Methodological Study on the Reporting of Doctorate-Holding Nonfaculty Researchers to the Survey of Graduate Students and Postdoctorates in Science and Engineering—Outcomes and Recommendations

March 2022

Prepared for

National Science Foundation
2415 Eisenhower Avenue
Alexandria, VA 22314

Prepared by

RTI International
3040 Cornwallis Road
Research Triangle Park, NC 27709

RTI Project Number 0216545.000.009.001

Contents

Figures

Number **Page**

Tables

Number **Page**

Executive Summary

The Survey of Graduate Students and Postdoctorates in Science and Engineering (GSS) collects data on the science, engineering, and health (SEH) workforce pipeline to include graduate students, postdoctoral researchers, and other doctorate-holding nonfaculty researchers (NFRs). 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 collected by GSS to help determine the research intensity of colleges and universities. The Carnegie designation of schools as very high research, high research, or doctorate/professional universities is partially informed by GSS data.

A 2012 GSS debriefing survey noted that institutional coordinators (ICs) struggled to define and identify NFRs on their campus and that institutional data systems often lacked the necessary information (e.g., highest degree obtained by employees) required to accurately report NFRs. Anecdotal evidence from the GSS help desk and recent site visits further supported the notion that NFR reporting presented challenges for ICs and that they held varying interpretations of the NFR definition provided by the survey.

RTI International conducted a series of investigations to validate reported NFR data against other federal data collections and to gain further insight into the challenges that ICs face when applying the GSS definition of NFR in their own institutional context, offering an alternative NFR definition in order to improve consistency of NFR reporting across the survey universe. This report offers results of investigations performed to date and makes recommendations for future methodological activities. Among the key findings are the following:

- Certain job titles collected in the Early Career Doctorates Survey might be used to provide GSS respondents with examples of research staff that fit the NFR definition.
- Comparison between data reported to the Higher Education Research and Development Survey (HERD) and the Integrated Postsecondary Education Data System (IPEDS) found limited but significant correlations between the number of NFRs reported to GSS and research expenditures and the number of research staff without faculty status.
- Research also suggests, however, that institutions report research staff inconsistently across these three surveys. For example, 115 institutions reported NFRs to the GSS but did not report any research staff without faculty status to IPEDS.
- Fifteen virtual site visits with selected GSS respondents reinforced previous findings that many institutions consider NFR reporting to be a complicated process that requires considerable judgment calls on what types of positions and job titles qualify as NFRs. Additionally:
  - Some coordinators strictly follow the definitions provided by GSS instructions regarding faculty status as a disqualifier, while others include research faculty job titles as being within the spirit of the rule when it comes to NFR reporting.
  - Support exists for an alternative definition of researchers offered by the project team that removed the term “faculty” from the definition. Specifically, *campus staff who are primarily engaged in research, applied research, or development; hold a doctorate or equivalent; and are not on the tenure track.*
  - This report recommends a record-keeping survey distributed to all institutions in the GSS universe that would explore data and reporting capabilities on research staff and would obtain additional feedback on the alternative definitions for researchers investigated in the site visits.

1. Introduction

The Survey of Graduate Students and Postdoctorates in Science and Engineering (GSS) has collected data on the science, engineering, and health (SEH) workforce pipeline annually since 1972. Data collection was revised in 1979 to offer more clarity regarding the workforce pool, separating postdoctoral researcher appointees (postdocs) and other doctorate-holding nonfaculty researchers (NFRs). This change was implemented to improve the counts of postdocs, with the NFR counts initially serving as a residual category to keep research staff from being erroneously reported as postdocs.

In many ways, the legacy of NFR as a residual category persists in the GSS data collection. 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 2019 survey reported information on 690,117 graduate students, 66,247 postdocs, and 30,349 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.

Historically, NFR data maintained stable growth, increasing by about 7% per year from 1979 through 2007 (see figure 1). Between 2007 and 2010, there were two large increases in reported NFRs—a 28% increase between 2007 and 2008 and a 52% increase between 2009 and 2010. A 2015 working paper attributed these large increases to methodological changes implemented during this period, including improving coverage of research centers and other non-degree-granting organizational units, recruiting postdoc- and NFR-specific coordinators at schools where the main graduate student coordinator had difficulty accessing the postdoc and NFR data, and separating the postdoc and NFR items onto separate forms in the survey instrument.[1] The working paper noted that while methodological changes led to the large increases in NFR counts between 2007 and 2010, NFR data were relatively stable from year to year within institutions and found that institutional NFR counts were highly correlated with research expenditures reported to the Higher Education Research and Development Survey (HERD).

However, a debriefing survey conducted after the 2012 GSS data collection noted that institutional coordinators struggled to define and identify NFRs on their campus and that institutional data systems often lacked the necessary information (such as highest degree obtained by employees) required to accurately report NFRs. For example, of the 91 responding schools in the debriefing survey, only 24 of them reported that they had a common definition of NFR in their programs.[2] Meanwhile, nearly 60% of responding schools that reported zero NFRs indicated that these researchers probably did exist on their campus, but they were unable to report on them.

RTI International conducted a series of investigations to better appreciate the challenges that institutions face when dealing with the GSS definition of NFR. As described above, the investigations were necessary because institutional interpretations of the GSS definition of NFR vary. This complicates comparisons of NFR data between institutions. Such comparisons are critical to our understanding of the SEH workforce, the Carnegie Classification ranking system, and institutional benchmarking work. The following questions orient this investigation:

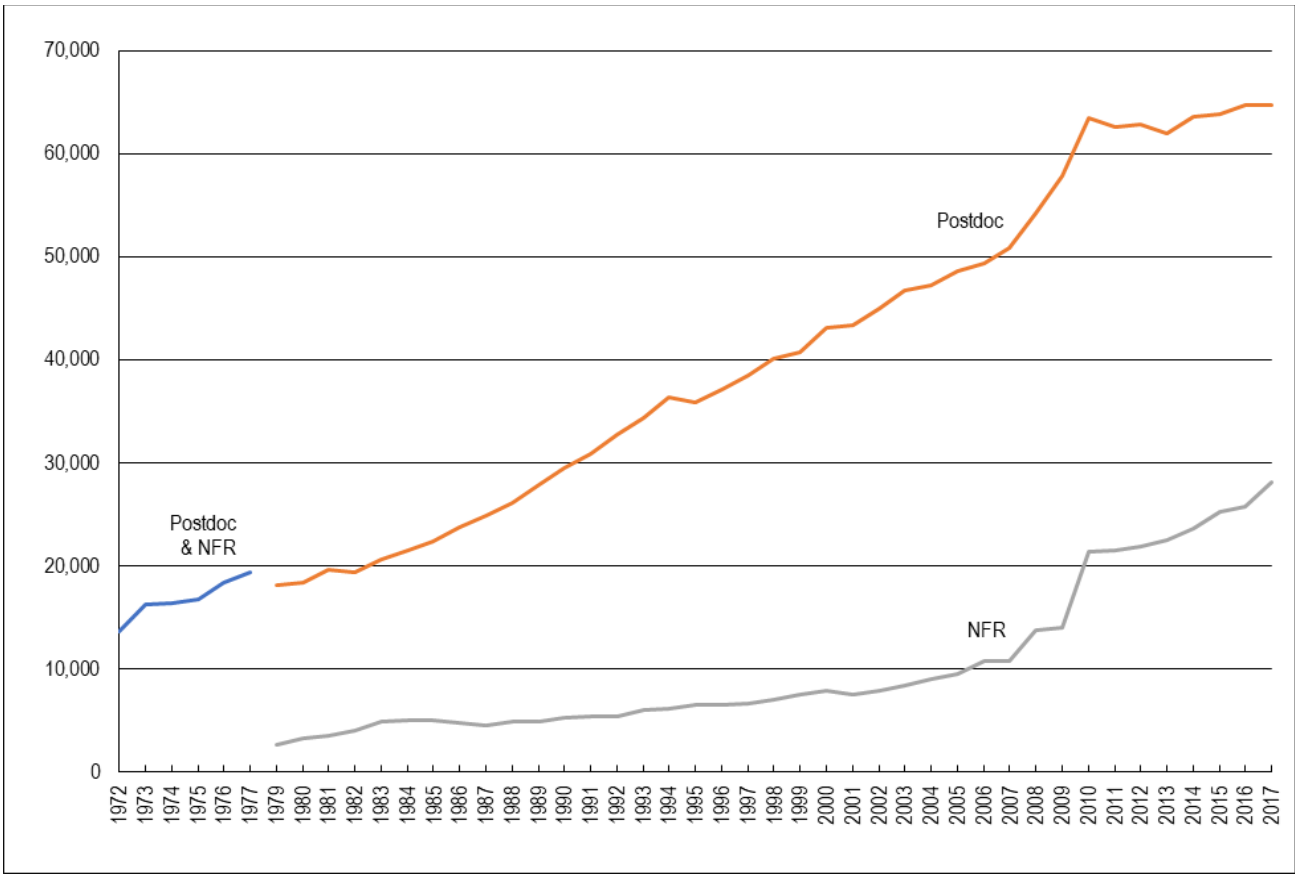
- Can NFR reporting be made more consistent across institutions?
- Can the definition of NFR be modified to make reporting easier for institutions?
- Are there ways NFR data can be made more useful to the National Center for Science and Engineering Statistics (NCSES) and its stakeholders?

