

B. Collections of Information Employing Statistical Methods

The agency should be prepared to justify its decision not to use statistical methods in any case where such methods might reduce burden or improve accuracy of results. When the question “Does this ICR contain surveys, censuses or employ statistical methods” is checked, “Yes,” the following documentation should be included in the Supporting Statement to the extent that it applies to the methods proposed:

- 1. Describe (including a numerical estimate) the potential respondent universe and any sampling or other respondent selection methods to be used. Data on the number of entities (e.g., establishments, State and local government units, households, or persons) in the universe covered by the collection and in the corresponding sample are to be provided in tabular form for the universe as a whole and for each of the strata in the proposed sample. Indicate expected response rates for the collection as a whole. If the collection had been conducted previously, include the actual response rate achieved during the last collection.*

The potential respondent universe includes principal investigators (PIs) serving on NSF-funded SEES and non-SEES projects. The respondents’ projects must start on or after January 1, 2010, the inception of the SEES initiative, and end on or before August 31, 2016, to allow sufficient information to be collected for this evaluation. SEES PIs will be asked to respond to the survey for the study if their SEES projects are funded by SEES programs selected for the comparative analyses. The selection of respondents to the survey includes three steps: (1) selecting comparable SEES and non-SEES programs, (2) selecting comparable SEES and non-SEES projects, and (3) selecting PIs to respond to the survey.

Selecting comparable SEES and non-SEES programs

There are 17 SEES programs, established at different time since 2010, that have different themes, focus on different scientific goals, and target different audiences; not all SEES programs have comparable programs at NSF and, as the Statement of Work suggests, a sample of SEES programs needs to be selected for the comparative analyses.

Several data sources have been used to select the SEES programs and their comparable non-SEES programs for the comparative analyses. The Statement of Work listed comparable programs for six SEES programs. Focus groups with 12 SEES program officers were held in September 2014, during which some non-SEES programs were suggested as comparable. In March 2015, a program officer survey was administered to program officers who have worked with SEES programs at NSF, to collect recommendations on comparable non-SEES programs. NSF program officers suggested 15 SEES programs with 47 potential comparables. These programs fall across two program categories: programs with projects focusing on domain-specific research activities and programs with projects focusing on education partnerships or collaboration activities.

Table 1 Potential Respondent Universe

Program	<i>J</i>	Program Activity	<i>j</i>	Projects (<i>n</i>)	Potential Projects and PI Respondents (<i>N</i>)
SEES Programs	1	Domain-specific scientific research activities	10	323	408
	5	Education, partnership, collaboration activities	5	85	
Comparable Non-SEES Programs	4	Domain-specific scientific research activities	38	2206	2629
	8	Education, partnership, collaboration activities	10	423	

The evaluation team further screened the SEES and comparable non-SEES programs against four criteria that reflect the overall programmatic and research goals of the SEES initiative:

1. Does the comparable program focus on environmental sustainability?
2. Does the comparable program focus on similar environmental topics?
3. Does the comparable program focus on interdisciplinary collaboration?
4. Does the comparable program focus on integrating social, economic, and behavioral dimensions of research?

With our evolving understanding of SEES through a historical review of SEES programs and a comparative analysis of SEES solicitation language, we selected 10 SEES programs and 10 primary comparable non-SEES programs for the comparative analyses as shown below. For some of the 10 selected SEES programs, there are secondary and tertiary comparable non-SEES programs. In the event that no or an insufficient number of comparable projects are found within the primary non-SEES programs, projects in the secondary or tertiary can be used.

Table 2 Selected Matched SEES and Non-SEES Programs (Primary Matches)

SEES Program	Number of SEES Projects	Non-SEES Comparative Program	Number of Non-SEES Projects
SEES Fellows	43	AGS Fellowship	33
Dimensions of Biodiversity	29	Population and Community Ecology	32
WSC	23	Hydrological Sciences	25
Cyber SEES	14	Cyber-Physical Systems	21
Hazard SEES	3	Engineering for Natural Hazards	25
EaSM	32	Climate and Large Scale Dynamics	47
Sustainable Energy Pathways	20	Energy for Sustainability	34
Ocean Acidification	48	Chemical Oceanography	48
Arctic SEES	3	Arctic System Science	8
Total	215		273

Selecting comparable SEES and non-SEES projects

Once SEES and comparable non-SEES programs were identified, projects within these programs were compiled for matching. All projects in the selected SEES programs are included in the sample. There are 215 projects in the 9 selected SEES programs, as shown in the table above. Each SEES project will have one or two matched non-SEES projects that focus on a related substantive domain with similar project duration and award size. Assuming half of the SEES projects (107 projects) have one matched non-SEES project each, and the other SEES projects (108) have two matches, the comparative analyses will have a sample of 538 projects, 215 SEES projects and 323 non-SEES projects (107*1+108*2).

Selecting PIs to respond to the survey

PIs on the selected SEES and non-SEES projects will be contacted to respond to the survey. In the case of multiple PIs serving on a project, we will select the PI NSF uses as the primary point of contact for the project. Therefore, the total number of respondents to be contacted for the survey is 538 PIs. With a response rate of 80%, the final analytic sample size is expected to be 430 PIs.

