

# **Examining the Reporting of Nonfaculty Doctorate Researchers in the Survey of Graduate Students and Postdoctorates in Science and Engineering**

Working Paper | NCSES 15-201 | August 2015

by Peter Einaudi, Ruth E. Heuer, Patricia J. Green, and Kelly Kang [1]

## **Disclaimer and Acknowledgment**

Working papers are intended to report exploratory results of research and analysis undertaken by the National Center for Science and Engineering Statistics (NCSES). Any opinions, findings, conclusions, or recommendations expressed in this working paper do not necessarily reflect the views of the National Science Foundation (NSF). This working paper has been released to inform interested parties of ongoing research or activities and to encourage further discussion of the topic.

The analysis for this research was conducted by RTI International on behalf of NSF's NCSES under contract number NSFDACS1063289.

## **Suggested Citation**

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 <http://www.nsf.gov/statistics/2015/ncses15201/>.

## **Contact**

Kelly Kang  
Project Officer  
Human Resources Statistics Program  
National Center for Science and Engineering Statistics  
[kkang@nsf.gov](mailto:kkang@nsf.gov)  
(703) 292-7796

## **Contents**

Executive Summary

Introduction

Analysis of the GSS Data

Validation and Imputations

Conclusions

Data Sources and Limitations

Notes

Appendix A – GSS NFR Question Change

Appendix B – GSS NFR Debriefing Survey Questions

## Executive Summary

This working paper examines the consistency of doctorate-holding nonfaculty researcher (NFR) reporting in the Survey of Graduate Students and Postdoctorates in Science and Engineering (GSS) since 2010, analyzes key reporting patterns, and attempts to validate the GSS data to better understand the overall quality of the NFR data. The accurate reporting of NFRs is of critical interest to the National Science Foundation and other stakeholders because it is the first step in understanding the use and impact of this increasingly significant part of the academic science and engineering workforce.

After growing by an annual average of 7% for the prior 2 decades, the count of NFRs reported in the GSS jumped by 28% in 2008 and by 52% in 2010. These increases were likely driven by methodological changes to the GSS focused on improving the reporting of postdoctoral appointees (postdocs) and NFRs (see <http://www.nsf.gov/statistics/infbrief/nsf13334/> for a summary of the impact of these changes on the postdoc data).

The methodological changes in 2010 included: a continued focus from 2008 to include centers and other non-degree-granting units; survey redesign to expand and separate the NFR items from the postdoc items; the designation of separate respondents for the graduate student and postdoc or NFR sections; and, a reminder to respondents of the importance of the postdoc and NFR data.

After implementing these changes, the number of schools and units reporting NFRs increased dramatically. Between 2009 and 2010, the percentage of GSS organizational units (academic departments, programs, research centers, or health care facilities) reporting NFRs increased from 19% to 26%, and the number of schools reporting NFRs increased from 38% to 48%. These large increases validated long-standing concerns that NFRs were being undercounted due to the lack of a common definition across institutions and limited access to these data among GSS respondents. Within the GSS, NFRs are defined as doctorate-holding researchers who are neither postdocs nor faculty members. The magnitude of the increases, however, raised the concern that some of the GSS institutions might have overreported their NFR counts.

From 2010 to 2012, NFR counts for most units (93%) have stabilized; 95% of all units had similar NFR counts in 2010 and 2011, and 96% had similar counts in 2011 and 2012. However, it is clear that some respondents have difficulty reporting NFRs. Of the units that reported having at least one NFR from 2010–12, 5% were unable to provide any data about them, and only half were able to provide complete responses to the NFR questions (see Appendix A for the NFR questions). Similarly, 122 of the 684 schools in the 2012 GSS reported having postdocs but no NFRs; this is fairly unlikely, and further follow-up with these schools is warranted to identify the cause of this discrepancy, the availability of relevant data at these institutions, and determine whether there are issues with the NFR definition.

In a short debriefing survey conducted following the 2012 data collection, 26% of respondents indicated that their institution had a common definition for NFRs. In addition, among the 32 responding schools that reported zero NFRs in 2012, the majority were unable to provide NFR

counts because their institutional data management systems did not have the necessary information to identify the NFRs.

Finally, GSS counts of postdocs and NFRs were compared to the Higher Education Research and Development Survey data on funding, salaries, and postdocs. This comparison showed that research expenditures tracked closely with the number of NFRs reported in the GSS, supporting the data collected in both surveys, and identified several institutions needing follow-up for the potential over- and underreporting of NFRs in the GSS.

Based on these analyses, the 2010–12 NFR counts are much more reliable and accurate in gauging the size and distribution of this population across the GSS academic institutions than prior estimates. However, NCSSES plans to continue working with those schools identified as having potential NFR data issues to improve their reporting practices and to ensure continuous quality improvement in the reporting of NFR data.

## Introduction

The Survey of Graduate Students and Postdoctorates in Science and Engineering (GSS) is an annual census of all U.S. academic institutions granting research-based master's degrees or doctorates in science, engineering, and selected health (SEH) fields as of fall of the survey year. The survey, sponsored by the NCSES of the National Science Foundation (NSF) and the National Institutes of Health, collects the total number of graduate students, postdoctoral appointees (postdocs), and doctorate-level nonfaculty researchers (NFRs) by demographic and other characteristics, such as source of financial support. Results are used to assess shifts in graduate enrollment and postdoc appointments as well as trends in financial support.

The concept of NFRs was introduced in the GSS in 1979. Before that time, counts were collected for a combined "postdocs and/or research associates" category. However, emphasis in 1979 was in collecting specific data on postdocs, and only limited data items were collected on NFRs. Also in 1979, "research associates" was changed to "nonfaculty research staff with doctorates," and data were collected separately for postdocs and NFRs. The NFR category was kept to ensure that institutions would not lump the previously reported research associates into the new postdoc category. Growth in postdoc employment led to a more extensive series of questions for postdocs than for the new NFR category. For example, sex and medical degree status were gathered for both groups, but sources of support and foreign status were collected for postdocs only. The definition of NFRs has remained relatively constant over time with two criteria: (1) NFRs are doctorate-holding staff who are neither postdocs nor faculty, and (2) NFRs are primarily involved in research.

Starting in 2008, NSF began addressing known issues with the postdoc and NFR data collections by emphasizing to respondents that non-degree-granting units, like centers or research institutes, were eligible for the GSS. In 2009, NSF conducted a Postdoc Pilot Study to determine if schools could provide more detailed postdoc and NFR data and whether having separate respondents provide data on graduate students and on postdocs and NFRs improved reporting of the postdoc and NFR data. Based on the results of this study, a new series of items about postdocs and NFRs was added to the 2010 GSS.[2] Appendix A shows the changes in the GSS NFR questions between 2009 and 2010.