This report presents the results of the investigations performed to date and makes recommendations for future methodological activities to enhance consistency across respondents, reduce the burden of the data request, and improve the utility of the collected data. RTI staff conducted the following work as part of this effort:

- Analysis of job titles involving research activity at institutions participating in the NCSES Early Career Doctorates Survey (ECDS)

- Review of websites and policy documents of GSS institutions to determine how respondents currently define NFR and other research staff
- Comparison of reported NFR data against research and employment data reported HERD and the Integrated Postsecondary Education Data System (IPEDS)
- Virtual site visits to 15 GSS participants across a variety of institutional characteristics and NFR reporting abilities

Figure 1. Postdoctoral researcher and NFR reporting in GSS: 1972–2017



GSS = Survey of Graduate Students and Postdoctorates in Science and Engineering; NFR = doctorate-holding nonfaculty researcher.

1.1 Analysis of relevant job titles from the ECDS

The methodological study investigated data from the ECDS to assess whether job titles captured by the survey could improve instructions to GSS respondents regarding whom on their campus to report as NFRs. The ECDS was a two-stage sample of individuals who earned their doctorate or equivalent between 2007 and 2017. The first stage of data collection sampled GSS institutions, including federally funded research and development centers, while the second stage sampled early career doctorates directly. This allowed for administrative data reported by institutions to be compared against responses and impressions of the actual employees. Because the administrative data reported to ECDS closely matches those collected by the GSS, the analysis offered some insight into the roles and responsibilities associated with administrative job titles.

A key component of the methodological study was an analysis of the job titles reported by institutions in the first stage of data collection compared to the job functions and responsibilities reported by the early career doctorates themselves. The goals were to use research-based responsibilities to identify research staff who might be relevant to the GSS. This analysis was supported by the over 10,000 ECDS responses from GSS institutions.

ECDS respondents could choose one of eight options as their primary or secondary professional responsibilities[3]. The job activity options that related to *research* included *basic research*, *applied research*, and *development*. Taken together, about 50% of all early career doctorate respondents listed one of the three research-related activities as the primary activity of their current position. When considered as a function of job title, research accounted for over 50% of primary activities among nine sets of job titles. These titles typically included the terms *research*, *scientist*, or *engineer*.

Further investigation into the key job title terminology found that more than 80% of those early career doctorates with "research" in their job title report research and development as their primary professional responsibility. Of this group, nearly 70% had some level of faculty status, either as non-tenure-track faculty, tenure-track faculty, or tenured faculty. The ECDS data indicate that job titles with *research*—and, to a lesser extent, *scientist* or *engineer*—index those staff performing research with high accuracy. Further, many of the staff engaged in research do have faculty status. Non-tenure-track faculty research positions often look similar to nonfaculty research positions, both in terms of job title and primary activity. The review of ECDS data suggests institutions are aware of their core research staff and apply specific job titles to identify them.

The job titles collected by ECDS might serve as examples that might be provided to respondents in instructional materials for GSS data collection of NFRs. These findings also informed the creation of alternative definitions of NFR that were subsequently discussed with site visit participants.

1.2 Review of institutional policies on employment of research staff

In parallel to the job title analysis, RTI also conducted a review of institutional websites to look for classifications and employment policies of NFRs and other research staff. The goal was to determine what publicly available employment policies and job information exist within various institutions that might guide their decisions about whom to include and exclude when reporting NFRs. The review included targeted internet searches of terms such as "nonfaculty research staff" or "university research staff" at GSS-participating institutions.

Internet searches helped us to identify a dozen NFR definitions and guidelines published online by institutions. Please note that the 12 examples do not represent an exhaustive overview of all available NFR information online. Instead, they represent the most relevant hits to our targeted search criteria.[4]

The Harvard School of Public Health published a document regarding "non-faculty research titles" that separates five job titles into three job categories, including research/postdoc, research associate, or research scientist.[5] Each job title includes a definition of role, responsibility, and promotional guidelines. The file also includes details on translational activities that are expected for each job title.

The University of Nevada, Las Vegas (UNLV), offered a different perspective in a document regarding policies on postdocs and research staff.[6] This policy document introduces the GSS definition of NFR and then notes that position/job titles for non-postdoc NFRs vary greatly across campus, continuing to report that "the most common NFR title is assistant or associate research professor." While *professor* often denotes faculty status, the UNLV document does not explicitly make that distinction. If that was the case, then this document would indicate that the institution knowingly reports ineligible research staff—a phenomenon that was noted in several site visits.

The Office of Institutional Research at the University of Memphis maintains the most clearly defined account of NFRs that this analysis encountered. Representatives from the institution published a short document that includes results of a peer analysis as well as a five-step flow chart that determines whether specific employees should be reported to the GSS (figure 2). The flow chart demonstrates how complex the categorization of research staff can be and is indicative of the challenges institutions face in determining whom to report as NFRs to the GSS.

