Section B: Statistical Methods

B.1. Respondent Universe and Sampling Methods

The respondents for which the clearance is sought includes two groups: (1) IGERT trainees (N=750) and (2) IGERT Principal Investigators (N=40). We provide a brief description of these respondent groups below. The target population includes:

1. All currently enrolled IGERT trainees (including those who had received IGERT funding in a prior year as well as those receiving funding in the current academic year) from the 2007 and 2008 cohorts of IGERT awards who were enrolled in graduate school as of fall 2011;
2. All Principal Investigators (PIs) of the 2007 and 2008 cohorts of awards whose projects are still active.

The sampling frame for projects was deliberately constructed to meet the following criteria: representation of projects funded across all directorates; projects that are active and that have been operational long enough to have moved beyond initial implementation; projects that are likely to have students at different points along the graduate education trajectory; and projects housed at a variety of higher education institutions. Most projects support ~5-10 students/year, and generally, trainees are funded for one or two years, typically in the first two years of their graduate studies. This study will include both trainees who are currently receiving funding and those who are still enrolled in their graduate programs but are no longer receiving IGERT funding. Trainees who have graduated from or left their doctoral degree program will not be included because the study's research questions address how (projects and their) students are experiencing training in the six specific skill areas described above.

We anticipate a response rate of at least 80 percent from the respondent groups based on response rates of similar surveys conducted with samples of currently enrolled graduate students and faculty members who participated in NSF programs. A recent study of NSF's Graduate STEM Fellows in K-12 Education (GK-12) Program had a 92 percent response rate from currently enrolled doctoral GK-12 fellows, and a previous study of the IGERT program had an 85 percent response rate from currently enrolled IGERT trainees and an 88 percent response rate from participating faculty members and department chairs. We do not anticipate having any issues contacting trainees or PIs for this study given that they will still be directly tied to their IGERT respective university and we will be able to take advantage of the contact information maintained in the Distance Monitoring System.

B.2. Information Collection Procedures/Limitations of the Study

The following steps will be taken to collect data from trainees and PIs described in the previous section.

Step 1: Create list of respondents using IGERT Distance Monitoring System data. NSF's Distance Monitoring System (DMS) contains a list of all ever-funded IGERT trainees and PIs which will be used to create a list of respondents for the study. PIs will be asked to update a list of funded trainees (based on the DMS data from spring 2011) to reflect names of trainees selected to participate in academic year 2010-2011 because the DMS will not have this information at the time that the survey will be fielded.

Step 2: Locate respondents. The IGERT DMS includes contact information for IGERT trainees and PIs (e.g., email address, mailing address, and phone number while in graduate school) which will be used to invite respondents to participate in the study.

Step 3: Web survey. Once approval is obtained from OMB, the survey will be programmed for online data collection. The study team will test the survey system to ensure functionality and accuracy of data capture. Survey data collection is scheduled to begin in late Fall 2011 or early 2012.

Step 4: Recruitment and Data collection.

- All subjects will be sent an invitation email by NSF, introducing the study and the contractors conducting the study (Abt Associates Inc. and ICF/Macro).

Trainee Survey **(see Attachment A)**:

- Abt will follow up with an email to the PIs explaining the process and the schedule and asking them to encourage trainees to participate.
- Trainees will be sent an email explaining the purpose of the study that also contains a unique link to the survey. Three email reminders and three telephone reminders will be used to boost response rates. The survey will be open for 1 month. If desired response rates have not been achieved at that time, Abt may decide to extend the survey deadline by one or two weeks. Throughout the data collection cycle, a project phone number and e-mail address will be available to allow potential respondents to easily and quickly obtain answers to questions or concerns about the study.

PI Interview (see Attachment B):

- Abt will send a follow-up email to the PIs asking them to provide convenient times over the next two or three weeks for a telephone interview. PIs will be sent an

email reminder and a telephone reminder, if necessary, to make sure we obtain a high response rate.

- Following receipt of the times for the interview, Abt will schedule the interview and on confirmation, will conduct the interview using the OMB-approved semi-structured protocol. While Abt staff will take notes during the interview, permission will also be sought to tape the interview to ensure accuracy.

- After developing the final interview protocols, all interviewers will be trained on the interview protocol so that questions are standard across interviews. Only Abt staff members with working knowledge of the program will conduct these semi-structured interviews. Interviews will be scheduled using contact information obtained from the PIs. Interviewees will be assured that information they provide will not be released in any form that identifies them as individuals and their responses will be kept confidential. Interviews will be tape recorded so that notes can be captured and analyzed using a combination of Microsoft Access and NVivo software packages.

There are three major limitations of the proposed study. First, there is no one commonly-accepted definition of interdisciplinarity, so the dimensions/traits being explored in the study are necessarily exploratory. A second limitation is that student outcomes are self-reported by the trainees—they will be asked about their perceived confidence in skills across these areas to conduct interdisciplinary research. The study is not measuring these skills directly. However, given that there is no commonly accepted definition of interdisciplinarity, nor a commonly accepted set of outcomes, a study that describes perceptions of importance of skills, preparation in these skill areas, and means by which trainees acquire capabilities in these areas can be useful.