In addition, the protocols for collecting data changed, and institution presidents were asked to appoint a postdoc coordinator in addition to the school coordinator responsible for reporting graduate student data. The letters to the presidents included information on counts of postdocs and NFRs reported by their school coordinators as well as information on how the data are used by external organizations, such as in the Carnegie Classification. The designation of a postdoc coordinator at institutions, the expanded set of questions, and the heightened awareness of the value and use of the items contributed to an increase in the number of postdocs and NFRs reported in the survey.

As shown in figure 1, the percentage of GSS organizational units (academic departments, programs, research centers, or health care facilities) reporting NFRs increased from 19% in 2009 to 26% in 2010, and the number of NFRs increased even more sharply from 14,059 to 21,345.

The large increase in NFRs and the sheer magnitude of this type of employment indicates the significance of NFRs to the science and engineering workforce. This working paper examines the consistency of NFR reporting since 2010, analyzes key reporting patterns, and attempts to validate the GSS data to better understand the overall quality of the NFR data. Schools and units, rather than institutions, were used to analyze the consistency of reporting to prevent masking of inconsistencies at the institution level. Some of the major concerns about the quality of NFR data that this working paper addresses include (1) inconsistent reporting from year to year, (2) the lack of a common definition for NFRs across respondents, and (3) the inability of reporting units to accurately differentiate postdocs and NFRs.

## **Analysis of the GSS Data**

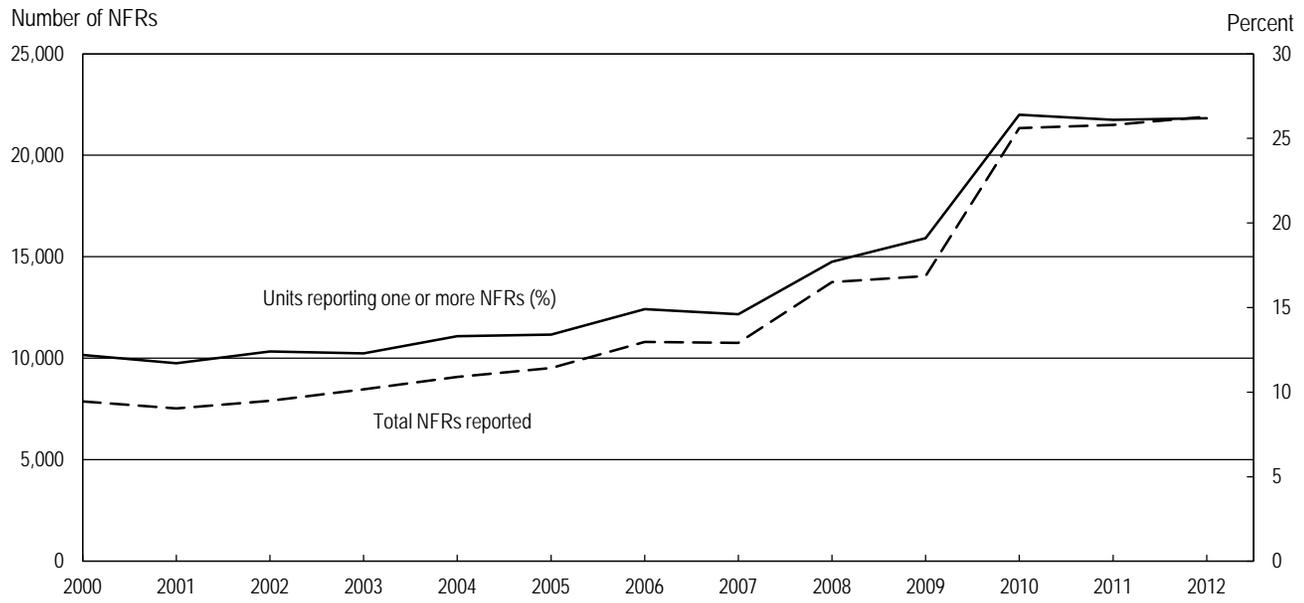
### *Consistency of NFR Reporting across Time*

Given improvements to the data collection protocol in 2010, considerable increase in NFR counts was expected between 2009 and 2010 followed by stability after 2010. Although this pattern was observed at the aggregate level, the NFR data from 2010 to 2012 showed some inconsistencies in reporting NFR counts at the unit level.

Tables 1 and 2 show the level of consistency of NFR counts across all schools and all units for three time periods: between 2010 and 2011, between 2011 and 2012, and between 2010 and 2012. At the *school* level, 91% of schools reported similar NFR counts in 2010 and 2011, 94% in 2011 and 2012, and 89% in 2010 and 2012 (table 1). A school's NFR count was considered to be similar if the difference in the number of NFRs was within one standard deviation of the mean difference observed from 2010 to 2012, or plus or minus 22 NFRs. The difference in counts columns show the aggregate change in NFR counts across all schools included in that category.

Table 2 provides similar information on the consistency of NFR reporting at the *unit* level and shows even greater stability over time: 95% of all units reported similar NFR counts in 2010 and 2011, 96% in 2011 and 2012, and 93% in 2010 and 2012. At the unit level, the NFR count was considered to be similar between 2 years if the number of NFRs was within plus or minus 3 NFRs (i.e., within one standard deviation of the mean difference observed from 2010 to 2012).

FIGURE 1. NFR counts and percentage of units reporting one or more NFRs: 2000–12



NFR = nonfaculty researcher.

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

TABLE 1. Distribution of schools, by difference in NFR counts between 2010 and 2012

School type	2010–11		2011–12		2010–12	
	Number of schools	Difference in counts	Number of schools	Difference in counts	Number of schools	Difference in counts
Total schools	698	153	698	410	698	563
Schools with 200 or more NFRs over previous year	2	479	1	224	3	708
Schools with 100 to 199 more NFRs over previous year	4	574	1	170	6	812
Schools with 50 to 99 more NFRs over previous year	6	447	7	435	10	655
Schools with 23 to 49 more NFRs over previous year	14	465	21	607	20	643
Schools with the same number or within 22 NFRs compared with previous year	638	-132	655	-136	622	-156
Schools with 23 to 49 fewer NFRs over previous year	25	-813	6	-211	20	-631
Schools with 50 to 99 fewer NFRs over previous year	6	-341	4	-319	13	-783
Schools with 100 to 199 fewer NFRs over previous year	2	-317	3	-360	2	-270
Schools with 200 or fewer NFRs over previous year	1	-209	0	0	2	-415

NFR = nonfaculty researcher.