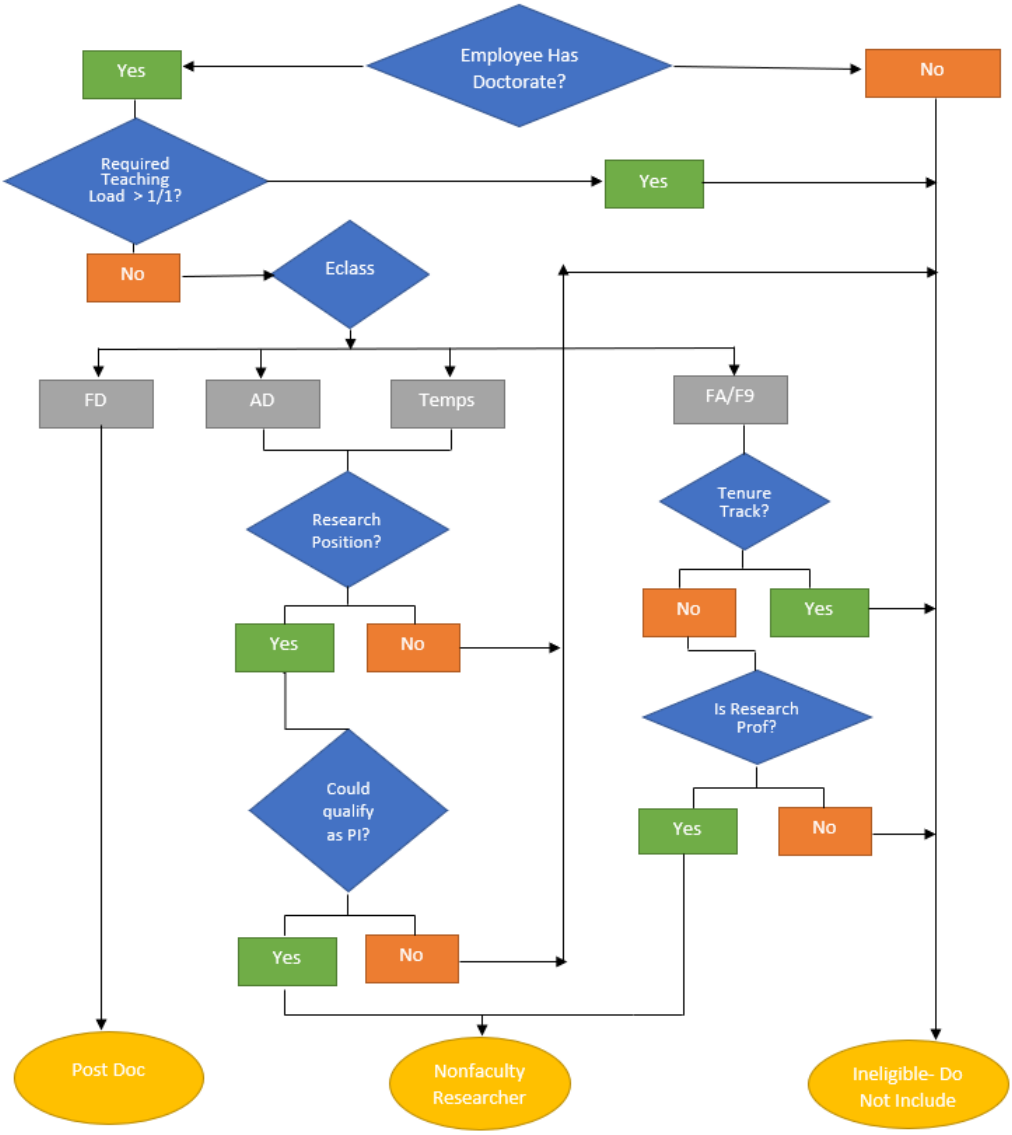
The review also included several high-profile institutions in the GSS universe that routinely state that they have zero eligible staff to report as NFRs. This exploration yielded mixed results. In a few cases, the reason behind zero NFRs reported was clear—research staff who held doctorates all had faculty status (i.e., *research professor* titles). Because faculty status is tied to the research position, the individuals do not qualify as NFRs and are not eligible for the GSS.

The exploration of institutional NFR definitions and the investigation of NFR nonreporters combine to suggest that institutions may not have a single, stand-alone category of employee that aligns with the GSS NFR data request. Instead, institutional representatives interpret the GSS definition of NFR and then build a crosswalk to identify eligible staff by job title or job description from their

institutional database. As evidenced by the University of Memphis flowchart, the interpretation of the GSS definition of NFR can be a complex process.

This online review uncovered NFR definitions that are inconsistent across the survey universe and that the data reported to the GSS do not necessarily reflect an institution's internal definition of this employment category. At the heart of this inconsistency is the fact that a feature of the job title or job description that can cause one institution to deem a staff member ineligible for reporting could be the same feature that a peer institution uses to define their staff as eligible.

Figure 2. University of Memphis GSS nonfaculty research decision tree



AD = 12-month administrative or academic professional; FA = 12-month faculty; FD = postdoctoral; PI = principal investigator.

Source(s): Office of Institutional Research. 2019. *Non-Faculty Researcher and Post Doc Definitions*, p. 3. Memphis, TN: University of Memphis. https://www.memphis.edu/aa/resources/docs/2019_nonfac_researcher_postdoc_definitions.docx.

1.3 Comparison of GSS NFR data against research and employment data in the HERD and IPEDS surveys

The current methodological study sought to extend the data validation analysis conducted in 2015—which compared GSS NFR counts to HERD R&D expenditures—by also comparing NFR counts to research staff reported to IPEDS. One of the primary goals of this effort was to identify institutions whose NFR counts aligned with data reported to HERD and IPEDS as well as to identify institutions that did not align. Institutions so identified became candidates for site visits with the GSS data collection team to further explore their

NFR reporting capabilities and the challenges they encounter in identifying appropriate research staff to GSS. A secondary goal was to assess whether the HERD or IPEDS data might serve as alternatives to NFR data collection by the GSS.

HERD was considered a relevant data source because it has the same federal sponsors as GSS and because there is considerable overlap between the universe of both surveys. Research expenditures reported by HERD can serve as a proxy for research staff in that the more money spent on research likely correlates with more staff employed to conduct that research.

The U.S. Department of Education relies on IPEDS surveys to collect data from every postsecondary education institution in the country, including virtually all those in the GSS survey universe. IPEDS reports the number of staff engaged in research as well as the amount of research expenditures at each participating institution. Though IPEDS staff definitions do not have a direct analog to NFR, a review of published documents as well as conversations with the IPEDS data collection team made it clear that *research staff without faculty status* represent the closest available category.

GSS, HERD, and IPEDS data from 2019 combine to represent 719 institutions. Institutions with multiple schools that report data separately to the GSS were merged in this analysis, a process that ensured that all 809 schools in the GSS survey universe were represented in the analysis. Each institution had both GSS and IPEDS data, while 626 schools also had HERD data. Research expenditure data from IPEDS were applied in cases where HERD data were not available. Taken together, all the institutions analyzed had NFRs from the GSS and had information on research staff from IPEDS, while 92% of the institutions had data on research expenditures from IPEDS or HERD.

The comparison included three steps: a quartile-based comparison of data in the surveys, a correlation analysis of variables between the surveys, and a regression analysis that included a predictive model regarding the number of NFRs that each institution should report was based on data reported in other variables.

The quartile-based assessment of NFRs reported to the GSS, research staff reported to IPEDS, and research expenditures reported to HERD and IPEDS tested the relationship between the three surveys. For each data element, institutions were assigned a quartile within the range of reported values. Table 1 presents an inventory of the data elements included in this analysis. The expectation was that if the variables were reported consistently, then similar variables (i.e., the number of NFRs reported to GSS, research staff reported to IPEDS, and research expenditures reported to IPEDS and HERD) would code into similar quartiles across the various surveys. For example, one would expect that an institution that is in the highest quartile for research expenditures would also be in the highest quartile for NFRs reported to GSS. If the reporting was not consistent across the surveys, then one would expect that the off-diagonal cases would identify schools where further investigation would help us understand issues that institutions face when reporting NFRs to the GSS.

Results of the quartile analysis (presented in appendix A) suggest that institutions report the data inconsistently across the surveys. The most apparent inconsistency is the comparison between research staff reported to the surveys. For example, 115 institutions reported NFRs to the GSS but did not report any research staff without faculty status to IPEDS. Among institutions that did report NFRs, there was some agreement among the highest-quartile cells. While the data sources do not fully align with one another, the limited agreement among elements suggest that HERD and IPEDS could serve as predictors for the number of NFRs that an institution reports to the GSS. The off-diagonal comparisons presented in appendix tables A-3–A-6 were informative in identifying how reporting at institutions differed across surveys.

Table 1. Data elements used to compare federal surveys

Data element	Source	Applied in quartile analysis	Applied in correlation matrix	Applied in regression analysis
Doctorate-holding nonfaculty researcher (NFR)	GSS	x	x	x
Postdoctoral researchers	GSS	x	x	-
All research staff	IPEDS	x	x	-
Research staff without faculty status	IPEDS	x	x	-
Research staff tenured or on tenure track	IPEDS	-	x	x
Research staff not tenured or not on tenure track	IPEDS	-	x	x
Research expenditures	HERD & IPEDS	x	x	x
Status of institution as public or private	GSS, HERD, & IPEDS	-	x	x
Status of institution as very high research activity that grants a medical doctorate	GSS, HERD, & IPEDS	-	x	x
Status of institution as very high research activity that does not grant a medical doctorate	GSS, HERD, & IPEDS	-	x	x
Status of institution as not very high research activity that grants a	GSS, HERD, & IPEDS	-	x	x

Data element	Source	Applied in quartile analysis	Applied in correlation matrix	Applied in regression analysis
medical doctorate				
Status of institution as not very high research activity that does not grant a medical doctorate	GSS, HERD, & IPEDS	-	x	x

GSS = Graduate Students and Postdoctorates in Science and Engineering; HERD = Higher Education Research and Development Survey; IPEDS = Integrated Postsecondary Education Data System.

