

# Supporting Statement (3145-0136)

**Request For Revision of Clearance: National Science Foundation, Directorate of Education and Human Resources, Division of Graduate Education**

**Distance Monitoring of Integrative Graduate Education and Research Training (IGERT) Program**

**Attachment E**

## Section A

### Introduction

This request for Office of Management and Budget (OMB) review asks for a renewal of clearance for the distance monitoring data collection for the Integrative Graduate Education And Research Traineeship Program (IGERT) that is part of the National Science Foundation's (NSF) Directorate for Education and Human Resources (EHR) Generic Clearance, OMB 3145-0136, which expires on January 31, 2008. The EHR Generic Clearance includes collections of information about NSF's education and training (E&T) activities.

### A.1. Circumstances Requiring the Collection of Data

The IGERT program was initiated in 1998 to catalyze a cultural change in graduate education, for students, faculty, and institutions, by establishing innovative new models for graduate education and training in a fertile environment for collaborative research that transcends traditional disciplinary boundaries. It is also intended to facilitate diversity in student participation and preparation, and to contribute to a world-class, broadly inclusive, and globally engaged science and engineering workforce. IGERT awards are made to academic institutions in the United States and its territories that grant the Ph.D. degree and have research and training programs in science and engineering. The awards to each institution may be up to \$500,000 per year for 5 years, with an additional \$200,000 available for appropriate instrumentation or research materials during the first year of the award. The majority of the funds must be used for doctoral student stipends, training and educational activities, and related expenditures, such as student travel, publication costs, and recruitment. More information about the IGERT program can be found at: <http://www.nsf.gov/crssprgm/igert/intro.jsp>

The IGERT program supports the NSF strategic outcome goal of cultivating "a world-class, broadly inclusive science and engineering workforce," and expanding "the scientific literacy of all citizens," labeled as 'Learning' on page 5 of the FY 2006-2011 Strategic Plan, <http://www.nsf.gov/pubs/2006/nsf0648/NSF-06-48.pdf>. In particular, the program will help promote NSF's Learning-related investment priorities to "develop methods to effectively bridge critical junctures in STEM education pathways" and "prepare a diverse, globally engaged STEM workforce" (reference page 7 of the plan).

Data collected from IGERT awards through the monitoring system are needed by NSF for project and program monitoring, to fulfill policy and program reporting needs, and to serve as preliminary work for future impact assessment and evaluation activities. The data collected as part of OMB 3145-0136 allow NSF officials to document the overall program investment in individual alliances, and make future funding and program policy decisions.

A crosswalk of questions on the data collection instruments is included in appendix A.

### A.2. Purposes and Uses of the Data

The information collected in this task is required for effective administration, communication, and program and project monitoring; for meeting reporting requirements; for measuring attainment of NSF's

program, project and strategic goals as laid out in NSF's Strategic Plan; and as a baseline for future program evaluations.

The primary purpose of this collection is program management, also known as program monitoring. This data collection activity is designed to track the extent to which IGERT awards meet the objectives of the program. Within the DGE division, this information is used to administer and monitor the progress of the program. The findings are used to recommend, among other things, administrative changes in program functions, level of award support, individual program focus and emphasis, and recruiting efforts. Serious lapses in adherence to program guidelines or administrative problems are flagged and can be addressed immediately (e.g., failure to recruit students or recruitment of ineligible students (support is restricted to U.S. citizens and permanent residents)).

The IGERT program also uses the data to fulfill reporting requirements. As a part of its performance assessment activities, NSF relies on the judgment of external experts to maintain high standards of program management. Directorate and Office advisory committees (ACs) meet twice a year, while Committees of Visitors (COVs) for divisions or programs meet once every three years. Data collected in the IGERT monitoring system may be used to report to these committees on program activities. In addition, NSF is required to measure the attainment of its program, project and strategic goals by the President's Management agenda as represented by the Office of Management and Budget's (OMB) Program Assessment Rating Tool (PART), by the Government Performance and Results Act (GPRA) of 1993, and by the NSF's Strategic Plan. Data collected in the IGERT monitoring system help NSF management examine their progress towards the Foundation's goals and respond to these reporting requirements.

Finally, the data can also be used as a preliminary step in more detailed evaluation efforts, such as the sort of rigorous evaluations described in the May 2007 Report of the Academic Competitiveness Council, which was established by the Deficit Reduction Act of 2005 (P.L. 109-171) to serve as a multi-agency effort to identify federal STEM education programs and establish their effectiveness. The full ACC report can be accessed at <http://www.ed.gov/about/inits/ed/competitiveness/acc-mathscience/index.html>. The data is currently used as a baseline for the program evaluations conducted by Abt Associates (OMB# 3145-0812).

