

## APPENDIX B: METHODOLOGY

We used a mixed methods approach that included secondary data analysis, primary data collection and analysis using a survey questionnaire, and interviews conducted during institutional site visits.

### *Analysis of Attendance Patterns and Completions*

We used two existing data sets to examine sending and receiving institutions for NSF fellows and doctoral completion rates—the National Science Foundation’s Cumulative Index (CI) and annual Survey of Earned Doctorates (SED).

*Cumulative Index (CI).* The CI is a file that contains records for every individual who applies for a GRF. Records include information on undergraduate institution and undergraduate performance, GRE scores, the outcome of the review process, and demographics. The CI goes back to the beginning of the fellowship program in 1952 and was updated each year until 1989. From 1989 to 1993, partial information was added each year, but it has not been maintained since then. In order to undertake complete analysis through 1993 awardees, we also obtained from NSF an updated data set extracted from the GRF Program’s internal management information system that identified NSF fellows and Quality Group 2 non-awardees. This enabled us to analyze data for the 1979 through 1993 cohorts of fellows.

If individuals applied more than once, they will be in the CI more than once. We therefore selected the latest applicant record for each individual in order to assess their success in receiving a fellowship. Demographic data for individual years, such as the 1979 and 1993 comparisons in Tables G6 and G7 are based on all applicants in those years. We matched each institutional Office of Scientific and Engineering Personnel (OSEP) identification code to Carnegie institutional categories (Carnegie Corporation, 1994) using a crosswalk provided by NSF contractor Quantum Research Corporation Inc.

*Survey of Earned Doctorates (SED).* The SED is administered annually to all new doctoral recipients from U.S. institutions. Graduate schools are responsible for submitting completed forms to the National Opinion Research Center (NORC), which administers the survey on behalf of NSF. A study undertaken by the National Research Council (NRC, 1996) confirmed high self-report response rates to this survey of 94%. There is excellent coverage of research doctorates because NORC is also able to create skeleton records for all those who do not return a questionnaire, based on information provided by institutions. Each year new data are added to the cumulative file, known as the Doctorate Records File (DRF). The most recently added year covers those who received doctorates between July 1998 and June 1999.

With NSF authorization, WestEd provided NORC with a file containing the following information for each NSF fellow and Quality Group 2 non-awardee from the CI (unduplicated) from 1979 through 1993: social security number; name (first, last, middle); date of birth (day,

month, year); gender; and baccalaureate institution OSEP code. The file included 10,104 individuals awarded the GRF (including 1295 who declined the fellowship) and 3379 QG2 non-awardees.

Using this file, NORC staff sequentially performed six matching tests, including visual review to eliminate false matches, with the 1999 DRF and returned to WestEd a file that indicated whether a doctorate had been granted, year of award, institution, and field. WestEd received data from NORC on 8589 matched cases indicating completion of the Ph.D. According to NORC staff, these cases represented the best possible match given the available data. The matched cases were then compared to the updated GRF data set to generate a database from which we undertook the completions analyses contained in the body of this report. This database included 6535 individuals awarded the GRF (including 727 who declined the fellowship) and 2054 QG2 non-awardees. Decliner information is not included in this report.

*Graduate Program Quality Ratings.* We used as our measure of the quality of programs attended by GRF fellows ratings from the National Research Council's 1993 study (NRC, 1995). This study collected information on 3634 research-doctorate programs at 274 U.S. universities. To generate reputational measures of quality, the study conducted the National Survey of Graduate Faculty in Spring 1993. We used the 93Q measure of program effectiveness, which is defined in the study as:

the 1993 trimmed mean for scholarly quality of program faculty. Dropping the two highest and two lowest scores on the survey obtain the trimmed mean before computing the average. For purposes of analysis, scores were converted to a scale of 0 to 5 with 0 denoting 'not sufficient for doctoral education' and 5 denoting 'distinguished'. (NRC 1995, p. 25)

The scale was converted to five groupings, with programs scoring 4.01+ categorized as Distinguished, 3.01-4.00 = Strong, 2.51-3.00 = Good, 2.00 – 2.50 = Adequate, 1.00 – 2.50 = Marginal, and less than 1 = Not Sufficient (p. 32).