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

TABLE 2. Distribution of units, by difference in NFR counts between 2010 and 2012

Unit type	2010–11		2011–12		2010–12	
	Number of units	Difference in counts	Number of units	Difference in counts	Number of units	Difference in counts
Total units	15,294	153	15,294	410	15,294	563
Units with 20 or more NFRs over previous year	32	1,169	13	355	35	1,185
Units with 10 to 19 more NFRs over previous year	64	827	62	832	107	1,422
Units with 4 to 9 more NFRs over previous year	264	1,455	269	1,452	413	2,332
Units with the same number or within 3 NFRs compared with previous year	14,556	-63	14,677	26	14,238	-31
Units with 4 to 9 fewer NFRs over previous year	284	-1,581	216	-1,159	376	-2,062
Units with 10 to 19 fewer NFRs over previous year	72	-924	40	-519	91	-1,212
Units with 20 or fewer NFRs over previous year	22	-730	17	-577	34	-1,071

NFR = nonfaculty researcher.

NOTE: Unit consists of departments, degree-granting programs, research centers, and health facilities at the institutions.

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

Overall, the reported counts of NFRs from year to year are reasonably stable, with a few exceptions (as noted at the extremes in these tables). In addition, the counts appear to be getting more stable as schools gain experience in collecting and reporting these counts and with the appointment of separate respondents who are better able to provide more accurate postdoc and NFR data.

Despite the relative stability of overall counts, a recurring problem for some respondents has been the inability to gather and report all requested NFR data. The school coordinator or unit respondent may know that the unit has NFRs but may be unable to provide any information (i.e., all data are unknown) or may be able to provide only partial information (e.g., total counts but not counts by sex or doctoral degree type). Table 3 provides a unit-level analysis of the extent of missing data within the NFR items from 2010 to 2012.

Including units that reported having zero NFRs for three years from 2010 to 2012, 78% of the units in the 2010–12 GSS were able to provide complete NFR data across all years; however, only 47% of the 6,478 units that reported having NFRs at least once from 2010 to 2012 were able to provide complete NFR data across all years. An additional 18% of units with NFRs were able to provide total counts in all three years but were unable to provide one or more detailed counts from 2010 to 2012, and 29% were missing the total count in at least 1 year. Finally, 5% (346 units) indicated that they had NFRs but were unable to provide any data from 2010 to 2012.

#### *Comparison of Postdoc and NFR Counts*

Table 4 provides a comparison of postdoc and NFR reporting at the school level from 2007 to 2012. Despite a reduction in the overall number of GSS eligible schools, the number and proportion of schools reporting NFRs increased substantially as expected. Although there was only a relatively modest increase in schools reporting NFRs in 2008, there was a much larger spike in 2010, coinciding with the expansion of the GSS postdoc and NFR items and the addition of the postdoc coordinators at many schools. Although the number of schools reporting NFRs declined in 2011, the 2011 and 2012 reporting levels remain substantially higher than those in 2009 and prior years.

Some respondents conveyed their difficulty in distinguishing differences between postdocs and NFRs. This difficulty resulted in schools not reporting counts for one or the other position or in their reporting the same individuals in both counts. Although the majority of these issues were usually resolved during data collection (see Data Review discussion below), some could not be resolved due to an inability to gather the requested data. As shown in table 4, more than 10 schools in each of the past 3 years did not report any counts on postdocs or NFRs, but indicated that they had postdocs or NFRs. (A check box to indicate the unit had postdocs and NFRs but could not report any data about them was added in 2010.) Similarly, more than 100 schools (15%–19% of schools) reported postdocs but no NFRs in each of the past 6 years. In 2012, 27 of these schools reported more than 100 postdocs, and 15,560 postdocs were reported in the 122 schools with postdocs and no NFRs. These schools are likely to have NFRs but are either not reporting them or are including them in their postdoc counts.

TABLE 3. Unit reporting status of NFR data: 2010–12

Reporting status	Including units with no NFRs		Excluding units with no NFRs	
	Number	Percent	Number	Percent
All units	15,294	100.0	6,478	100.0
NFR data complete in all years <sup>a</sup>	11,858	77.5	3,042	47.0
NFR totals provided all years, but some data missing in 1 or more years	1,182	7.7	1,182	18.2
NFR total count missing in 1 or more years	1,908	12.5	1,908	29.5
NFR total count missing in all years	346	2.3	346	5.3

NFR = nonfaculty researcher.

<sup>a</sup> Includes units reporting valid zero NFRs.

NOTE: Unit consists of departments, degree-granting programs, research centers, and health facilities at the institutions.

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

TABLE 4. School reporting status of postdocs and NFRs: 2007–12

School reporting status	Number of schools						Reported 2012 counts	
	2007	2008	2009	2010	2011	2012	NFRs	Postdocs
Total	700	708	701	690	686	684	19,639	60,452
Reported NFR and postdoc counts	240	258	255	296	279	281	19,484	44,892
Number of postdocs > number of NFRs	212	217	217	236	225	231	14,950	42,031
Number of postdocs = number of NFRs	6	24	13	17	6	11	211	211
Number of postdocs < number of NFRs	22	17	25	43	48	39	4,323	2,650
Reported NFR counts only	7	6	3	19	7	6	155	0
Reported postdoc counts only	122	113	134	107	121	122	0	15,560
Reported that they had postdocs and/or NFRs, but could not provide counts	-	-	-	16	15	12	0	0
Reported no postdocs and no NFRs	315	323	305	247	257	261	0	0
Nonresponding school	16	8	4	5	7	2	0	0
Percent reporting NFRs	35.3	37.3	36.8	45.7	41.7	42.0	na	na

- = data not collected prior to 2010; na = not applicable.

NFR = nonfaculty researcher.

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

The opposite—reporting NFRs and no postdocs—is also potentially problematic, but the associated error would be quite limited. Only 6 schools reported having NFRs and no postdocs in 2012, and only 155 NFRs were reported from these schools. Schools reporting identical counts of NFRs and postdocs are also of concern, but inaccuracies are similarly limited by the total numbers of NFRs and postdocs in these schools. Of the 11 schools reporting equal numbers of each position in 2012, 8 schools have very low counts (6 or fewer) and the remaining 3 schools reporting equal numbers of postdocs and NFRs of 25 or more.

Lastly, there were 39 schools reporting more NFRs than postdocs in 2012. Although some schools might legitimately have more NFRs than postdocs, this is the exception rather than the rule. All these schools will be followed up with in the next GSS cycle to verify the postdoc and NFR counts.