The distance monitoring system includes two Web-based survey forms: the principal investigator (PI) survey and the trainee survey. The PI survey collects basic information on program participants and program implementation and management. The IGERT trainee survey collects information on trainee demographics, activities and accomplishments during the funding period, as well as information that will allow NSF to conduct future follow-ups with trainees. In order to encourage frank responses to questions, trainees have sole access to their own surveys, although PIs can monitor trainees' completion of each section of the survey via a management screen. Together these surveys allow for a comparison of the project as designed and implemented by the PI and faculty, and the project as it is experienced by the participant trainees.

The following is an overview of the types of information collected:

- **Project Characteristics**: The PI provides basic information about the program: what NSF disciplines are supported in the project, what institution faculty are advising students, if any consortial arrangements have been established and what additional funding has been received.
- **Outputs**: The PI describes the level of success in meeting the goals for the trainees as a group. Additionally, the PI describe what impacts the program has had on the primary institutions (e.g., institutionalized programs, faculty development).
- **Project Features**: The PIs provides detailed information about activities developed and used by the program (e.g., recruitment strategies, research training, training for future faculty). They also comment on the effectiveness of all planned activities.
- **Trainee Data**: The PI provides a list of trainees and some basic information about their progress towards a Ph.D and employment after leaving the program. The trainees themselves provide demographic data, information on educational background, achievements during the award period, and detailed information about activities related to training goals.
- **Trainee Commentary**: Trainees are given the opportunity to comment on training activities and the IGERT trainee program as a whole. PIs cannot access the trainee surveys and cannot read trainee comments.

Several changes were made to the data elements during preparation of the survey for the 2008 collection. Specifically, principal investigators (PIs) of IGERT awards are now asked to provide more information on their recruitment practices, on international opportunities, and the information requested on partnerships and project evaluation strategies has been modified. Several of these questions now ask PIs to provide information on what they believe to be the best practices of their IGERT award, to indicate how productive certain practices have been, and to describe a demonstrable institutional change that has occurred due to the IGERT program. Trainees are also now asked what they believe were the most beneficial parts of different aspects of their IGERT experiences. During a telephone conference on October 15, 2007, John Kramer of OMB confirmed that these questions could be submitted under the EHR Generic Clearance, with the understanding that the data collected could not be used to draw conclusions about the success of the program. These questions will be used only to identify activities of individual projects within the IGERT program and will not be used for evaluation purposes.

These changes to the IGERT system are currently in progress and screenshots of the system are not yet available. However, please see Appendix A1 and Appendix A2 for detailed crosswalks of data elements and details on changes to the data elements since the last clearance.

### **A.3. Use of Information Technology To Reduce Burden**

EHR tends to favor Web-based systems because they can facilitate respondents' data entry across computer platforms. One innovative feature of many of the individual Web systems is the thorough editing of all submitted data for completeness, validity and consistency. Editing is performed as data are entered. Most invalid data cannot enter the system, and questionable or incomplete entries are called to respondents' attention before they are submitted to NSF. Web based surveys employ user-friendly features such as automated tabulation, data entry with custom controls such as checkboxes, data verification with error messages for easy online correction, standard menus and predefined charts and graphics. All these features facilitate the reporting process, provide useful and rapid feedback to the data providers and reduce burden.

The data for this monitoring effort are collected by 508-compliant Web-based surveys. The survey structure allows respondents to move between a menu screen and a screen addressing individual topics. The question format is primarily quick-response checkboxes, with text boxes provided for the addition of specific, outstanding examples. Respondents may enter and leave their surveys as often as they desire and continue to change their responses until they submit their surveys. Respondents have access to an online glossary to assist them in understanding the specific meaning of terms in the context of these surveys. Additionally, since the collection is Web-based, minor changes in wording and displays have been easily made in response to user feedback.

### **A.4. Efforts To Identify Duplication**

This system does not duplicate other NSF efforts. Comparable data are not currently being collected on an annual basis for the IGERT program. In addition, the collection is coordinated with the NSF FastLane Project Reports system (OMB 3145-0058) to ensure that the two collections do not collect similar data. As much as possible, data from other NSF monitoring collections are used to pre-fill IGERT items, further minimizing overall response burden. Additionally, aggregate data are being shared with NSF-funded researchers as appropriate, thereby minimizing the possibility that other researchers will duplicate these efforts in their own future collections.