The field code contained on the SED file can be crosswalked to the NRC ratings file, and so we were able to attach to each record the program Quality Rating for each doctoral completer based on their institution of graduation and field code. Field codes on the CI, however, were not compatible with the NRC categorization, and so we were unable to attach a Quality Rating to programs in which NSF fellows enrolled. We used as a proxy for this measure information from a study undertaken by Webster and Skinner (1996). These researchers used the NRC program ratings to develop discipline group ranking by institution. Institutions were included when they had specific numbers of programs included in the NRC study by discipline grouping as follows: Biological Sciences = 7; Engineering = 8; Physical Sciences and Mathematics = 4; Social and Behavioral Sciences = 3. Only the top twenty institutions in each group were ranked. This more general grouping allowed us to relate field categories from the CI to the discipline group ranking.

## *The Graduate Student Follow-up Survey*

The Graduate Student Follow-up Survey was administered to three samples—the Disciplinary sample, the MGF sample and the WENG sample.

### Samples

*Disciplinary Sample.* We wanted to compare NSF fellows to non-fellows (program peers) enrolled in the same graduate programs. We identified a database that allowed the selection of a comparison group of graduate students who were enrolled in the same programs as many GRF recipients. This was the American Association of Universities (AAU)/Association of Graduate Schools (AGS) database on doctoral students, administered by the Educational Testing Service (ETS). The AAU/AGS Project for Research on Doctoral Education database was established in 1989 to collect student-level data from AAU institutions. Forty institutions have participated although some have not done so on a regular basis. The first fields included were Biochemistry, Economics, English, Mathematics, and Mechanical Engineering. Four of these are SMET fields included in NSF's mission. In 1992, the fields of Chemical Engineering, History, Physics, Electrical Engineering, and Psychology were added.

We obtained authorization from the AAU/AGS Project's Steering Committee to use the database. We also received permission from each of the participating institutions to use their data from the database. We needed first to match the CI to the AAU/AGS file to see how many NSF fellows who received awards between 1989 and 1993 were also in the latter file. ETS undertook this match and returned to us the tabulated results. They found approximately 500 fellows in the four NSF disciplines of Biochemistry, Economics, Mathematics, and Mechanical Engineering. From these tabulations we established criteria for inclusion in the study.

We included all institutions that had at least two beginning NSF fellows in the match. The Massachusetts Institute of Technology, which enrolls high numbers of NSF fellows had not participated in the AAU/AGS database but agreed to work with WestEd directly to enable its NSF fellows and their program peers to participate in this evaluation. Because of the concentration of NSF fellows by field in a few institutions, we estimate that we included in the survey samples approximately 61% of 1989-1993 fellows in Biochemistry, 81% in Economics, 62% in Mathematics, and 71% in Mechanical Engineering.

When we had identified the 15 institutions that met our criteria (plus MIT), we contacted their chief academic affairs officers to seek permission to use the AAU/AGS database to draw a sample of peers who began the same programs as the NSF fellows at the same time. Appendix D contains the list of participating institutions. We also asked them to help us locate addresses for these NSF fellows and program peers.

After we obtained this permission, ETS provided WestEd with an extract of the AAU/AGS database for these institutions. This allowed us to match the file to the CI to identify NSF fellows in the AAU/AGS database. Our final count was 480. We then selected a probability sample of peers, stratified by institution and the four disciplines. Our aim was to select two peers for each fellow surveyed (a 2X sample), anticipating that we might have more difficulty in locating peers

and that their response rate may be lower since the questionnaire clearly is related to the GRF. There were a few cases where there were insufficient program peers for a 2X match. However, since we did not analyze the four groups as a single set, these differences did not disrupt the design of the study. Discounting the individuals we could not contact (194 or 17.15%), our response rate for the Disciplinary Sample is 41.41%. Completed surveys received from 200 NSF fellows (55.71%) and 188 program peers (32.53%) have been included in the analysis for this report.

*MGF Sample.* We also administered the same Graduate Student Follow-up Survey to 200 MGF recipients (35% sample), regardless of discipline or institution enrolled in. The Minority Graduate Fellow sample was randomly drawn from the Cumulative Index from 1989-1993. The MGF sample included fellows in 33 disciplines at 62 institutions. Discounting individuals we could not contact (25 or 12.50%), the response rate is 49.71% and includes questionnaires received from 88 MGF recipients. The MGF sample was analyzed independently of the Disciplinary sample. There was no comparison group of peers for this sample.