## **Validation and Imputations**

### *Debriefing Survey*

To better understand the large increase in NFR counts from 2009 to 2010 and to confirm the reliability of the NFR data, the NFR Debriefing Survey was conducted following the 2012 data collection to obtain information about how schools define and identify NFRs. This Web survey was e-mailed to a purposive sample of schools based on school's reporting of NFR data. A school was selected if: NFR count increased by 15% or more between 2009 and 2010; NFR count is higher than postdoc count in 2012; or they reported no NFRs in 2012 (see Appendix B). The NFR Debriefing Survey was sent to 103 GSS schools, and 91 schools responded.

Schools indicated that they had definitional and other data issues (e.g., school or postdoc coordinator did not have access to these data, NFRs are not classified uniformly, school considered NFRs and postdocs to be faculty, school combined postdocs and NFRs for reporting). An analysis of the responses found the following:

- Of 91 responding schools, only 24 schools reported that they had a common definition for NFRs in their programs.
- Thirty-nine schools reported that they considered "postdoctoral fellow or researcher" to be also an applicable job title for a doctorate-holding NFR.
- Of the 30 schools that reported a larger count of NFRs than postdocs in 2012, only 1 school reported double-counting some individuals as both postdocs and NFRs, which suggests that the vast majority of schools were able to separate the reporting of postdocs from other NFRs as instructed.
- For schools that reported a larger number of NFRs in 2010 than in 2009, the most common reason cited was that the 2010 data were more accurate. For example, some schools indicated that they did not provide NFR counts in 2009 or made changes to the NFR definition or GSS data reporting procedures (e.g., switching from individual unit respondents to reporting from a central data source). Other schools cited increased

research funding in 2010, which led to more hiring of NFRs. However, several schools were unsure of the reasons for the change in their NFR counts from 2009 to 2010.

- Of the 32 schools that reported zero NFRs in 2012, 19 indicated it was likely that their school had NFRs but was unable to provide counts of these individuals, and 3 indicated that they did not know if their school had NFRs. The typical reason was the schools' systems not having the information needed to identify the NFR population properly (e.g., type of degree (master's or doctorate), descriptive job titles).

Based on NFR Debriefing Survey responses, the schools that had specific reporting issues will be contacted to determine whether they will be able to report the NFR data correctly in the future.

### *Quality Measures*

To assess the quality of the GSS NFR data further, the institution-level NFR counts were compared with funding and personnel data from the NSF/NCSES's Higher Education Research and Development Survey (HERD). The HERD is an annual census of U.S. colleges and universities with at least \$150,000 in research and development (R&D) expenditures during the prior fiscal year. It collects information on the R&D expenditures by field of research as well as counts of R&D personnel supported by those funds. Given that the majority of R&D funds in academia are expended in science, engineering, and selected health (SEH) fields, substantial overlap exists between the HERD and GSS-eligible institutions. Therefore, the number of NFRs reported in the GSS is expected to be highly correlated with the level of R&D expenditures and counts of R&D personnel in the HERD at a given institution.

Table 5 presents the results of the institution-level match for the 2012 GSS and 2012 HERD data. A total of 454 institutions were in both the GSS and the HERD, 111 institutions were in the GSS only, and 182 institutions were in the HERD only. Institutions that are in the GSS only had less than \$1,000,000 in R&D expenditures (i.e., not eligible for the full version of the HERD that collects data on R&D personnel and postdocs). Although the nonmatching institutions represent potential undercoverage for the surveys, the magnitude of this undercoverage is small, as 99% of R&D expenditures reported in the HERD were reported in institutions found in both the GSS and the HERD. Similarly, 99% of all postdocs and NFRs reported in the GSS were in the institutions found in both the GSS and the HERD.

Research expenditures are expected to track closely with the number of NFRs in the GSS; as an institution's funding for R&D increases, the number of NFRs in the GSS is expected to increase. A plot of research expenditures (from the HERD) and number of NFRs (from the GSS) for each institution in both the HERD and GSS confirms this expectation (figure 2). An ordinary least squares regression analysis found that 61% of the variance in NFR counts is explained by the amount of research funding; the correlation between the number of NFRs and funding is .78 ( $p < .0001$ ). These findings provide substantial support for the quality of GSS NFR data.

TABLE 5. Results of GSS-HERD data match: 2012

GSS-HERD match	In both GSS and HERD	In GSS only <sup>a</sup>	In HERD only
Number of institutions	454	111	182
Total R&D funds (HERD Q1) (US\$)	64,456,987	NA	1,317,537
Salaries, wages, benefits (HERD Q13a) (US\$)	27,822,699	NA	518,325
HERD R&D personnel (HERD Q16)	890,117	NA	26,705
HERD postdocs (HERD Q17)	65,213	NA	1,459
GSS postdocs	62,377	474	261
GSS NFRs	21,773	135	51

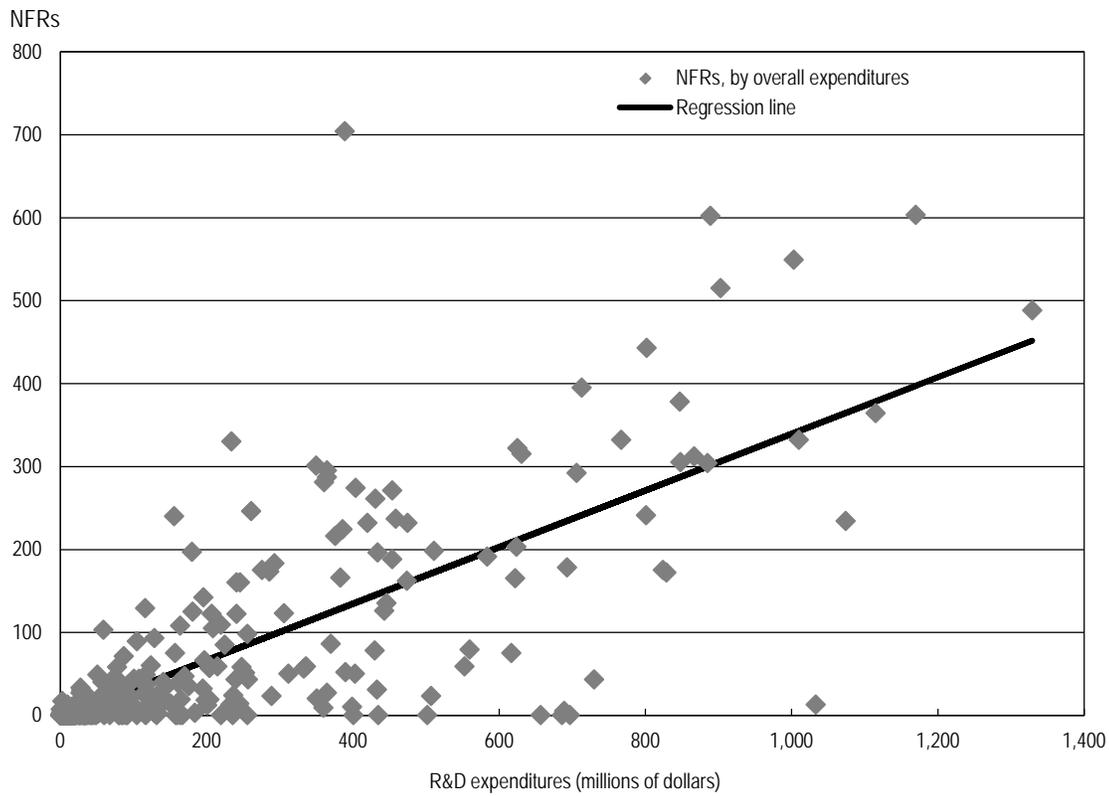
NA = not available.