### **A.5. Small Business**

No information is to be collected from small businesses.

### **A.6. Consequences of Not Collecting the Information**

Without this information NSF will be unable to document the effectiveness or outcomes of the IGERT

program. The Foundation will be unable to disseminate information to other projects and institutions about successful approaches to the integration of research and education and graduate student training. Additionally, without this feedback NSF would have no way of making systematic modifications to the IGERT program (e.g., adequacy of funding amount, duration of award, institutional supports needed). Moreover, NSF will be unable to comply fully with the GPRA and PART reporting requirements or with the congressional mandate that the Foundation evaluate its science, technology, engineering, and mathematics (STEM) education programs.

### **A.7. Special Circumstances Justifying Inconsistencies with Guidelines in 5 CFR 1320.6**

The data collection will comply with 5 CFR 1320.6.

### **A.8. Consultation Outside the Agency**

The notice inviting comments on the EHR Generic Clearance (OMB 3145-0136) was published in the Federal Register August 24, 2007, Volume 72, Number 164, page 48694. No comments were received.

During the initial system development principal investigators (PIs) from IGERT awards reviewed the system; their responses to the PI survey and their assessments of the trainee survey were taken into account in the development of the system. Changes in the system since initial development are informed by ongoing consultations with the respondents, Macro International Inc. (the contractor that designed the Web interface and database system) and Abt Associates, Inc. (the contractor that performs program evaluations). Macro currently maintains the surveys and survey databases and provides technical support to respondents as needed.

### **A.9. Payments or Gifts to Respondents**

No payments or gifts will be provided to respondents.

### **A.10. Assurance of Confidentiality**

Data collected under this task are only available to the respondents, NSF, and the firms hired to manage the data and data collection software. Data are processed according to Federal and State privacy statutes. To protect privacy, only composite data or graphical representations will be released to the public.

For the collection covered by this clearance request, when respondents are presented with the first screen of the survey, they are additionally instructed as follows:

"Information from this data collection system will be retained by the National Science Foundation, a federal agency, and will be an integral part of its Privacy Act System of Records in accordance with the Privacy Act of 1974 and maintained in the Education and Training System of Records 63 Fed. Reg. 264, 272 (January 5, 1998). These are confidential files accessible only to appropriate National Science Foundation (NSF) officials, their staffs, and their contractors responsible for monitoring, assessing, and evaluating NSF programs. Only data in highly aggregated form, or data explicitly requested as "for general use," will be made available to anyone outside of the National Science Foundation for research purposes. Data submitted will be used in accordance with criteria established by NSF for monitoring research and education grants, and in response to Public Law 99-383 and 42 USC 1885c. The Social Security number (SSN) and date of birth will be maintained in accordance with the requirements of the Privacy Act of 1974. Submission of the SSN is voluntary. It is used for survey quality control, program evaluation, and for matching with other data sets maintained in the Education and Training System of Records 63 Fed. Reg. 264, 272 (January 5, 1998)."

## A.11. Questions of a Sensitive Nature

The IGERT survey requests that each trainee provide his or her name, SSN, phone number, a personal e-mail address (e.g. aol.com or hotmail.com), and contact information for a person likely to know how to reach him or her in three years. These data are collected in order to ensure consistent monitoring and to permit follow-up studies that examine the long-term effect of the IGERT program on individuals' success. The IGERT survey also collects the name, telephone number, e-mail address, fax number, disability status, and citizenship of each PI, co-PI or advisor. Additionally, trainee Graduate Record Exam (GRE) scores are collected. Respondents have the option of not providing information that they consider privileged, such as disability status, by marking the "not reported" option on the form or by leaving their SSN blank. Because the program requires that all IGERT trainees be U.S. citizens or permanent residents, the question of IGERT trainee citizenship is directly addressed both in the PI survey (by asking the PI to verify the IGERT-required citizenship/residency requirement for each trainee) and in the trainee survey (by not providing a "not reported" option for citizenship). All information will be maintained in accordance with the requirements of the Privacy Act of 1974. Individualized data are provided only to IGERT program staff and to contractors from Abt Associates, Inc. conducting site evaluations authorized by NSF. Any public reporting of the data is in aggregate form.