*WENG Sample.* In order to see how the Women in Engineering GRF recipients have fared in comparison to other fellows, we also administered the Graduate Student Follow-up Survey to 143 WENG recipients from 1990-1993 (a 50% sample).

The WENG sample was randomly drawn from the Cumulative Index from 1990-1993 and included fellows in various sub-fields of engineering (not just the Mechanical Engineering area that was the focus of the Disciplinary sample) at 46 institutions. After using the same follow-up and search procedures and discounting individuals we could not contact (18 or 12.58%), the response rate is 68.00% and includes questionnaires received from 85 WENG recipients. The WENG sample was analyzed independently of the Disciplinary sample analysis. There was no comparison group of peers for this sample.

### Questionnaire Design and Administration

WestEd developed the questionnaire in 1998. We pilot tested an alpha version of the survey with four former NSF fellows and two non-fellows. These respondents represented several disciplines and institutions, different enrollment years, and programs with different quality ratings. The panel of experts and NSF staff also reviewed the instrument. A beta version was then pilot tested with two respondents to confirm that the changes worked and establish the amount of time to complete. Approval for its administration was obtained from the Office of Management and Budget (OMB) in December 1998. The questionnaire was mailed in January 1999 and continuously thereafter as possible addresses were obtained for additional recipients. Those living abroad received the questionnaire by Federal Express, where possible, or by U.S. mail. Appendix C contains the survey questionnaire.

*Location strategies.* For Disciplinary fellows and peers, we sent lists to the 16 institutions for help in locating respondents. We also sought institutional help in locating fellows drawn for the MGF and WENG samples from 69 institutions. Many of the addresses given to us by the institutions were no longer valid, and two institutions, citing privacy concerns did not provide addresses for some or all students. We sent postcard follow-ups to individuals whose

questionnaires were not returned either completed or as undeliverable. We hired a private investigation firm to further search for current addresses, but some individuals still could not be found, and more questionnaires were returned as undeliverable. We counted as “located” individuals whose questionnaires were not returned as undeliverable. It is highly unlikely that in fact all these surveys reached their intended recipients. So our response rate is calculated on a base that removes only those whom we know that we could not locate and may in fact be higher than we have reported here.

Our experience demonstrates the difficulty of research involving graduate students also experienced by other researchers. Institutions do not keep track of their former graduate students, and NSF does not keep track of its former fellows.

As expected, the response rate is higher for NSF fellows than program peers. Comparing the distribution of respondents by disciplinary area to the total population of NSF fellows, we found them to be quite similar (Table B1). This is especially true for MGF respondents, whose distribution by discipline area is almost identical to the population of fellows in 1993. Because of the small number of cases, we do not separate out most of the analyses of MGF responses by discipline (Engineering being the exception). The similarity of the disciplinary distribution between the total MGF population and the MGF sample enables us to be confident that the results are representative.

**Table B1**  
*Disciplinary Distribution of Survey Respondents*

	Engineering (Mechanical Engineering)	Math / Comp Science / Phys. Sciences (Math)	Biological/Life Sciences (Biochemistry)	Behavioral & Social Sciences (Economics)
GF 1993 Population	31%	23%	24%	23%
Disciplinary Fellows	27%	25%	16%	32%
Disciplinary Peers	22%	27%	26%	26%
WENG	100% *			
MGF 1993 Population	34%	16%	23%	27%
MGF	41%	16%	21%	22%

\*5 WENG fellows received their graduate degrees in Math. They are all treated as Engineering fellows for purposes of this study since they were awarded fellowships in Engineering.

*Data Analysis.* Responses were entered into a database. A set of derived variables was designed to simplify analysis of responses to many of the questions. The more detailed data collected through the questionnaire enabled us to choose appropriate categorizations and investigate responses that appeared inconsistent. Differences observed did not attain statistical significance.

We found three significant problems with respondent accuracy. First, 34 respondents did not give us information beyond their bachelor’s degree in response to question 1.17. The second problem occurred in the transition from Section 1 of the questionnaire to Section 2. We did not clearly phrase the transition question to emphasize that only those who had been continuously enrolled in graduate school as their primary activity should skip Section 2.