These discoveries precipitated the next two analytical steps: a correlation matrix, and a regression analysis. The correlation matrix identified three variables with a moderate correlation to the number of NFRs reported to the GSS: research expenditures, research staff without faculty status, and institutional status of very high research activity that offers a medical doctorate. These results from the correlation matrix informed the variables included in the regression analysis (see appendix tables A-7 and A-8 for details).

The third and final step in this analysis was a regression analysis of reported NFR counts. The variables included in this step had been identified through the quartile analysis and verified in the correlation matrix. The regression captured a collective significant effect between the number of NFRs reported to the GSS and the variables chosen for consideration because the 2017 survey data yielded an R^2 value of 0.5797. When repeated on the 2018 and 2019 survey data, each regression yielded an R^2 value of at least 0.52.

The stepwise approach to the identification of variables with strong correlations to reported NFR counts, and the high R^2 values achieved by the regression analysis created a systematic method of identifying candidates for institutional site visits. The expectation was that those institutions that consistently did, or did not, meet expectations of the model would qualify for a site visit follow-up. The predicted NFR values were compared against actual reported values for the 2017–19 GSS survey cycles to identify the degree to which institutions' NFR reporting to GSS aligned with what was expected given their responses to HERD and IPEDS. For any given year, institutions were categorized in the following manner:

Aligned institutions were those where

- the institution was predicted to have reported more than 15 NFR, *and*
- the difference between reported and predicted NFRs was within 10%, *or*
- the overall difference between reported and predicted NFRs was less than 15 employees.

*Misaligne*d institutions were those where

- the model predicted the institution would have more than 15 NFRs, *and*
- the institution did report research expenditures to either HERD or IPEDS, *and*
- the number of NFRs predicted by the model exceeded the number reported to GSS, *and*
- the difference between reported and predicted NFRs was greater than 10%, *or*
- the overall difference between reported and predicted NFRs was more than 15 employees.

It should be noted that the misaligned group also included *nonreporters*, institutions that did not report any NFRs to the GSS between 2017 and 2019 but where at least 15 NFRs were predicted each year. Of the initial set of 719 institutions, only 140 met the requirements enumerated above. Across the three years of analyses, 81 institutions were consistently aligned or misaligned for all three years. An additional 59 institutions were either aligned or misaligned in at least one survey cycle. The 140 institutions identified through this process formed the pool of candidates for subsequent site visits.

2. Virtual Site Visits

Site visits allowed for GSS respondents to share their institutional definitions of NFR, methods of identifying NFRs from their institutional databases, and challenges with data reporting directly with representatives from the GSS data collection team. The site visits also afforded an opportunity to obtain feedback on two alternative definitions for NFRs under consideration by NCSES.

2.1 Site visit approach

Table 2 presents details regarding the 15 institutions included in this analysis. In selecting site visit candidates, the data collection team sought to include geographic diversity, representation of a variety of research intensities (as represented by Carnegie Classification of research intensity), public or private control, and participation from institutions that offer medical doctorates. Additionally, for those reporting NFR data to GSS, we initially characterized each institution's alignment with data reported to HERD and IPEDS based on the predicted outcomes of the regression model described above. Institutions whose NFR counts were within range of the predicted values for 2017–19 were considered *aligned* with the model, while those whose counts were outside the range in any of these three years were considered *not aligned*.

While categorizing alignment of reported NFR counts with the predicted values assisted in identifying appropriate candidates for institutional site visits, these labels did not offer much value in understanding the resulting conversations from the site visits. Institutional representatives from schools whose NFR counts aligned with the regression model and those that did not align both expressed some doubt about what NFR data were to be reported to GSS. A stakeholder from Auburn University, a school whose counts consistently aligned with the predicted model, went so far as to express their surprise when learning that the data collection team considered their institution a consistent aligned reporter. The institution agreed to participate in a site visit to better understand how we thought they were meeting expectations. Further, regardless of alignment with data reported to HERD and IPEDS, institutions described similar methodological steps and decision-making processes when collecting and reporting NFRs. Because categories of aligned with the model and not aligned with the model did not differentiate institutions by methodology or confidence in the data, we are choosing not to use these heuristics in summarizing the results of these discussions.

However, it is useful to distinguish institutions by whether they reported any NFR to the GSS—with the terminology of *nonreporter* (4 institutions) versus *reporter* (11 institutions). The site visit summary below explores these processes further, but nonreporters clearly cited one of two roadblocks to NFR reporting: all research staff are given faculty status and thus are ineligible as nonfaculty researchers, or educational attainment data are not reliably available, so doctorate-holding status cannot be verified. On the other hand, reporter institutions often work through issues that otherwise may preclude participation. This sometimes includes knowingly reporting researchers with faculty status or using job titles as proxies for educational attainment.

Table 2. Characteristics of institutions participating in NFR site visits

Institution name	Institution control	Carnegie R1: very high research activity	Has medical center	Reports NFRs?	Reported NFR aligned with HERD/IPEDS 2017–19
Auburn U.	Public	Yes	No	Yes	Aligned
Boston C.	Private	Yes	No	Yes	Not aligned
Georgia State U.	Public	Yes	No	No	Did not report NFRs
Johns Hopkins U.	Private	Yes	Yes	Yes	Not aligned
Massachusetts Institute of Technology	Private	Yes	No	No	Did not report NFRs
New Mexico State U.	Public	No	No	Yes	Aligned
North Carolina State U.	Public	Yes	No	Yes	Not aligned
Northwestern U.	Private	Yes	Yes	Yes	Not aligned
San Diego State U.	Public	No	No	No	Did not report NFRs
Texas A&M U.	Public	Yes	Yes	Yes	Not aligned
U. Alaska	Public	No	No	Yes	Not aligned
U. California, Davis	Public	Yes	No	Yes	Not aligned
U. Kentucky	Public	Yes	Yes	No	Did not report NFRs
U. Tennessee Health Science Center	Public	No	Yes	Yes	Aligned
U. Virginia	Public	Yes	Yes	Yes	Aligned

HERD = Higher Education Research and Development Survey; IPEDS = Integrated Postsecondary Education Data System; NFR = nonfaculty researcher.

NCSES encouraged the participation of various campus stakeholders—including provosts, deans of graduate and postdoctoral study, vice presidents for research, and institutional research staff. Four site visits included five or more institutional stakeholders, while only

one included a single representative. A total of 15 site visits were conducted as 1-hour virtual meetings on Zoom. The site visits were conducted between mid-August and mid-October 2021.

The GSS project officer from NCSES attended each site visit, with the GSS project director from RTI International acting as facilitator, guiding institutional representatives through a theme route designed to spark conversations around several topics (see appendix B):

- How the institution currently defines NFR.
- Characteristics of research staff for each campus—including which research staff are on the tenure-track, how definitions of faculty apply to research personnel, and the degree to which institutions maintain formal policies around the hiring and promotion of research staff.
- How the institution gathers data on NFRs for the GSS—including which institutional databases and other campus resources must be consulted to obtain the requested information for NFR.
- The type of data available regarding research staff that could be reported in the future.
- How campus stakeholders currently use GSS data.

The site visits also offered an opportunity to explore alternative definitions of NFRs or research staff that GSS could work toward in the future. Each visit included a discussion of two alternative definitions. The first proposed alternative, *doctorate-holding non-instructional researcher*, was derived in part from the analysis of ECDS job titles and the exploratory research on institutional policies regarding research staff. The other proposed alternative definition was taken directly from HERD because this NCSES survey also assesses research activities at many of the same institutions. HERD collects research expenditure data and, since at least 2016, added items regarding the number of personnel engaged in R&D activities. Based on definitions from the Organisation for Economic Co-operation and Development (OECD) *Frascati Manual*, in 2020 HERD began asking its respondents to differentiate three types of research staff—*researchers*, *R&D technicians*, and *R&D support staff*.[7] GSS would only be concerned with the first category of researchers as an alternative to NFRs. Prior to the site visit, participants were provided with definitions of all three R&D functions to properly contextualize the discussions. Please refer to table 3 for an overview of the definitions.