2. *Describe the procedures for the collection of information including:*
 - * *Statistical methodology for stratification and sample selection,*
 - * *Estimation procedure,*
 - * *Degree of accuracy needed for the purpose described in the justification,*
 - * *Unusual problems requiring specialized sampling procedures, and*
 - * *Any use of periodic (less frequent than annual) data collection cycles to reduce burden.*

To estimate the sample size needed for this study, the research team conducted a model-based power analysis based on a two-level linear model. The research team assumed a fixed effect two-level model where projects at Level 1 are modeled with the treatment effect and a vector of covariates. The matched programs will be the Level 2 unit of analysis with up to three program grouping dummy covariates.

Level 1 (project level):

$$y_{ij} = \beta_{0j} + \beta_{1j}(SEES) + \sum_2^p \beta_{pj} + e_{ij}$$

Level 2 (program level):

$$\beta_{0j} = \gamma_{00}$$

$$\beta_{1j} = \gamma_{10} + \sum_1^q \gamma_{1q} w_{1q} + u_{0j}$$

$$\beta_{pj} = \gamma_{p0}$$

where, in the Level 1 model, y_{ij} is the outcome measure of the i^{th} project in the j^{th} program, β_{0j} is the conditional project-specific intercept, β_{1j} is the effect of SEES, and β_{pj} is the effects of the p^{th} project-level predictor on the outcome. The Level 2 model specifies the SEES program effects as a function of program-level covariates, w_{1q} .

The power analysis based on the two-level linear model shows that, given an average of 20 projects per program and 18 programs, we will detect effect sizes of 0.25 with 80% confidence, assuming 40% projects are SEES and project-level covariates explain 30% of variance in project outcomes.

3. *Describe methods to maximize response rates and to deal with issues of non-response. The accuracy and reliability of information collected must be shown to be adequate for intended uses. For collections based on sampling, a special justification must be provided for any collection that will not yield "reliable" data that can be generalized to the universe studied.*

When developing the survey instrument, the research team mapped the survey questions to specific research questions to ensure all questions will collect information needed for the study. The survey questionnaire has been revised based on feedback from eight PIs in a pretest. To onboard PIs to respond to the survey, a pre-survey email message will be delivered to PIs of projects selected for the comparative analyses. The email message will be sent through the NSF server in order to avoid messages being blocked by the recipients' email servers. Email from NSF also tends to encourage timely responses from PIs. The message will provide PIs with information on the evaluation project, as well as an overview of the online survey. Unique survey URLs will be created for each respondent, allowing the research team to track survey responses and provide target follow-up, as needed. After the completion of the survey, the research team may conduct follow-up interviews with PIs to clarify their responses to open-ended questions and ask additional questions to check the reliability of the responses as well as the accuracy of our interpretations of the responses.

4. *Describe any tests of procedures or methods to be undertaken. Testing is encouraged as an effective means of refining collections of information to minimize burden and improve utility. Tests must be approved if they call for answers to identical questions from 10 or more respondents. A proposed test or set of test may be submitted for approval separately or in combination with the main collection of information.*

The research team conducted a pretest of the contact procedures and the online survey with eight PIs on SEES projects. The pretest allowed us to better understand the time it took respondents to complete the survey. We also asked respondents about their perceptions of the questions, whether any wording created confusion about specific questions, whether certain questions would not be applicable to their situation, recommendations to clarify certain terms, and other valuable feedback. Specifically, during the pretest, we probed the following eight questions to ensure high quality of the survey instrument:

- 1) How long does it take you to complete the survey? Is there any question unnecessary or redundant?
- 2) Is there any question that causes confusion? How did you interpret these questions? How would you change the question to make it clear?
- 3) Is there any question hard to answer because you couldn't recall?

- 4) Is there any question hard to answer because you were not sure about the accuracy of the information?
- 5) Do you feel any questions are out of order? Or, do you have any suggestions on the sequence of the questions?
- 6) Is there any question you think may lead to biased responses?
- 7) Have we overlooked any important questions?
- 8) Are there other revisions you'd like to suggest?

With the input received from the pilot surveys, the instrument was revised to better capture data needed for the comparative analyses.

5. *Provide the name and telephone number of individuals consulted on statistical aspects of the design and the name of the agency unit, contractor(s), grantee(s), or other person(s) who will actually collect and/or analyze the information for the agency.*

The National Science Foundation has contracted with the Manhattan Strategy Group (MSG) to conduct the Evaluation of the SEES Portfolio of Programs. MSG will be responsible for all data collection and analysis. The network analysis portion of the study will be conducted by a subcontractor, the National Opinion Research Center (NORC) at the University of Chicago.

Dan Geller, Ph.D., Director of Evaluation Services, Manhattan Strategy Group

Email: dgeller@manhattanstrategy.com

Office: 301-828-1348

Ying Zhang, Ph.D., Project Manager, Manhattan Strategy Group

Email: yzhang@manhattanstrategy.com

Office: 301-828-1346

Kevin Brown, Ph.D., Senior Research Scientist, NORC

Email: Brown-Kevin@norc.org