We estimate that about one-quarter of respondents incorrectly failed to answer Section 2. Third, at the beginning of Section 3, we asked for information about financial support in graduate school, but our respondents like those in other studies, such as the SED, did not grasp the intricacies of their sources of funding. Responses to these questions have not been included in the report.

We also added descriptive information to some data elements. All institutions were given their appropriate OSEP code and through that their Carnegie Classification was added. Respondents were asked to choose from a table that was provided the field code for their study. These fields, along with the OSEP code, were designed to map to the NRC Ph.D. program ratings, and we added the 93Q value for each designated field of study.

### *Institutional Site Visits*

Two-person teams conducted interviews over a 2- to 4-day period at six research universities. With advice from the panel of experts and NSF staff, these institutions were selected on the basis of enrollment of significant numbers of NSF fellows as well as geographical location and institutional type. In addition to administrators and staff responsible for graduate studies and fellowships at each university, we interviewed faculty, staff, and students in a total of 19 departments corresponding to the four disciplines selected for the Disciplinary sample for the survey: Biochemistry (5), Economics (3), Mathematics (6), and Mechanical Engineering (5). Interview protocols were approved by OMB in December 1998 and may be found in Appendix E. Interviews were tape recorded to ensure accuracy of the accounts.

Teams created site reports for each institution that were combined in Hyperqual2, a qualitative data analysis program (Padilla, 1993). The data was then sorted using a coding plan (Appendix F) to identify patterns and issues. This sorting made it possible to read across the six site reports and compare responses by departments and by type of person interviewed. Each code produced a set of findings supported by data from the interviews. Both site report text and direct quotations have been used in this report to illustrate findings.

At the six institutions, we interviewed 75 administrators, faculty, and staff. We interviewed 149 students (73 NSF fellows and 76 peers). Only one student interviewed (a peer) indicated having a disability. There was more gender balance in the NSF fellows interviewed (Table B2).

**Table B2**

*Comparison of NSF Fellows and Peers Interviewed: Gender*

Gender	Fellows N=73	Peers N=76
Men	41 (56.2%)	52 (68.4%)
Women	32 (43.8%)	24 (31.6%)

Only 5.5% of NSF fellows and 6.6% of peers interviewed were Hispanic, but the NSF fellows were more racially and ethnically diverse than peers interviewed (Table B3). Since only U.S. citizens and Permanent Residents are eligible for the GRF, there were no international students

among NSF fellows; however, 14% of the peers interviewed were international students (Table B4). Most of the NSF fellows and peers interviewed were in the second through fourth year of their graduate program (Table B5).

**Table B3**  
*Comparison of NSF Fellows and Peers Interviewed: Race and Ethnicity*

Race and Ethnicity	Fellows N=73	Peers N=76
Ethnicity		
Hispanic	4 (5.5%)	5 (6.6%)
Not Hispanic	67 (91.8%)	69 (90.8%)
Unknown	2 (2.7%)	2 (2.6%)
Race		
American Indian/Alaskan Native	0	0
Asian	12 (16.4%)	8 (10.5%)
Black/African American	3 (4.1%)	3 (3.9%)
Native Hawaiian/Pacific Islander	0	1 (1.3%)
White	51 (69.9%)	63 (82.9%)
Multiple Races Identified	5 (6.8%)	1(1.3%)
Unknown	2 (2.7%)	0

**Table B4**  
*Comparison of NSF Fellows and Peers Interviewed: Citizenship Status*

Citizenship Status	Fellows N=73	Peers N=76
U.S. Citizen	72 (98.6%)	59 (77.6%)
Permanent Resident	1(1.4%)	3 (3.9%)
International Student	N/A	14 (18.4%)

**Table B5**  
*Comparison of NSF Fellows and Peers Interviewed: Year in Program*

Year in Program	Fellows N=73	Peers N=76
1	9 (12.3%)	11 (14.5%)
2	22 (30.1%)	21 (27.6%)
3	22 (30.1%)	13 (17.1%)
4	9 (12.3%)	21 (27.6%)
5, 6 or just completed	11(15.1%)	10 (13.2%)