GSS = Survey of Graduate Students and Postdoctorates in Science and Engineering; HERD = Higher Education Research and Development Survey; NFR = nonfaculty researcher.

<sup>a</sup> In 2011, the GSS began collecting data from potentially eligible new frame institutions from the Integrated Postsecondary Education Data System and other data sources. These new frame data are not included in the published GSS data while the final eligibility determination is made. These institutions are included for this comparison with the HERD data to maximize the institutional match.

SOURCES: National Science Foundation, National Center for Science and Engineering Statistics, Higher Education Research and Development Survey, and Survey of Graduate Students and Postdoctorates in Science and Engineering.

FIGURE 2. NFRs per institution, by overall R&D expenditures: 2012



NFR = nonfaculty researcher.

NOTE: The figure excludes two outlier institutions; one had an unusually high number of NFRs and the other had unusually high R&D expenditures.

SOURCES: National Science Foundation, National Center for Science and Engineering Statistics, Higher Education Research and Development Survey, and Survey of Graduate Students and Postdoctorates in Science and Engineering.

Potential NFR reporting problems could be identified by an institution with a large R&D expenditures (in the HERD data) but few or no NFRs (in the GSS data) or, conversely, small R&D expenditures and a large number of NFRs. Further follow-up will be made in such cases in the next GSS cycle.

#### *Data Review Results and Imputation of Missing NFR Data*

Table 6 summarizes the results of the post-submission review of the NFR data from 2010 to 2012. The most common issue identified was a substantial change in total NFR counts; 404 units were reviewed in 2010, and about half (204) of these were followed up with the respondent. For the majority of these large changes, the respondent verified that the reported count was accurate. Combined with the increasing proportions of issues accepted in review or verified as correct by the respondent in 2011 and 2012, the decline in large count changes across years provides further evidence that the NFR data quality is improving and becoming more stable.

Despite these improvements in reporting, respondents still had a harder time providing NFR counts than providing graduate student and postdoc counts in 2012. The imputation rates for each NFR item, shown in table 7, ranged from 7.3% to 7.6%. The NFR imputation rates are higher than the rates for other GSS data items (which are typically around 4.6%) and reflect the relative difficulty in gathering and reporting the NFR data.

In 2012, the total NFR count was 19,639 before imputation and 21,908 after imputation. Table 8 shows that the majority of the count changes were due to changes in reporting rather than from imputation. Moreover, the increased correlation and decreased standard deviation of the 2010 to 2012 counts by field show that imputation tends to stabilize trends by field. Table 8 also provides the reported values and the imputed values of total NFR counts by field for 2009 to 2012.

## **Conclusions**

This working paper assesses the reliability of institutional reporting of NFRs in the GSS and also the effect of 2010 data collection changes to improve the accuracy of NFR reporting. Overall, the 2010 to 2012 reporting of NFRs was fairly consistent and became more stable over time. Many institutions appear to have definitions and procedures for identifying NFRs. As expected, there was a considerable increase in numbers and percentages of units reporting NFRs from 2009 to 2010; this coincided with the expansion of data collection on postdocs and NFRs and with the appointment by some institutions of postdoc coordinators who were more knowledgeable about their institution's postdoc and NFR data. Comparison of the NFR data from 2010 through 2012 showed that the reporting has been fairly stable at both the school and unit levels since the expansion of the postdoc and NFR data collection in 2010.

However, some issues in NFR reporting remain. One issue is that the definition of NFRs is not consistent across, or even within, schools. The comparison of postdoc and NFR counts from 2009 through 2012 revealed that some school coordinators are still unable to differentiate postdocs and NFRs, resulting in not reporting one or the other or, occasionally, reporting the

TABLE 6. Review of NFR data reported by the institution: 2010–12

Type of check	Data reviewed	Data accepted in review	Data updated in review	Followed up with institution	Data verified	Data revised	Data unresolved
2010							
Total NFR count	404	170	30	204	141	37	26
NFR gender distribution	13	9	0	4	2	0	2
NFR medical degree distribution	32	19	0	13	8	0	5
Subtotal	449	198	30	221	151	37	33
Percent of all	100.0	44.1	6.7	49.2	33.6	8.2	7.3
Percent of retrieved	-	-	-	100.0	68.3	16.7	14.9
2011							
Total NFR count	301	207	0	94	51	35	8
NFR gender distribution	12	12	0	0	0	0	0
NFR degree type distribution	51	44	0	7	3	4	0
Units with same (nonzero) NFR data as prior year	83	61	0	22	18	1	3
Units with same (nonzero) NFR data in the current year	106	83	0	23	15	3	5
Other NFR comment	115	107	1	7	2	4	1
Subtotal	668	514	1	153	89	47	17
Percent of all	100.0	76.9	0.1	22.9	13.3	7.0	2.5
Percent of retrieved	-	-	-	100.0	58.2	30.7	11.1
2012							
Total NFR count	243	195	0	48	33	13	2
NFR gender distribution	2	0	0	2	1	1	0
NFR degree type distribution	26	23	0	3	1	1	1
Units with same (nonzero) NFR data as prior year	88	62	0	26	21	4	1
Units with same (nonzero) NFR data in the current year	66	44	0	22	17	3	2
Other NFR comment	151	143	3	5	4	1	0
Subtotal	576	467	3	106	77	23	6
Percent of all	100.0	81.1	0.5	18.4	13.4	4.0	1.0
Percent of retrieved	-	-	-	100.0	72.6	21.7	5.7

NFR = nonfaculty researcher.