Table 3. Research staff definitions discussed during site visits

Source	Definition
GSS (current definition)	Doctorate-holding nonfaculty researcher: All doctorate-holding researchers who are (1) not considered either postdoctoral researchers or members of the faculty, and (2) involved primarily in SEH-related research activities.
Alternative #1	Doctorate-holding non-instructional researcher: This includes staff on your campus who are primarily engaged in research, applied research, or development; hold a doctorate or equivalent; and are not on the tenure track.
Alternative #2	Researcher: Professionals engaged in the conception or creation of new knowledge, products, processes, methods, and systems and in the management of the projects concerned. Include R&D managers in this category. R&D technicians and R&D support staff are to be reported separately. For context, two additional staff categories should not be reported as researcher. Research and development technicians: Persons whose main tasks require technical knowledge and experience in one or more fields of science or engineering but who contribute to R&D by performing technical tasks such as computer programming, data analysis, ensuring accurate testing, operating lab equipment, and preparing and processing samples under the supervision of researchers. Research and development support staff: Not directly involved with the conduct of a research project but who support the researchers and technicians. These employees might include clerical staff, financial and personnel administrators, report writers, patent agents, safety trainers, equipment specialists, and other related employees.

GSS = Graduate Students and Postdoctorates in Science and Engineering; SEH = science, engineering, and health.

3. Site Visit Results

The discussion of results and reactions to the theme route questions are presented at a general level with a focus on the consensus or common themes that emerged from each conversation. Where appropriate, the responses of the 11 institutions that report NFRs to the GSS (reporters), and the 4 institutions that do not report NFR to the GSS (nonreporters) are contrasted in the summaries.

3.1 Does your institution have a formal definition for “nonfaculty researcher”?

No institution included in the site visits had a formal job category of *doctorate-holding nonfaculty researcher* on their campuses. Instead, institutions that report NFRs interpret the guidance from the GSS and identify these staff through reference to job titles and job requirements available in their institutional databases. All stakeholders in the reporter group shared the methods that their institutions use to identify and report eligible employees with data that are available. The processes through which respondents crosswalk an institution’s data to identify eligible NFRs are time-consuming tasks that require staff to apply their own interpretations of the GSS definition of NFR to judge whether a researcher qualifies as an NFR.

3.2 How are research staff classified at your institution? Who constitutes your core research staff?

Stakeholders responded to these questions in one of two ways: (1) the institution has too many relevant job titles with too much variation to determine a single core of research staff, or (2) research faculty without tenure status represent the core research staff. Institutions that report too many job titles to determine a single core of research staff include Auburn University, Boston College, Johns Hopkins University, and New Mexico State University. Institutions that identified research faculty as representing core researchers include those that do report NFRs—like North Carolina State University and the University of Tennessee Health Science Center—and those that do not report NFRs to the GSS, like Georgia State University and the University of Kentucky. When comparing reporters to nonreporters, it became clear that the differentiator between the two groups was the ways in which institutional representatives interpret the GSS definition of NFR.

3.3 How does your institution report NFRs to the GSS?

Each site visit included a conversation regarding the process by which stakeholders approach NFR data collection and reporting. Participants noted various methods to achieve the task.

Coordinators at Auburn University, Boston College, North Carolina State University, and the University of Tennessee Health Science Center manually review employees to identify eligible NFRs and begin the process with a list of staff pulled from their institution’s human resources (HR) database. The potentially eligible staff are typically identified by job title, typically by querying those titles that include “research.” The manual review then focuses on whether staff with particular job titles have faculty status and had earned a doctorate degree, thus meeting the requirements of doctorate-holding nonfaculty researcher. Our conversations pressed those coordinators whose institutional databases did not include reliable information on educational attainment data. In these cases, manual review included direct outreach to employees, a review of professional profiles online, and conversations with departmental representatives that might have additional insights into the specific job duties of the researchers in question. There is considerable subjectivity involved in this process; and, in addition to being time-intensive, respondents often stated that the process does not yield systematic results across the institution.

Institutional coordinators at New Mexico State University, University of Alaska, and University of California, Davis, elect not to conduct the manual review outlined above, instead choosing to report all staff in research job titles that require a doctoral degree. These respondents acknowledged that their institutions may be reporting research staff with faculty status as NFRs. The reasoning was that some data on NFRs were preferable to reporting no data on such research staff to the GSS. The general argument from these institutions is that they felt they were complying with the spirit of the law rather than the letter of the law. Other coordinators noted that all doctorate-holding researcher staff have faculty status, and are thus ineligible, reporting zero NFRs as a result. Coordinators from Northwestern University and University of Virginia indicated that a manual review process was prohibitively time-consuming, so they pulled data that were readily available and only reviewed those data if GSS staff posed questions.

Johns Hopkins University and Texas A&M University elect not to report data on NFRs to the GSS from a central office. Instead, the coordinators outsourced the responsibility to unit respondents who complete portions of the survey relevant to their department on behalf of the institution. Both institutions consult over 50 such departmental representatives, each of whom determines whether their staff qualify as NFRs. This is responsible for a disconnect between institutional policy and reported GSS data at one of the site visit participants. In this case, the institutional coordinator said in our site visit conversations that the institution had no eligible staff because all doctorate-holding researchers have faculty status. This was not conveyed in the survey because despite the institutional

policy relayed during the visit that the institution has no eligible NFRs, unit respondents who complete portions of the survey on behalf of this institution report more than 60 NFRs annually.

The four institutions that do not report NFRs each articulated one of two barriers to their institution's capacity to provide data on NFRs to the GSS. Either the data on degree attainment are not available (as at Georgia State University and Massachusetts Institute of Technology), or all eligible researchers have faculty status (as at San Diego State University and University of Kentucky). No additional steps were taken to tease out eligible staff (e.g., considering job title as a proxy for degree attainment).

Examination of the methods stakeholders use to report NFRs to the GSS makes it clear that this process is often contingent on complex sets of judgments. The 15 conversations make it clear that institutions do not share a systematic approach to data collection. Instead, the current definition of NFR includes enough latitude that the same set of circumstances can cause stakeholders to make completely different decisions regarding eligibility of research staff as NFRs. For example, in cases like University of Kentucky, respondents follow the letter of the law and do not report any NFRs because all doctorate-holding research staff have faculty status. In other cases, like Northwestern University, some respondents follow the spirit of the law and report doctorate research staff with faculty status as NFRs, reasoning that the GSS is probably interested in that type of researcher because they are not primarily instructional faculty. In the case of Johns Hopkins University, institutional leaders that participated in the site visit follow the letter of the law, explaining that all eligible staff have faculty status, although some of the unit respondents tasked with completing the survey follow the spirit of the law. This creates a situation with an institution where all doctorate-holding staff that have faculty status can still report approximately 60 nonfaculty research staff to the GSS each year.

3.4 What data elements are available to report regarding research staff?

The GSS currently collects 15 data elements regarding NFRs, including sex and type of terminal degree. The site visits assessed whether institutions had the capacity to report additional data regarding research staff, with specific attention to demographic information (e.g., citizenship, race, ethnicity). The conversations also touched on the availability of job title, tenure status, and highest educational attainment data. In some cases, this set of questions led to a discussion of the institution's current HR or data reporting system. Seven different data systems were named by the nine institutions that chose to share this information. Responses to these questions did not vary by the data system utilized by the institution, suggesting that the seven systems represented in the site visits have similar capacities for data collection and reporting. What did vary was the degree to which those database capacities were fully utilized by institutions.

