revised ethnicity and race categories to correspond to IPEDS by combining
“Hispanic/Latino, One race only” and “Hispanic/Latino, More than one race” categories into
“Hispanic/Latino (one or more races).”

- **Citizenship.**

  2010: Began collecting citizenship data on postdocs using the same categories as used for
graduate students. In previous years, only counts of postdocs who are foreign nationals holding
temporary visas were collected.

  2008: Clarification made for “non-U.S. citizens” to exclude non-U.S. citizens residing outside of
the United States who are enrolled in an online degree program at a U.S. institution.

- **Financial support.**

  2010: Began collecting data on the largest source of financial support and on the largest
mechanism of support separately for postdocs. For mechanism of support, “nonfederal sources”
was replaced with “other support.”

  2008: Graduate student data no longer collected for NIH teaching assistantships because NIH
does not offer financial support for students through this mechanism.

  2008: Began collecting number of full-time graduate students whose largest source of support
came from a non-U.S. source via teaching assistantship.
• **Doctoral degree.**

2010: Began collecting more detailed information on postdocs’ and NFRs’ doctoral degree type. Categories were added for those holding a doctoral degree (e.g., PhD, ScD, DEng), a professional degree (e.g., MD, DVM, DO, DDS), and dual degrees (MD-PhD, DVM-PhD), as well as for those for whom type of degree was unknown. In previous years, the degree-type question simply asked the respondents to report the postdocs with MD, DO, DDS, DVM from the total postdoc count.

2010: Began collecting postdocs’ doctoral degree type by citizenship and by country of origin (United States, foreign, unknown) of doctoral degrees. Also began collecting NFRs’ doctoral degree type by sex.

*Changes in Web survey instrument.* Modifications were made to the following four areas of the 2015 GSS Web survey instrument:

- **Update Unit List:** An option was added at the bottom of the screen to view any units deleted in the current cycle, and provided coordinators with the ability to review and to “undelete” units that had been inadvertently deleted.

- **Full-time student financial support table:** An addition was made to the table collecting full-time student financial support data. If a coordinator indicated that any students were supported by “other federal sources,” a dialog box would appear that asked coordinators to select the specific federal source from a list provided. If the source was not on the list, a comment box was provided for respondents to enter the name of the other federal source. This dialog box could be accessed again by clicking the “other federal source” title in row H of the survey.

- **Unit Status Menu:** Revised to allow coordinators to “lock” individual columns within the unit (i.e., specific data grids); previously it was only possible to lock an entire unit at one time.

- **Nonfaculty researcher and postdoc demographic tables:** In an effort to prevent double-counting of postdocs, a check was added for the scenario in which a unit reported the same number of male and female postdocs and NFRs. When the postdoc and NFR counts matched by sex, respondents were required to check a box to confirm that the postdoc and NFR counts were for unique individuals.

*Changes in survey procedures.*

- **2013:** Three different versions of the launch e-mail were used. One version was sent to coordinators who used the data file upload feature in 2012, informing them of updates. An alternate version was sent to coordinators who might benefit from using the upload feature (i.e., coordinators with a large number of units and who were not currently using the upload feature). The third version was the standard launch e-mail, with no mention of the upload. In addition, the data review and retrieval efforts began in January, earlier than in prior years.

- **2010:** Significant effort was made to ensure that appropriate personnel were providing postdoc and NFR data. As a result, it is unclear how much of the increase reported in 2010 represented actual growth in postdocs and how much resulted from improved data collection. For information on the improved data collection and changes in postdoc data, see *Counts of Postdoctoral Appointees in Science, Engineering, and Health Rise with Reporting Improvements* at [http://www.nsf.gov/statistics/infbrief/nsf13334/](http://www.nsf.gov/statistics/infbrief/nsf13334/), and for changes in NFR data, see *Examining the Reporting of Nonfaculty Doctorate Researchers in the Survey of Graduate Students and Postdoctorates in Science and Engineering* at [https://nsf.gov/statistics/2015/ncses15201/](https://nsf.gov/statistics/2015/ncses15201/).
Historical changes. Changes have been made over the years to the coverage and content of the GSS to keep it relevant to the needs of data users. Such changes impact analysis of trend data and data comparisons across years should be made with caution. This is especially true for counts; however, proportions or shares are typically robust enough to allow for such comparisons.


Definitions

Enrollment status.

- **Full-time and part-time**—Coordinators were instructed to use their institution’s definitions.

- **First-time, full-time**—Students enrolled for credit in a graduate degree program in an organizational unit for the first time in fall 2014. This may include graduate students previously enrolled in another graduate degree program at the institution or at another institution and students who already hold another graduate or professional degree.

Ethnicity and race—The GSS uses definitions of ethnicity and race that are based on the OMB’s “Standards for the Classification of Federal Data on Race and Ethnicity.”