SOURCES: National Science Foundation, National Center for Science and Engineering Statistics, Higher Education Research and Development Survey, and Survey of Graduate Students and Postdoctorates in Science and Engineering.

TABLE 7. NFR item imputation rates: 2012

Item	Number imputed	Imputation rate (%)
NFRs: Total	1,018	7.3
NFRs: Men	1,025	7.3
NFRs: Women	1,025	7.3
NFRs: With medical degrees (MD, DO, DDS, DVM)	1,057	7.6
NFRs: With medical degrees (MD, DO, DDS, DVM), men	1,054	7.6
NFRs: With medical degrees (MD, DO, DDS, DVM), women	1,045	7.5
NFRs: With nonmedical degree	1,057	7.6
NFRs: With nonmedical degree, men	1,059	7.6
NFRs: With nonmedical degree, women	1,050	7.5
NFRs: With dual medical/research doctoral degree	1,057	7.6
NFRs: With dual medical/research doctoral degree, men	1,054	7.6
NFRs: With dual medical/research doctoral degree, women	1,045	7.5
NFRs: With unknown degrees	1,057	7.6
NFRs: With unknown degrees, men	1,054	7.6
NFRs: With unknown degrees, women	1,045	7.5

NFR = nonfaculty researcher.

SOURCES: National Science Foundation, National Center for Science and Engineering Statistics, Higher Education Research and Development Survey, and Survey of Graduate Students and Postdoctorates in Science and Engineering.

TABLE 8. Reported and imputed NFR counts, by field: 2009–12

GSS code and description	Reported totals				Imputed totals			
	2009	2010	2011	2012	2009	2010	2011	2012
Total NFRs	13,807	19,095	19,372	19,639	14,059	21,345	21,498	21,908
Aerospace engineering, GSS CODE 101	40	53	32	49	40	58	35	49
Agricultural engineering, GSS CODE 102	52	64	62	63	52	70	62	65
Biomedical engineering, GSS CODE 103	150	211	198	258	153	250	247	295
Chemical engineering, GSS CODE 104	215	248	184	179	224	265	204	211
Civil engineering, GSS CODE 105	181	213	252	251	181	256	278	298
Electrical engineering, GSS CODE 106	292	366	339	362	296	395	406	405
Engineering science, GSS CODE 107	124	114	97	159	124	114	119	170
Industrial engineering, GSS CODE 108	76	87	70	56	76	108	87	70
Mechanical engineering, GSS CODE 109	245	335	282	298	246	355	318	389
Metallurgical and materials engineering, GSS CODE 110	172	214	223	236	180	224	233	245
Mining engineering, GSS CODE 111	1	7	4	10	1	7	4	10
Nuclear engineering, GSS CODE 112	28	37	40	29	28	39	44	30
Petroleum engineering, GSS CODE 113	17	23	36	40	17	23	36	40
Engineering nec, GSS CODE 114	113	160	222	196	113	227	228	209
Astronomy, GSS CODE 201	352	400	381	380	352	400	381	380
Chemistry, GSS CODE 202	753	899	898	814	766	944	997	911
Physics, GSS CODE 203	601	759	754	831	611	825	857	916
Physical sciences nec, GSS CODE 204	44	72	74	89	44	82	87	89
Atmospheric sciences, GSS CODE 301	190	250	299	306	190	250	310	319
Geosciences, GSS CODE 302	312	465	474	420	312	502	529	523
Oceanography, GSS CODE 303	178	282	390	319	178	300	405	334
Earth, atmospheric, and ocean sciences nec, GSS CODE 304	94	306	180	159	94	310	381	337
Computer sciences, GSS CODE 401	314	305	284	313	331	318	326	349
Mathematics and applied mathematics, GSS CODE 402	151	146	135	172	153	159	162	190
Statistics, GSS CODE 403	7	14	12	19	7	14	12	19
Agricultural sciences, GSS CODE 501	424	533	574	533	431	572	581	567
Anatomy, GSS CODE 601	75	92	103	95	75	96	106	101
Biochemistry, GSS CODE 602	503	573	609	670	523	619	652	708
Biology, GSS CODE 603	401	688	622	566	434	790	639	598
Biometry and epidemiology, GSS CODE 604	177	227	312	348	177	345	322	348
Biophysics, GSS CODE 605	11	44	30	32	15	53	30	32
Botany, GSS CODE 606	154	173	205	236	154	187	216	260
Cell biology, GSS CODE 607	443	539	527	598	448	611	607	627
Ecology, GSS CODE 608	66	61	67	72	66	62	67	72
Entomology and parasitology, GSS CODE 609	50	74	78	73	50	83	83	80
Genetics, GSS CODE 610	245	268	295	306	254	329	312	320
Microbiology, immunology, and virology, GSS CODE 611	555	602	703	728	562	779	800	822
Nutrition, GSS CODE 612	66	111	148	145	66	149	160	153
Pathology, GSS CODE 613	309	449	360	437	309	484	438	445
Pharmacology, GSS CODE 614	353	445	445	399	374	470	515	431
Physiology, GSS CODE 615	337	445	496	408	339	507	532	479
Zoology, GSS CODE 616	17	29	18	12	17	29	18	18
Biological sciences nec, GSS CODE 617	347	585	665	705	350	678	727	755
Anesthesiology, GSS CODE 701	69	114	115	122	69	137	140	138
Cardiology, GSS CODE 702	139	143	133	176	139	180	167	202
Oncology and cancer research, GSS CODE 703	418	429	625	631	437	465	631	665
Endocrinology, GSS CODE 704	94	98	81	99	94	106	87	110
Gastroenterology, GSS CODE 705	59	58	55	74	59	73	60	77
Hematology, GSS CODE 706	89	96	99	108	89	121	117	110
Neurology, GSS CODE 707	179	272	173	180	179	302	205	221
Obstetrics and gynecology, GSS CODE 708	48	62	64	65	48	80	80	72
Ophthalmology, GSS CODE 709	127	200	180	222	127	210	206	246
Otorhinolaryngology, GSS CODE 710	29	107	51	77	29	110	56	78
Pediatrics, GSS CODE 711	208	521	474	307	211	573	533	485
Preventive medicine and community health, GSS CODE 712	178	363	542	546	185	562	563	601
Psychiatry, GSS CODE 713	122	213	180	218	140	252	218	234
Pulmonary disease, GSS CODE 714	50	80	63	71	50	92	96	107
Radiology, GSS CODE 715	188	247	221	295	188	273	293	347