## A.12 Estimates of Response Burden

### A.12.1. Number of Respondents, Frequency of Response, and Annual Hour Burden

The total number of annual respondents is 2,136 (136 project PIs and 2,000 IGERT trainees and associates) and the total annual person-hours is 9,440. The Web-based collection is an annual activity of the IGERT program. There are currently 136 IGERT awards and data is collected from each award site. PIs complete the PI survey; all IGERT trainees and associates are required to complete the IGERT trainee survey. We anticipate that new awards may be added, but at about the same rate that active awards expire; thus, on average, the number of respondents will remain constant over time. The annualized burden was computed by taking the number of respondents from the current survey cycle and estimating their response burden, based on a question in the Web-based data collection asking how long it takes respondents to complete the survey. The burden estimates for each type of respondent are outlined below:

Type of Respondent	Average Number of Respondents	Burden Hours Per Respondent	Annual Person-Hours
Project PIs	136	40 hours	5,440
IGERT trainees	2,000	2 hours	4,000
Total respondents	2,136	Total estimated hours	9,440

### A.12.2. Hour Burden Estimates by Each Form and Aggregate Hour Burdens

As mentioned above respondents will be project PIs and IGERT trainees. The estimated total annual response burden is 9,440 person-hours. There is a different Web-based form for each respondent. The annual burden by form was calculated as follows:

Form Type	Respondent Type	Number of Respondents	Burden Hours Per Respondent	Total Person-Hours
Principal Investigator survey	Project PIs	136	40 hours	5,440
Trainee Survey	Trainees	2,000	2 hours	4,000

Total	2,136	9,440
-------	-------	-------

### A.12.3. Estimates of Annualized Cost to Respondents for the Hour Burdens

The overall annualized cost to the respondents is estimated to be \$241,280. The following table shows the annualized estimates of costs to respondents. The estimated hourly rates for PIs are based on a report in the April 20, 2007, edition of *The Chronicle of Higher Education* (2007. "What Professors Earn." *The Chronicle of Higher Education*, 53(33), Washington, D.C.: The Chronicle of Higher Education, Inc.). According to the report, the average salary of an associate professor across all types of doctoral-granting institutions (public, private, church-related) was \$76,639. When divided by the number of standard annual work hours (2,080), this calculates to \$37.00 per respondent hour.

Respondents	Number of Respondents	Hours per Respondent	Average Hourly Rate	Total Annual Costs
Project PIs	136	40 hours	\$37	\$201,280
IGERT trainees and associates	2,000	2 hours	\$10	\$40,000
<b>Total estimated costs</b>				<b>\$241,280</b>

### A.13. Estimate of Total Capital and Startup Costs/Operation and Maintenance Costs to Respondents or Record Keepers

There is no overall annual cost burden to respondents or record-keepers that results from the distance monitoring of the IGERT program other than the time spent responding to the data collection instrument.

It is usual and customary for individuals involved in education and training activities in the United States to keep descriptive records. The information being requested is from records that are maintained as part of normal educational or training practice. Furthermore, the majority of PIs are active or former grantees or participants in programs or projects once funded by NSF. In order to be funded by NSF, institutions must follow the instructions in the NSF Grant Proposal Guide (GPG) that is cleared under OMB 3145-0058. The GPG requires that all applicants submit requests for NSF funding and that all active NSF awardees do administrative reporting via FastLane, an Internet-based forms system. Thus, PIs who are the primary respondents to the IGERT data collection make use of standard office equipment (e.g., computers), Internet connectivity that are already required as a startup cost and maintenance cost under the NSF GPG. The information requested of trainees is typical of doctoral student educational and research portfolios and would be maintained as par of normal practice.

### A.14. Estimates of Costs to the Federal Government

Computing the annualized cost to NSF for the IGERT data collection was done by taking the budgets for 3 years and calculating the costs for each of the following operational activities involved in producing, maintaining, and conducting the IGERT data collection:

Operational Activities	Cost Over 3 Years
System development (includes initial development of the database and Web-based application and later changes requested by the program, e.g., increased reporting tools, additional validations)	\$400,125

System maintenance, updates, and technical support (system requires updates each year before opening the collection; maintenance is required to keep the system current with technology, e.g., database servers, operating systems)	\$200,040
Data collection opening and support (e.g., online and telephone support to respondents and contacting respondents to encourage completion of the questions), reporting (as defined by the Division of Graduate Education), and followup activities (e.g., providing data to other consultants)	\$240,000
<b>3-Year Total for All Operational Activities</b>	<b>\$840,165</b>

The annual costs of the clearance was computed as one-third of the total 3-year costs; therefore, the annualized cost to NSF for the IGERT data collection is \$280,055.