- **Hispanic/Latino ethnicity (one or more races)**—All individuals of Cuban, Mexican, Puerto Rican, South or Central American, or other Spanish culture or origin, regardless of race. This category includes individuals who are Hispanic or Latino and any other race(s).

- **Not Hispanic/Latino**—Individuals who are not of Hispanic or Latino descent, regardless of race.

  - **American Indian or Alaska Native**—A person of only one race having origins in any of the original peoples of North and South America (including Central America) and who maintains tribal affiliation or community attachment.

  - **Asian**—A person of only one race having origins in any of the original peoples of the Far East, Southeast Asia, or the Indian subcontinent—for example, Cambodia, China, India, Japan, Korea, Malaysia, Pakistan, the Philippine Islands, Thailand, and Vietnam.

  - **Black or African American**—A person of only one race having origins in any of the black racial groups of Africa.

  - **Native Hawaiian or Other Pacific Islander**—A person of only one race having origins in any of the original peoples of Hawaii, Guam, Samoa, or other Pacific islands.

  - **White**—A person of only one race having origins in any of the original peoples of Europe, the Middle East, or North Africa.

  - **More than one race**—A person of two or more of the race categories listed above.

- **Unknown ethnicity or race**—A person whose ethnicity or race is unknown or not stated.
Graduate student mechanisms of financial support.

- **Fellowship**—A competitive award (often from a national competition) given to a graduate student that requires no work of the recipient.
- **Traineeship**—A financial award given to a graduate student selected by the institution.
- **Research assistantship**—A financial award given to a graduate student where most of the student's responsibilities are devoted primarily to research.
- **Teaching assistantship**—A financial award given to a graduate student where most of the student's responsibilities are devoted primarily to teaching assistant activities.
- **Other support**—All other mechanisms of support for graduate students.

Graduate student sources of financial support.

- **Federal sources**—Financial support provided by the federal agencies. Excludes federally guaranteed student loans.
- **Nonfederal sources**—Financial support from state and local governments; support from the institution, such as tuition waivers and stipends; support from foreign sources, such as foreign governments, foreign firms, and agencies of the United Nations; and other U.S. sources, such as support from nonprofit institutions, private industry, and all other nonfederal U.S. sources.
- **Self-support**—Loans (including federal loans) or personal or family financial contributions.

Historically black colleges and universities (HBCUs)—Institutions of higher education that were established prior to 1964, whose principal mission was, and is, the education of black Americans. The list of HBCUs is maintained by the White House Initiative on HBCUs at [http://sites.ed.gov/whhbcu/](http://sites.ed.gov/whhbcu/).

Nonfaculty researchers—All doctorate-holding researchers who (1) are not considered either postdocs or members of the faculty, and (2) are involved principally in SEH research activities. Also referred to as Other doctorate-holding nonfaculty researchers.

Postdoctoral researchers (postdocs)—The definition of a postdoc varies by institution. Respondents were instructed to use their institution’s definition. NCSES defines a postdoc as meeting both of the following qualifications: (1) holds a recent doctoral degree, generally awarded within the last 5–7 years, such as PhD or equivalent (e.g., ScD, DEng); or first-professional degree in a medical or related field (e.g., MD, DDS, DO, DVM); or foreign degree equivalent to a U.S. doctoral degree and 2) has a limited-term appointment, generally no more than 5–7 years, primarily for training in research or scholarship, and working under the supervision of a senior scholar in a unit affiliated with the institution.

Postdoc mechanisms of financial support.

- **Traineeship**—A financial award given to a postdoc selected by the institution.
- **Research grant**—A financial assistance award given to an organization or an individual postdoc that supports specific research goals.
- **Other support**—All other mechanisms of support for postdocs.

Postdoc sources of financial support.

- **Federal sources**—Financial support provided by U.S. federal agencies.
- **Nonfederal sources**—Financial support from state and local governments; support from the institution; support from foreign sources, such as foreign governments, foreign firms, and agencies of the United Nations; and other U.S. sources, such as support from nonprofit institutions, private industry, and all other nonfederal U.S. sources.
• **Personal resources**—Personal and family financial resources, including federal and other loans.

• **Unknown or not stated**—Sources of financial support for the postdoc are unknown or cannot be determined.

**Notes**

1 In this report, the term school refers to a graduate school, medical school, dental school, nursing school, or school of public health; an affiliated research center; a branch campus; or any other organizational component within an academic institution that grants an SEH degree.


3 The OMB standards designate Hispanics as an ethnic group rather than a racial group. Following these standards, Hispanic is not counted as a race in GSS. Cognitive interviews with respondents have shown that this is a source of considerable confusion. For example, black Hispanics and white Hispanics may be counted as “Hispanic, More than one race” rather than “Only one race, Hispanic.” The ethnicity and race categories were aligned to IPEDS by combining the “Hispanic/Latino, More than one race” and “Hispanic/Latino, One race only” categories. In 2008 these two Hispanic categories were collapsed into one: “Hispanic/Latino ethnicity (one or more races).”