B.3. Estimation Procedure

Simple descriptive analyses of the data will use means and standard deviations of continuous measures and percentages for ordinal or binary measures. These will provide an overview of the data and characteristics of programs and students in the sample. To the extent that the study contractor needs to correct for nonresponse, this will be explained along with possible implications for the analysis in terms of biases introduced into the estimates by the particular methods selected.

Our data will encompass both closed-ended and open-ended questions. At a broad level, the analysis will consist of the following:

The closed-ended questions will be analyzed through descriptive cross-tabulations (for example, the frequency of PIs who report that trainees conduct an interdisciplinary team research project during their first year; or the frequency of trainees who report that participating in interdisciplinary team research projects was the most helpful way to develop

knowledge of a discipline(s) other than one's primary discipline). We plan to use cross-tabulations to explore whether patterns of responses differ by institutional or student characteristics and to the extent feasible, by discipline or IGERT thematic focus.

The open-ended responses will be reviewed to develop a set of coding bins into which the data can be usefully categorized. The coded data will then be analyzed to see what patterns emerge and the extent to which we perceive common themes across projects or unique to some projects (for example, the majority of PIs reported that, in their opinion, interdisciplinary coursework represented one strategy to provide trainees a working knowledge of multiple discipline(s) other than their primary discipline but opinions varied about how best to implement these courses...). We will also analyze whether the patterns differ by institutional or student characteristics and to the extent feasible by discipline or IGERT thematic focus.

The open-ended responses from the PI and the trainees will be back-coded, where possible, and will also supplement the cross-tabulations and simple descriptive measures (such as percentages).

If feasible, we will conduct subgroup analyses to examine whether trainees' perceived outcomes vary by demographic and personal characteristics (e.g., discipline of undergraduate or graduate degree, timing and duration of IGERT participation).

B.4. Methods for Maximizing the Response Rate and Addressing Issues of Nonresponse
Method to maximize response rate are described in detail in section B.2. Briefly, these will include the following procedures:

1. Web format of the survey
2. Minimization of spam filtering
3. Invitation from NSF to participate in the study
4. Skip patterns, to reduce burden on respondents
5. Extensive email and telephone follow-up
3. Availability of a toll-free phone number and email address for questions.

Given past experience, the PI interviews should attain a high response rate and it is unlikely that we will need to adjust for nonresponse.

We will use the procedures described below to examine the bias in estimates because of nonresponse. Based on the analysis we will adjust the weights of responding students to account for student nonresponse.

1. *Examination of Response Rates.* The first step will be to monitor the overall response rate, as well as by year and by relevant subgroups (e.g., by discipline, or by gender and

race/ethnicity). High response rates (over 80 percent) for the entire sample as well as for subgroups might indicate no need for further analysis of bias due to nonresponse. Large differences in the response rates by strata and for subgroups serve as indicators that potential biases may exist. For example, if the response rate from an important subgroup is very low, then measures calculated for this group would lack precision. In addition, any difference in the characteristic of interest between this subgroup and other subgroups would result in a bias in the estimates. From the survey results we will examine whether there are differences in the characteristics in the subgroups, especially in a stratum where the response rate is low.

2. *Nonresponse Propensity Model.* Finally, should the response rate fall below 80 percent we will construct a propensity model to estimate the probability of a trainee responding to the survey both for responding and nonresponding trainees; this is called a propensity score. The estimated propensity scores come from a logistic regression model. The model will be based on variables which are available both for nonresponding and responding students. Trainees will be grouped using the estimated propensity scores. Within each group we will compare the frame characteristics of responding and nonresponding trainees. This grouping in addition to assessing the bias will also provide a method of forming weighting classes for adjusting the weights of responding trainees to reduce the bias due to nonresponse.

B.5. Tests of Procedures or Methods

The study team conducted an extensive review of the literature on interdisciplinary graduate training and did not find any pre-existing validated instruments or scales that were relevant to the research questions this study addresses—that is, the relationship of individual skills to interdisciplinary research training. Therefore the trainee survey and PI interview protocol were developed specifically for this study. An External Advisory committee (EAC), comprised of four individuals with expertise in interdisciplinary graduate training, STEM higher education, and program evaluation reviewed the study's overall design plan and data collection instruments.[1] The committee members consulted about the wording, order, clarity and relevance of the individual items for both the student survey and the PI interview protocol, and the study team revised these instruments based on that feedback. The study team is pilot testing the PI interview and trainee survey with up to nine respondents each who do not fall in the sampling criteria for this study. Pilot-testers of the trainee survey complete a paper version of the survey and comment on the clarity, sequencing, and content of the questions as well as the amount of time required to complete the survey. The PI interview pilot test protocol includes the questions and probes about question wording, content, and sequencing. Responses and general feedback obtained during pilot tests will inform revisions to both instruments (content, wording, sequencing etc.).