### **A.15. Changes in Burden**

The previously reported total annual response burden for this collection was 7,200 hours for 1,700 PIs and trainees; the current request for 9,440 hours for 2,136 respondents is an increase of 2,240 hours. This increase is due to gradual growth in the numbers of awards and trainees involved in the IGERT program.

Some changes have been made to the instruments since the last clearance, as some data elements were added or revised and others were deleted. However, the overall length of the survey has not changed and we do not anticipate that these changes will affect the respondents' burden. See the crosswalk of data elements in the appendices for more information on the changes to the survey.

### **A.16. Plans for Publication, Analysis, and Schedule**

Data collection is scheduled to begin in March, and award sites will have 90 days to enter data; extensions are granted by NSF program officers as necessary. Once the data collection has been completed, agency staff can access the data through the on-line system as needed.

Like many agencies, NSF is reducing its reliance on formal (i.e., traditional) publication methods and publication formats. Macro International Inc., the contractor that manages the data collection Web site and database, is forbidden contractually from publishing results unless NSF has made a specific exception. In short, all products of the collections are the property of NSF and NSF is the exclusive publisher of the information being gathered. Often it is only after seeing the quality of the information collected that NSF decides the format (raw or analytical) and manner (in the NSF-numbered product Online Document System (ODS) or simply a page on the NSF Web site) in which to publish.

The data from this collection will be used for internal review purposes and to monitor the IGERT projects, as well as for reporting to Congress (e.g., the GPRR Annual Performance Plan). Reports to NSF management, PIs, and Congress dealing with characteristics and performance of the IGERT program will include statistical tables and charts generated from the database. At this time, NSF has not set timeline for publishing interim reports from this study.

### **A.17. Approval to Not Display Expiration Date**

Not Applicable

### **A.18 Exceptions to Item 19 of OMB Form 83-I**

No exceptions apply.

## Section B

### Introduction

#### B.1. Respondent Universe and Sampling Methods

The sample size is the entire universe of respondents. There are currently 2,136 PIs and trainees working on IGERT awards and this number is expected to remain stable throughout the clearance period.

Population	Estimated Universe Size	Sample Size
IGERT respondents	2,136	2,136

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

This data collection uses a Web-based survey. Each IGERT project provides descriptive data each year for the duration of their NSF funding. The data are primarily useful for program management, monitoring and descriptive analysis.

NSF understands the limitations of the data collection, particularly in terms of using the data to determine program effectiveness. Data collected through the IGERT monitoring system are not used to determine the ultimate effectiveness of its STEM educational interventions, but are used in program planning and management, to report on agency activities and goals, and to lay the groundwork for future evaluations.

##### B.2.1. Statistical Methodology for Stratification and Sample Selection

This data collection is a census, so no sampling is required

##### B.2.2. Estimation Procedure

Not Applicable

##### B.2.3. Degree of Accuracy Needed for the Purpose Described in the Justification

Not Applicable

##### B.2.4. Unusual Problems Requiring Specialized Sampling Procedures

Not Applicable

##### B.2.5. Use of Periodic (Less Frequent Than Annual) Data Collection Cycles

Not Applicable

#### B.3. Methods for Maximizing the Response Rate and Addressing Issues of Nonresponse

Past collections have had 100 percent response rates and NSF anticipates that the rate will remain the same. This is achieved by sending emails every three weeks to award sites that have not logged into the system, and by notifying all award sites still entering data when the system closing date is one week away. Approximately 60% of award sites receive at least one of these follow-up emails. Examples of the emails announcing the opening of the system and reminding awardees to log in and enter data are included in appendix B. The collection is part of reporting required of awardees; principal investigators are responsible



for ensuring that individual data are collected from trainees, and will have access to status information on the Web site indicating which trainees have not responded.

#### **B.4. Tests of Procedures or Methods**

This system has been operational since 2000 and was tested extensively. In addition, many of the items and response categories follow formats that are already in place at NSF.

#### **B.5. Names and Telephone Numbers of Individuals Consulted**

##### **Agency**

Carol Van Hartesveldt, National Science Foundation, (703) 292-8696.

##### **Contractors**

Macro International, Inc. will be responsible for data collection and analysis under the direction of Lea Mesner, (301) 657-3077.