All site visit participants noted the ability to report demographic information such as citizenship, race, ethnicity, and job titles. Most were able to report field of research activity, although the origins of these data might need to be derived based on descriptions of the organizational unit in which the researcher was employed rather than something specifically coded in administrative data systems. Representatives from three institutions registered some hesitation regarding their ability to accurately report tenure status—Massachusetts Institute of Technology, Northwestern University, and University of Tennessee Health Science Center. Similarly, only Texas A&M University and University of Virginia noted the capacity to report field of doctoral study.

Regarding data on educational attainment of research staff, 13 institutions confirmed their database's ability to collect educational attainment data, but many also explained that the practice of actually recording this data element was inconsistent to the point that they expressed concern in its accuracy. Several examples were offered to explain these inconsistencies. For example, at University of California, Davis, among other institutions, only staff who had earned their doctorate at the institution where they were employed had reliable degree information. At other institutions, like Boston College and University of Alaska, degree attainment was stored in secondary, shadow, systems maintained by the hiring departments but were not centrally available. Georgia State University and Massachusetts Institute of Technology cited the fact that their institutional systems do not reliably store educational attainment as the motivation to not report NFRs to the GSS. Their inability to verify the staff as doctorate holding was their barrier to providing these data. Several institutions also noted that while degree data are collected at the time an employee is hired, that information is never updated. In such cases, an individual hired as they complete the requirements for a doctorate would not have their highest degree accurately reflected in the institutional database.

The site visits made it clear that job title and, to a lesser extent, tenure status were both more reliable administrative data elements than educational attainment. Job titles were included in this discussion of available data because as the ECDS work demonstrated, they have the capacity to serve as a proxy to identify eligible research staff. Eight GSS coordinators already apply job titles to their data collection steps, querying databases to populate a list of potentially eligible research staff. All 15 institutions reported the capacity to

access and report the job title; as addressed above, however, several representatives noted that their institutions have a multitude of different research job titles. Similarly, tenure status is an important factor for HR databases to maintain. Nearly all 15 institutions noted the ability to identify those staff who are on tenure track and those who are not.

3.5 Reactions to alternative researcher definitions

3.5.1 Doctorate-holding non-instructional researcher

Alternative Definition #1: Doctorate-holding non-instructional researcher: Staff on your campus who are primarily engaged in research, applied research, or development; hold a doctorate or equivalent; and are not on the tenure track.

This alternate definition was favorably received by many participants, and there was general consensus among stakeholders that *doctorate-holding non-instructional researcher* could be an acceptable substitute for NFR from a practical standpoint and from a data validity standpoint. Stakeholders were relieved that this alternative no longer excluded research staff with faculty status because that had been a pain point expressed in a majority of the site visits. This adjustment alone would allow Georgia State University and University of Kentucky, two of the four nonreporters included in site visits, to report this category of research staff to the GSS.

The alternate definition offered several sources of relief. Representatives who perform manual data reviews noted that this would make it easier to query their institution's database and would require that they consider one less factor when determining eligibility. Respondents generally indicated that their data systems could distinguish between faculty on the tenure track and faculty not on the tenure track.

The need for additional guidance was a common theme in the discussion of this proposed definition. Many stakeholders noted that the definition could only be a successful alternate if additional clarification and examples of eligible staff were included with the survey instructions. Further discussion allowed respondents the opportunity to describe the types of information or additional details that would be relevant. Specific inquiries about the definition are listed in bold below, including paraphrases of the questions and concerns that were raised:

- Doctorate or equivalent
  - Can workplace experience qualify as equivalent?
    - Paraphrased from comments made by Massachusetts Institute of Technology and New Mexico State University.
  - Are all terminal degrees eligible (e.g., MFA)?
    - Paraphrased from comments made by San Diego State University.
- Non-instructional researcher
  - Researchers who manage laboratory spaces are constantly teaching students and postdocs...so, would research staff who manage labs be ineligible?
    - Paraphrased from comments made by Johns Hopkins University.
- Primarily engaged in research
  - This was too open-ended for several stakeholders, who requested that the GSS offer explicit guidance—should they judge on the proportion of workload responsibilities, the source of

salaries, or some other factor?

- Paraphrased from comments by Boston College and San Diego State University.

- Staff
  - Stakeholders asked whether visiting scholars, honorary appointments, and/or doctoral dissertation committee members from outside the institution would qualify as staff.
    - Paraphrased from comments by Auburn University, New Mexico State University, and North Carolina State University.

Even with additional guidance, the use of doctorate-holding non-instructional researcher does not solve all issues that respondents encounter when collecting and reporting NFR data. Both those who report and those who do not report NFRs noted that "doctorate-holding" remained a barrier to data reporting. Those institutions that report NFR data made it clear that degree attainment status would need to be confirmed for all staff deemed newly eligible under the revised definition; this was particularly true among representatives from University of Virginia. Georgia State University and Massachusetts Institute of Technology, the two nonreporting institutions that do not have reliable degree attainment information, stated they would still be uncomfortable reporting data on this category of researcher. That said, these representatives also noted that if the definition was adopted, then they would investigate whether job titles or tenure status could serve as a proxy for educational attainment.

3.5.2 Reactions to the HERD definition of "researcher" and "research and development staff" categories

Alternative Definition #2: Researcher: Professionals engaged in the conception or creation of new knowledge, products, processes, methods, and systems and in the management of the projects concerned. Include R&D managers in this category.

The conversation began with the researcher category before extending to a discussion of the other types of research staff that the HERD collects. Discussion of the other two groups was intended to help clarify which staff qualify as a researcher (as opposed to R&D technician or R&D support staff) and would be reported as part of this alternative definition. The HERD definitions generated more mixed responses among the institutions—with one representative from University of California, Davis, initially stating "it is not possible" before relenting to say the definitions would be very difficult to implement. That reaction is indicative of a consensus that these definitions would be more challenging to accommodate. Coordinators did appreciate that more staff qualify for inclusion with this definition, but found the delineation between research staff types that HERD includes (other categories described in the OECD *Frascati Manual*[8]) difficult to operationalize on their campuses.

An operational definition is further complicated by the expectation that counts of researchers be reported with the attribute of highest degree obtained. With this attribute, GSS would be able to distinguish among various types of researcher—those holding a doctoral degree, and those with bachelor's or master's training. This is particularly relevant since the HERD definition of researchers is agnostic regarding educational attainment, while GSS has historically focused only on doctorate-holding research staff.

Stakeholders, including those at Georgia State University, did appreciate that the HERD definition is broader and more inclusive, which would allow them to report more research staff than is currently possible. However, the representatives did note that the increase in numbers would also decrease the clarity regarding the type of researcher being reported, as expressed by representatives from University of Kentucky. Specific comments about the definition are listed in bold below, including paraphrases of the questions and concerns that were raised:

- Professional
  - This term is too ambiguous and would lead to confusion on which staff to report or to exclude.

Also suggests that additional collaboration with HR and other campus stakeholders would be necessary.

    - Paraphrased from comments by North Carolina State University.

- Three categories of research staff
  - The multiple categories of research staff presented here are not helpful because the ineligible categories have plenty of "gray area" and accurate delineation of researcher is "too fuzzy."
    - Representatives from University of Virginia were responsible for "gray area," while stakeholders from University of Alaska deemed the categories "too fuzzy."
- Highest degree information
  - Fear the current issues in confirming doctoral degree status would be exacerbated if highest degree information was required because detail on nondoctoral degrees is not reliably available in institutional databases.
  - Instead, result stakeholders were concerned that many would simply report "highest degree unknown."
    - These concerns were paraphrased from conversations with representatives of Massachusetts Institute of Technology as well as University of Tennessee Health Science Center.

The overall consensus was that this definition would neither decrease reporting workload nor would it increase data reliability. Instead, the delineation of researcher among research and development technicians and research and development support staff would require additional steps for data collection while leaving more ambiguous loose ends up for individual interpretation.

3.5.3 Does your institution use GSS data? If so, in what ways? How can the utility of GSS data be improved?

The site visit discussions culminated with a question about the degree to which institutions use GSS data. Eight institutions report that they do not use the data, though seven of those expressed interest in the data moving forward. Among the seven institutions that do use the data, a majority indicated that these data have a role in their institutional benchmarking processes.