[1] The EAC includes: Monica Cox – Director, Pedagogical Evaluation Laboratory and Associate Professor of Engineering/Purdue University; Irwin Feller – Professor Emeritus of Economics/Pennsylvania State University; Lisa Lattuca – Professor of Higher Education/University of Michigan; and Nancy Nersessian – Regents' Professor of Cognitive Science/Georgia Institute of Technology.

B.6. Names and Telephone Numbers of Individuals Consulted

Key personnel who have been involved in the statistical aspects and who will be involved in collecting and analyzing data are presented in the table below. The contractor for collection and analysis of data in this study is Abt Associates Inc. Staff have experience in evaluation of research programs, expertise in scientific research, and knowledge of statistical methods, were involved in the design. NSF program staff members familiar with the programs have been included in the design of the evaluation.

Table B.5 Individuals Consulted

Name	Role	Phone
Abt Associates Inc.		
Beth Gamse	Project Director, Principal Associate	617-349-2808
Amanda Parsad	Director of Analysis, Senior Associate	301-634-1791
Kristen Neishi	Senior Analyst	301-634-1759
Radha Roy	Senior Analyst	301-347-5722
National Science Foundation		
Carol Stoel	Program Officer, Division of Graduate Education	703-292-8630
Melur Ramasubramanian	Program Director, Division of Graduate Education	703-292-5089
Maura Borrego	AAAS Fellow (former) Program Director for Math and Science Partnership (MSP) and Science, Technology, Engineering, and Mathematics Talent Expansion Program (STEP) (current)	703-292-7855

References

Aboelela, S.W., Larson,E., Bakken, S., Carrasquillo, O., Formicola, A., Glied, S.A., Haas, J., and Gebbie, K.M. (2007). Defining Interdisciplinary Research: Conclusions from a Critical Review of the Literature. Health Services Research 42: 1, Part 1 (February), 329-346. Published by Health Research and Educational Trust.

Boix Mansilla, V. (2006). Quality assessment of interdisciplinary research: Toward empirically grounded validation criteria. Research Evaluation 15 (1), April, 69–74.

Borrego, M. & Newswander, L. (2010). Definitions of Interdisciplinary Research: Toward Graduate-Level Interdisciplinary Learning Outcomes. The Review of Higher Education, 34(1), 61-84.

Carney, J.G., Chawla, D., Wiley, A., & Young, D. (2006). Evaluation of the Initial Impacts of the National Science Foundation's Integrative Graduate Education and Research Traineeships (IGERT) Program. Prepared for the National Science Foundation.

Carney, J.G., Martinez, A., Dreier, J., Neishi, K., Parsad, A. (2010). Evaluation of the National Science Foundation's Integrative Graduate Education and Research Traineeships (IGERT) Program: Follow-up Study of IGERT Graduates. Prepared for the National Science Foundation.

Chase, A. & Giancola, J. (2001). NSF Integrative Graduate Education and Research Traineeships Monitoring Report: Boston University, The Bioinformatics Project. Prepared for the National Science Foundation.

Chase, A., Giancola, J., Smith, C., Boulay, B., Gamse, B., Horst, L., Moss, M., Goldsmith, S., Haviland, D., Tushnet, N. (2002). *IGERT Annual Cross-Site Report: 1998 Cohort*. Prepared for the National Science Foundation.

Coppola, B..P., Banaszak Holl, M.M., Karbstein, K. (2007). Closing the Gap between Interdisciplinary Research and Disciplinary Teaching. American Chemical Society/Chemical Biology, 2(8), 518-520.

Dauphine´e, D., & Martin, J.B. (2000). Breaking Down the Walls: Thoughts on the Scholarship of Integration. Academic Medicine 75 (9) , 881-886.

Feller, I. (2006). Multiple actors, multiple settings, multiple criteria: issues in assessing interdisciplinary research. Research evaluation 15 (1), 5—15.

Jacobs, J. & Frickel, S. (2009). Interdisciplinarity: A Critical Assessment. Annual Review of Sociology. 35, 43-65.

Lattuca, L. & Knight, D. (2010). In the Eye of the Beholder: Defining and Studying Interdisciplinarity in Engineering Education. American Society for Engineering Education.

Martinez, A., Chase, A., Boulay, B., Chawla, D., Layzer, C., Litin, L., Zotov, N. (2006). Contractor Annual Report and Summary of the Cross-Site Monitoring of the NSF Integrative Graduate and Education Research Traineeship (IGERT) Program. Prepared for the National Science Foundation.

Nersessian, N.J. (2009). How do engineering scientists think? Model-based simulation in biomedical engineering laboratories, Topics in Cognitive Science, 1:730-757.

Schilling, K. L. (2001). Interdisciplinary assessment for interdisciplinary programs. In B. L. Smith & J. McCann (Eds.), Reinventing ourselves: Interdisciplinary education, collaborative learning and experimentation in higher education (pp. 344–54). Bolton, MA: Anker.

Van Hartesveldt, C.., & Giordan, J. (2008). Impact of Transformative Interdisciplinary Research and Graduate Education on Academic Institutions: National Science Foundation/Education and Human Resources Directorate/Division of Graduate Education, Integrative Graduate Education and Research Training Program Workshop Report (May).