TABLE 8. Reported and imputed NFR counts, by field: 2009–12

GSS code and description	Reported totals				Imputed totals			
	2009	2010	2011	2012	2009	2010	2011	2012
Surgery, GSS CODE 716	305	435	315	395	305	465	355	429
Clinical medicine nec, GSS CODE 717	633	903	931	921	634	1,010	1,023	952
Dental sciences, GSS CODE 718	52	113	92	66	55	116	94	85
Nursing, GSS CODE 719	27	113	79	89	31	113	79	92
Pharmaceutical sciences, GSS CODE 720	200	296	250	277	205	305	270	307
Veterinary sciences, GSS CODE 721	183	283	254	239	183	300	254	250
Other health nec, GSS CODE 722	114	256	253	290	114	307	258	302
Speech pathology and audiology, GSS CODE 723	53	36	38	33	53	36	38	37
Psychology, general, GSS CODE 801	104	205	200	211	105	220	219	230
Psychology, GSS CODE 802	154	212	174	164	154	216	194	184
Clinical psychology, GSS CODE 803	32	31	21	17	32	31	21	17
Agricultural economics, GSS CODE 901	31	34	20	19	31	34	22	21
Anthropology (cultural and social), GSS CODE 902	71	67	75	88	75	67	81	101
Economics (except agricultural), GSS CODE 903	30	38	58	59	30	44	61	62
Geography, GSS CODE 904	35	37	39	48	35	37	49	61
History and philosophy of science, GSS CODE 905	8	9	15	18	8	9	15	18
Linguistics, GSS CODE 906	10	29	11	12	10	29	13	13
Political science, GSS CODE 907	37	108	183	200	37	143	221	229
Sociology, GSS CODE 908	15	46	65	68	16	48	66	69
Sociology and anthropology, GSS CODE 909	3	0	0	1	3	0	0	1
Social sciences nec, GSS CODE 910	124	186	144	159	132	206	144	165
Family and consumer sciences and human sciences, GSS CODE 920	31	36	101	33	31	38	101	43
Communication, GSS CODE 930	9	19	15	12	9	24	17	14
Architecture, GSS CODE 940	6	12	0	6	6	15	11	11
Neuroscience, GSS CODE 950	77	180	325	322	77	191	378	356
Multidisciplinary and interdisciplinary studies, GSS CODE 980	231	440	479	330	231	467	509	497
Mean standard deviation of 2009–12 field counts	-	-	-	42	-	-	-	51
Mean standard deviation of 2010–12 field counts	-	-	-	28	-	-	-	22
Correlation of current and prior year field counts	-	0.958	0.970	0.982	-	0.948	0.981	0.992

GSS = Survey of Graduate Students and Postdoctorates in Science and Engineering; nec = not elsewhere classified; NFR = nonfaculty researcher.

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

same individuals in both counts. Another problem uncovered during data collection and from the NFR Debriefing Survey was that school coordinators had data access issues. In particular, several school coordinators explained that their databases do not capture the information needed—such as type of degree (master's or doctorate) and descriptive job titles—to determine whether researchers have the credentials to meet the NFR definition. In many schools, NFRs are simply not tracked well. More than 100 schools reported having postdocs and no NFRs in each of the past 6 years. These schools likely have NFRs but are not reporting them for various reasons. Other concerns with lower potential impact on the NFR and postdoc data include reporting identical counts of NFRs and postdocs, reporting NFRs and no postdocs, and reporting more NFRs than postdocs. The schools with specific reporting issues will be followed up with in the next GSS cycle.

The number of schools and units providing incomplete NFR data is high compared to the other GSS data. Although GSS respondents were able to provide NFR counts for most units, they were unable to provide any NFR data for about 7% of all units in each year. Of the units that reported having NFRs in 2010, 2011, or 2012, only 47% were able to provide complete NFR information in each of these years, and 5% (346 units) were able to report that they had NFRs but could not provide the count. The fact that 29% of these units with NFRs reported a total NFR count in some but not all years suggests that these data are available but may be difficult to obtain within the institution.

Additional work is needed to address these NFR reporting issues. For example, data reviews with both internal and external sources will continue along with further follow-up efforts with school and postdoc coordinators. Similarly, some schools might benefit from the assignment of a different coordinator—especially one who could distinguish between the faculty/nonfaculty status of staff researchers—to report NFRs. Data collection and quality control protocols will also be improved to notify school coordinators when some units are able to report NFRs and others are not; additional information can be shared between the unit respondents about how to access the NFR data.

NCSES is committed to deepening our understanding of this increasing segment of the academic workforce. Through additional qualitative research, NCSES will continue to seek input from the academic researchers and will work with the GSS institutions to improve the collection of NFR-related data.

## **Data Sources and Limitations**

This working paper is based on the GSS institutions' reported and imputed data in 2007 to 2012. Results presented herein, with the exception of tables 3 and 4 and associated text, use imputed data and can therefore be reproduced using the GSS public-use data. Tables 3 and 4, which look at the ability of schools and units to report the data, use reported, rather than imputed, values. The analysis comparing the institution-level GSS NFR counts with the R&D expenditure and personnel data uses the 2012 GSS and the 2012 HERD data.

The 2010 GSS postdoc expansion began to address NSF's concerns about the reporting of NFRs and postdocs as well as the need for additional data on both. The expansion of the GSS Code List in 2007, the change in focus from degree-granting graduate programs to eligible units regardless of degree-granting status, and the appointment of separate postdoc coordinators improved coverage of postdocs in SEH fields. The expansion of the NFR and postdoc data items and the appointment of more knowledgeable postdoc coordinators improved data reporting within extant units.

In addition to improving the reporting of overall NFR and postdoc counts, the GSS in 2010 and subsequent years provides more detailed information on postdocs, including race and ethnicity for U.S. citizens and permanent residents, federal and nonfederal sources of financial support and support mechanism, type of doctoral degree (such as MD, PhD, or MD/PhD), and origin of doctoral degree (United States or a foreign country). Breakdowns of these data are available in *Graduate Students and Postdoctorates in Science and Engineering* at <http://www.nsf.gov/statistics/gradpostdoc/>.

## Notes

1. Peter Einaudi and Ruth E. Heuer are research analysts, and Patricia J. Green is a survey director at RTI International, Research Triangle Park, NC. For more information, contact Kelly H. Kang, Human Resources Statistics Program, National Center for Science and Engineering Statistics, National Science Foundation, 4201 Wilson Boulevard, Suite 965, Arlington, VA 22230 (kkang@nsf.gov; 703-292-7796).
2. For more information on the Postdoc Pilot Study and associated efforts to improve postdoc reporting, see Einaudi P, Heuer R, Green P. 2013. *Counts of Postdoctoral Appointees in Science, Engineering, and Health Rise with Reporting Improvements*. InfoBrief NSF 13-334. Arlington, VA: National Science Foundation, National Center for Science and Engineering Statistics. Available at <http://www.nsf.gov/statistics/infbrief/nsf13334/>.