The most salient and popular suggestion to improve GSS data was to improve the website that hosts data. RTI International and NSF responded to this suggestion with some details about the revamped NCSES Interactive Data Tool. A link to this tool as well as its accompanying tutorial was shared with all site visit participants to accommodate this suggestion as well as to serve those institutions that had expressed interest in using GSS data.

4. Takeaways from the Site Visits

- Many institutions consider NFR reporting to be a complicated process that requires considerable judgment calls on what types of positions and job titles qualify as NFRs. These decisions are not always made centrally but rather are made at the organizational unit level.

- Some coordinators strictly follow the definitions provided by GSS instructions regarding faculty status as a disqualifier, while others include research faculty job titles as within the spirit of the rule when it comes to NFR reporting.

- Regardless of NFR definitions, most institutions maintain additional demographic data on employees in their administrative data systems, including race, ethnicity, and citizenship.

- Administrative data on highest degree obtained is far more variable among institutions—while some consistently maintain this information, others maintain this information in secondary systems or only maintain it for certain types of employees (such as faculty).

- Job title and tenure status were widely available and could be used as proxy data sources to help stakeholders identify eligible NFRs. There was consensus that administrative data systems could distinguish between faculty on the tenure track and those who were not.

- Among the alternative definitions of researchers discussed, there was more consensus around doctorate holding non-instructional researcher (alternative 1) than the second alternative. However, either definition will require additional context and examples for coordinators to appropriately report data.

5. Next Steps

This study has demonstrated that there is both a need to improve data collected on research staff at GSS institutions as well as to develop an appetite among the GSS respondents that participated in site visits to contribute toward those necessary adjustments. The findings suggest that a more inclusive definition of research staff is possible and practicable—specifically, a definition that includes space for doctorate-holding researchers who may be faculty but are not on the tenure track. While several institutions identified educational attainment data as a stumbling block, further conversations revealed that additional data elements could represent reliable proxies for these data. Specifically, job titles, associated minimum educational qualifications, and tenure status are data elements that could be used to identify eligible research staff. Instructions regarding the reporting of eligible research staff to GSS would benefit from guidance that references job titles and tenure status, both variables that institutional representatives can seemingly access with limited difficulty. Expanding the data collection to include demographics such as race, ethnicity, and citizenship would increase the utility of NFR data.

Stakeholders were optimistic about the feasibility of collecting doctorate-holding non-instructional researchers as an alternative definition for research staff. This optimism was paired with requests for additional clarification and concrete examples of job titles that qualify. Regardless, the definition received a more positive reception than did the current GSS definition or the current definition employed by the HERD Survey. The lack of educational attainment data may lead to some underreporting at specific institutions, although this definition would largely avoid the systemic issues caused by the variability in the use of the term “faculty” to describe similar types of staff performing research at different institutions. As an additional note, given the relatively cool reception and measurement issues raised by site visit participants to the HERD researcher definitions, it would be interesting to compare counts obtained by recent HERD Surveys to IPEDS data on researcher personnel and NFR counts obtained by GSS.

We recommend as a next step a record-keeping survey distributed to all institutions in the GSS universe. This would gather additional data from those institutions that currently report research staff to the GSS, either postdocs or NFRs, and include those that have research staff who are currently not reporting them. The survey would explore data and reporting capabilities on research staff and would obtain additional feedback on the alternative definitions for researchers investigated in the site visits. If a decision to revise NFR definitions is warranted, results from the record-keeping survey could then be used to inform a pilot study to further assess the feasibility of additional data elements and the usability of a new definition for research staff.

Appendix A: Results of Quartile, Correlation, and Regression Analyses

A.1 Results of quartile analysis

Quartile analysis was conducted on the 2017 data.

Table A-1. Key to interpret the color codes applied to tables

Color	Meaning
Red text	Highest value in both row and column
Green text	Highest value in column
Blue text	Highest value in row
Grey cell	Cell where overlap is expected

Table A-2. Key to the quartile values applied in the tables

Quartile	NFRs reported to GSS	NFRs and PDs reported to GSS	All research staff reported in IPEDS	Research staff without faculty status reported in IPEDS	Research expenditures
0	0	0	0	0	0
1	1–8	1–10	1–8	1–5	1–1,416,000
2	9–29	12–52	9–38	6–31	1,416,001–8,274,000
3	30–116	53–268	39–238	32–218	8,274,001–74,062,000
4	> 116	269–6,855	239–4,790	> 218	> 74,062,000

GSS = Graduate Students and Postdoctorates in Science and Engineering; IPEDS = Integrated Postsecondary Education Data System; NFR = doctorate-holding nonfaculty researcher; PD = postdoctoral researcher appointee.

Table A-3. NFRs reported to GSS as compared with research staff without faculty status reported in IPEDS: 2017

Staff		Research staff without faculty status reported in IPEDS					
NFRs reported to GSS	Quartile	0	1	2	3	4	Total
	0	358	44	36	12	9	459
	1	36	8	11	8	2	65
	2	33	5	6	18	0	62
	3	30	2	2	18	10	63
	4	21	0	2	3	36	61
	Total	478	59	57	59	57	710

GSS = Graduate Students and Postdoctorates in Science and Engineering; IPEDS = Integrated Postsecondary Education Data System; NFR = doctorate-holding nonfaculty researcher.

Table A-4. NFRs and postdocs reported to GSS as compared with all Research staff reported in IPEDS: 2017

Staff		All research staff reported in IPEDS					
NFRs and PDs reported to GSS	Quartile	0	1	2	3	4	Total
	0	280	61	27	4	1	373
	1	31	22	27	6	0	86
	2	14	8	26	33	2	83
	3	12	3	9	37	23	84
	4	8	1	1	9	65	84
	Total	345	95	90	89	91	710

GSS = Graduate Students and Postdoctorates in Science and Engineering; IPEDS = Integrated Postsecondary Education Data System; NFR = doctorate-holding nonfaculty researcher; PD = postdoctoral researcher appointee.

Table A-5. NFRs reported to GSS as compared with research expenditures reported to HERD and IPEDS: 2017

Staff		Research expenditures					
NFRs reported to GSS	Quartile	0	1	2	3	4	Total
	0	57	156	141	77	28	459
	1	0	5	19	33	8	65
	2	0	2	3	42	15	62
	3	0	0	0	11	52	63
	4	2	0	0	0	59	61
	Total	59	163	163	163	162	710

GSS = Graduate Students and Postdoctorates in Science and Engineering; HERD = Higher Education Research and Development Survey; IPEDS = Integrated Postsecondary Education Data System; NFR = doctorate-holding nonfaculty researcher.

Table A-6. NFRs and postdocs reported to GSS as compared with research expenditures reported to HERD and IPEDS: 2017

Staff		Research expenditures					
NFRs and PD reported to GSS	Quartile	0	1	2	3	4	Total
	0	57	150	119	44	3	373
	1	0	10	40	34	2	86
	2	0	2	4	68	9	83
	3	0	1	0	17	66	84
	4	2	0	0	0	82	84
	Total	59	163	163	163	162	710

GSS = Graduate Students and Postdoctorates in Science and Engineering; HERD = Higher Education Research and Development Survey; IPEDS = Integrated Postsecondary Education Data System; NFR = doctorate-holding nonfaculty researcher

A.2 Select results of correlation analyses

Table A-7. Select results of correlation analyses: 2017–19

Variable	NFRs reported to GSS, 2017	NFRs reported to GSS, 2018	NFRs reported to GSS, 2019
HERD/IPEDS research and development expenditures	0.71	0.69	0.30
IPEDS research staff with tenure or on tenure track	0.05	0.08	0.06
IPEDS research staff without tenure or not on tenure track	0.25	0.34	0.32
IPEDS research staff without faculty stature	0.62	0.52	0.57
Status as public institution	0.03	0.05	0.06
Not very high research institution; does not grant medical doctorates	-0.43	-0.44	-0.43
Very high research institution; does not grant medical doctorates	0.12	0.18	0.18
Not very high research institution; does grant medical doctorates	-0.05	-0.06	-0.07
Very high research institution; does grant medical doctorates	0.59	0.59	0.59

GSS = Graduate Students and Postdoctorates in Science and Engineering; HERD = Higher Education Research and Development Survey; IPEDS = Integrated Postsecondary Education Data System; NFR = doctorate-holding nonfaculty researcher.