## Appendix A – GSS NFR Question Change

### 2009 GSS

**In fall 2009, how many postdoctoral researchers (postdocs) and other doctorate-holding nonfaculty researchers did this organizational unit have in the categories below?**

*[Postdoc instructions omitted]*

**Doctorate-holding nonfaculty researchers (U.S. and foreign) to be included:**

- Those not considered postdoctoral researchers or members of the faculty and
- Who are involved principally in research activities

Postdocs by support and citizenship and other doctorate-holding nonfaculty research staff (report individuals in whole numbers)	Sex and selected degree field			
	Men 1	Women 2	Total 3	Of the total in Col 3, how many have an MD, DO, DDS, or DVM? 4
<b>Postdocs by largest mechanism of support</b>				
• Federal fellowship..... A				
• Federal traineeship..... B				
• Federal research grant..... C				
• Nonfederal sources..... D				
<b>Total postdocs (sum Rows A-D)..... E</b>				
• Of the total postdocs (Row E), how many are foreign nationals holding temporary visas..... F				
<b>Doctorate-holding nonfaculty researchers..... G</b>				

### 2010 GSS

**Please report the number of other doctorate-holding nonfaculty researchers in this organizational unit in fall 2010 by sex and type of degree. This includes individuals who are not considered postdocs or members of the faculty, and who are primarily involved in research.**

Type of doctoral degree	Nonfaculty Researchers		
	Men 1	Women 2	Total 3
Number of nonfaculty researchers with a <b>professional degree</b> (such as MD, DVM, DO, or DDS)..... A			
Number of nonfaculty researchers with a <b>doctoral degree</b> (such as PhD, ScD, DEng) ..... B			
Number of nonfaculty researchers with <b>both a professional degree and a doctoral degree</b> (such as MD-PhD, DVM-PhD)..... C			
Number of nonfaculty researchers with <b>doctoral degree type unknown</b> ..... D			
<b>Total nonfaculty researchers (sum rows A-D)..... E</b>			

## Appendix B – GSS NFR Debriefing Survey Questions

As part of the methodological research on the Survey of Graduate Students and Postdoctorates in Science and Engineering (GSS), NSF is interested in learning more information about how the institutions define and identify doctorate-holding nonfaculty researchers (NFRs), other than postdocs.

We have a few questions that we would like to ask you about the NFR data you reported to GSS.

- A1. Does your organizational unit have a commonly used definition for doctorate-holding nonfaculty researchers, other than postdocs?

Yes (*go to A2*)                       No

- A2. [IF YES] What is the definition?

---



---



---

- A3. Please see the job titles below and provide additional job titles for the research staff employed in your organizational unit.

*Check the job title used to identify NFRs and provide any additional criteria used to report NFRs by each applicable job title.*

	Job title	Check if this job title applies to NFRs	Comments or Additional criteria Used
A4a	Research scientist (any level – Principal, Senior, Full, Associate, Assistant)		
A4b	Research associate (any level)		
A4c	Research assistant		
A4d	Visiting scholar or fellow		
A4e	Postdoctoral fellow or researcher		
A4f	Instructor (any level)		
A4g	Research investigator		
A4h	Other research staff – specify job title:		

- A5. How did you determine who is to include in your count as a NFR (a nonfaculty doctorate-holding researcher, other than a postdoc)?

---



---



---

A6. Below is a table showing how you reported NFRs for your organizational unit.

Doctoral degree type	Men	Women	Total
Number of nonfaculty researchers with a <b>professional degree</b> (such as MD, DVM, DO, or DDS)			
Number of nonfaculty researchers with a <b>doctoral degree</b> (such as PhD, ScD, DEng)			
Number of nonfaculty researchers with <b>both a professional degree and a doctoral degree</b> (such as MD-PhD, DVM-PhD)			
Number of nonfaculty researchers with <b>doctoral degree type unknown</b>			
<b>Total nonfaculty researchers</b> (sum rows A-D)			

a. Did you have any difficulty in reporting doctoral degree type for the NFRs?

Yes (*go to A6b*)       No (*go to A6c*)

b. [IF YES] What were the difficulties involved in reporting NFRs' doctoral degree type?

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

c. Was it clear that the only professional degrees to be reported in the first doctoral type category were the doctorate-level professional degrees?

Yes       No

[IF 2010 NFR # GTE (2009 NFR # \* 1.15), GO TO SECTION B; OR ELSE GO TO SECTION C]

B1. It appears that the number of NFRs at your institution has increased significantly in the past few years. Your institution reported:

<2009 NFR #> NFRs in 2009

<2010 NFR #> NFRs in 2010

<2011 NFR #> NFRs in 2011

<2012 NFR #> NFRs in 2012

Has the definition and/or process you use to identify NFRs changed in the past few years?

Yes (*go to B2*)       No (*go to B3*)

B2. [IF YES] Can you please explain what has changed? *Please check all that apply.*

B2a. More units were added that employ NFRs [ONLY FOR COORDINATORS]

B2b. NFR identification process was changed, e.g., pulling information from a central database

B2c. Institution adopted new definitions for NFRs

B2d. Other changes – please specify \_\_\_\_\_

B3. Can you explain the relatively large increase in the number of NFRs your institution reported in 2010 from 2009?

---

---

---

[IF 2010 NFR # IS GT 2010 POSTDOC #, GO TO SECTION C; OR ELSE GO TO SECTION D]

C1. In 2010, you reported: <2010 NFR #> NFRs compared to <2010 POSTDOC #> postdocs.

Is it possible that your institution may have included some postdocs in the number of NFRs in 2010?

Yes  No

C1a. [IF YES] What are some of the difficulties your institution may have in reporting postdocs separately from NFRs?

---

---

---

C1b. [IF NO] Can you provide any information that may help explain the higher number of NFRs than postdocs at your institution?

---

---

---

[IF ZERO NFRS REPORTED, GO TO SECTION D]

D. Your school reported data on postdocs, but did not report other nonfaculty doctorate-holding researchers.

D1. Do you know whether your school has other nonfaculty doctorate-holding researchers (NFRs)?

Yes (go to D2)  No (go to D3)

D2. [IF YES] Why were you unable to report the NFRs?

---

---

---

D3. [IF SCHOOL HAS GTE 40 POSTDOCS] Most schools that report postdocs also report having NFRs. Although your institution reported <2010 POSTDOC #> postdocs in 2010, you did not report any NFRs. Can you provide any information that might be helpful to explain the differences?

---

---

---