A.3 Select results of regression analyses

Table A-8. Select results of regression analyses: 2017–19

Year	Number of observations read	Degrees of freedom	Sum of squares	Mean square	*F*-value	R^2 value
2017	710	8	5,340,911	667,614	120.88	0.58
2018	718	8	4,453,787	556,723	97.68	0.52
2019	719	8	4,984,023	623,003	106.28	0.55

Table A-9. *P*-values of variables in the regression analyses: 2017–19

Variable	*P*-value, 2017	*P*-value, 2018	*P*-value, 2019
Intercept	0.9144	0.9951	0.8281
HERD/IPEDS research and development expenditures	< 0.0001	< 0.0001	< 0.0001
IPEDS research staff with tenure or on tenure track	0.1346	0.4239	0.1343
IPEDS research staff without tenure or not on tenure track	0.0661	0.4862	0.6772
IPEDS research staff without faculty stature	< 0.0001	< 0.0001	< 0.0001
Status as public institution	0.6711	0.5796	0.2651
Very high research institution; does not grant medical doctorates	0.0681	< 0.0001	0.0001
Not very high research institution; does grant medical doctorates	0.9738	0.8018	0.5832
Very high research institution; does grant medical doctorates	< 0.0001	< 0.0001	< 0.0001

HERD = Higher Education Research and Development Survey; IPEDS = Integrated Postsecondary Education Data System.

Appendix B: Theme Route for Virtual Site Visits

B.1 Prologue

Doctorate-holding nonfaculty researchers (NFRs) have been collected for a long time—we are attempting to determine how we can improve the value of these data. We are engaging in a series of conversations to understand how you approach reporting these data, what challenges you face in providing NFR data, and how these data might of more use to you.

We will be taking notes during our meeting to help us recall key points—nothing you say will be directly attributed to you or your institution in any published materials stemming from this meeting.

B.1.1 Current definition used by the Survey Graduate Students and Postdoctorates in Science and Engineering (GSS)

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

B.2 First set of questions: Definitional questions

- How are research staff classified and counted at your institution? Do research staff overlap with faculty? If so, In what ways?
- Does your institution have a formal definition for “nonfaculty researcher”?
- How is the term “faculty” defined on your campus?

B.2.1 For both types of reporters

- Please tell us about the researchers (individuals primarily engaged in research, applied research and development) on your campus.
  - Who are your core research staff?
    - What are the major centers in which they work?
  - What data do you hold regarding research staff?
    - Demographics?
    - Job titles?
    - Educational attainment (e.g., highest degree obtained)?
    - Field of doctorate?
    - Tenure status?
    - Instructional responsibilities?

- PROBE: Job title as proxy for educational attainment, tenure status, or instructional responsibilities.

B.3 Second set of questions: Current reporting practices and capabilities

B.3.1 For those reporting NFRs

How do you and your colleagues obtain the NFR data you report to the GSS?

- How does your institution identify NFRs?
- What institutional databases need to be consulted?
- What other staff/offices are required to obtain this information?
  - Do you formally or informally consult specific academic or research units (i.e., *unit respondents*) for this data?

B.3.2 Field of study/organizational unit

- For those reporting NFRs: How do you associate reported NFRs with disciplinary fields?
  - Does this method differ from how you associate postdocs with disciplinary fields?

B.3.3 For those <u>not</u> reporting NFRs

Given the current definition of NFR, are there staff at your institution that fit this description?

- If yes: What are the challenges you encounter in being able to report these researchers?
  - PROBE: Database issues, different offices involved in accessing data, data not actually collected.
- If no: What job titles are common among those staff at your institution whose primary responsibility is to conduct research

B.4 Third set of questions: Alternative NFR definition

NCSES is considering changes to how NFRs are defined. What are your thoughts on the following definition?

B.4.1 Doctorate-holding non-instructional researcher

This includes staff on your campus who are primarily engaged in research, applied research or development, hold a doctorate or equivalent and are not on the tenure-track.

- In what ways would this change in definition impact the source from which you obtain these data on your campus (e.g., different database, office or staff required to report these data)?
- Based on knowledge of your institution and expertise regarding academic employment, could you speak toward what you see as the advantages and problems with this definition?

B.4.2 Higher Education Research and Development Survey (HERD) definition of research staff for comparison

***Researcher*:** Professionals engaged in the conception or creation of new knowledge, products, processes, methods and systems also in the management of the projects concerned. Include research and development (R&D) managers in this category (appendix table B-1).

Table B-1. Categories of researcher and research and development staff used in the HERD Survey

Description of R&D functions		
Researchers	R&D technicians	R&D support staff
Professionals engaged in the conception or creation of new knowledge, products, processes, methods, and systems and also in the management of the projects concerned. Include R&D managers in this category.	Persons whose main tasks require technical knowledge and experience in one or more fields of science or engineering, but who contribute to R&D by performing technical tasks such as computer programming, data analysis, ensuring accurate testing, operating lab equipment, and preparing and processing samples under the supervision of researchers.	Not directly involved with the conduct of a research project but support the researchers and technicians. These employees might include clerical staff, financial and personnel administrators, report writers, patent agents, safety trainers, equipment specialists, and other related employees.
Researcher vs. R&D technician		
Researchers contribute more to the creative aspects of R&D whereas technicians provide technical support. For example, a researcher (scientist or engineer) would design an experiment and a technician would run the experiment and assist in analyzing results.		

HERD = Higher Education Research and Development Survey.

B.5 Fourth set of questions: Utilization of GSS data

- Does your institution regularly utilize reported GSS data?
  - If yes, in what ways?
  - If no, what suggestions might you have for the National Center for Science and Engineering Statistics to make the data more useful to your institution?
  - Are there any specific reports you would be interested in that involve NFR data?
- Are there any other suggestions you have for improving the GSS data collection process or for improving the utility of its reports?

Endnotes

[1].Einaudi P, Heuer R, Green P, Kang K. 2015. *Examining the Reporting of Nonfaculty Doctorate Researchers in the Survey of Graduate Students and Postdoctorates in Science and Engineering*. Working Paper NCSES 15-201. Arlington, VA: National Science Foundation, National Center for Science and Engineering Statistics. Available at https://www.nsf.gov/statistics/2015/ncses15201/.

[2].Einaudi et al. (2015).

[3].The eight categories of primary or secondary professional responsibilities available to respondents of the ECDS include basic research, applied research, development, management or administration, supervising students, teaching, clinical or professional services, and other activity.

[4].Relevance is determined by Google's rankings of search output as well as our manual review of the content published on the websites.

[5].Harvard School of Public Health. 2019. *Guidelines for Non-faculty Research Titles*. Boston, MA: Harvard T.H. Chan School of Public Health. https://cdn1.sph.harvard.edu/wp-content/uploads/sites/40/2019/04/nonfacresearchgrid-4-30-2019.pdf.

[6].DiBela S, Thomson J. 2016. UNLV Top Tier Research, Scholarship and Creative Activities: 2015–2016 Updates. *Path to Top Tier: Increasing Nonfaculty Research Staff and Postdoctoral Fellows*, p. 1. Las Vegas, Nevada: University of Nevada, Las Vegas. https://www.unlv.edu/sites/default/files/assets/toptier/plans/RSC/RSC-Policies-Processes-4-5-Post-Docs.pdf.

[7].Organisation for Economic Co-operation and Development (OECD). 2015. *Frascati Manual 2015: Guidelines for Collecting and Reporting Data on Research and Experimental Development.* The Measurement of Scientific, Technological and Innovation Activities. Paris: OECD Publishing.

[8].See OECD (2015).