

SUPPORTING STATEMENT
FOR OMB CLEARANCE
PART B

NASA Summer of Innovation FY2012

2012 PROGRAM DATA COLLECTION

National Aeronautics and Science Administration

December 16, 2011

Contents

Part B: Collection of Information Employing Statistical Methods.....	3
Introduction.....	3
B.1 Respondent Universe.....	4
B.2. Procedures for the Collection of Information.....	5
<i>Implementation Forms</i>	8
B.2.1 Statistical Methodology for Sample Selection.....	9
B.2.2 Analytic Approach.....	11
Student Descriptive Analyses: Single Time Point.....	12
Teacher Descriptive Analyses: Single Time Point.....	13
Student Descriptive Analyses: Change Over Time Analyses.....	13
Teacher Descriptive Analyses: Change Over Time Analyses.....	16
B.3 Methods to Maximize Response Rates.....	16
B.4 Test of Procedures.....	19
B.5 Individuals Consulted on Statistical Aspects of Design.....	20
References.....	21

Part B: Collection of Information Employing Statistical Methods

Introduction

The purpose of the national formative study of NASA's Summer of Innovation (SoI) project is to gather data that will inform NASA's continued development of (SoI) as well as to assess whether a summative, impact evaluation would be warranted at a future date. This decision will engage a group of external experts—Robert Tai, Phoebe Cottingham, Laura LoGerfo and Thomas Cook—who will review the current evaluation activities and findings and be involved in the development of a research design that would answer questions of impact. To this end, an important component of the current evaluation will focus on identifying significant changes between baseline and follow-up student and teacher surveys. While measuring outcomes at multiple points in time can provide evidence of whether the outcomes of interest change, it will not allow us to rule out the possibility that something other than the program is affecting this change. As such, we emphasize that the national evaluation is a formative study, focused on gathering information to inform promising practices.

The Summer of Innovation (SoI) data collection for FY2012 will involve the eight national awards from FY2011 (submit to funding review in February 2012), new and continuing NASA Center partnerships.

The national awardees and NASA Centers have different programming requirements. SoI programming requirements for students are the following:

- Each national awardee is required to reach 2,500 middle school students and provide 40 hours of student STEM activities utilizing NASA content over the summer and an additional 25 hours by March 2012.
- Each NASA Center should reach 1,500 students through partnerships that provide a minimum of 20 hours of student STEM activities utilizing NASA content during the summer and an additional two STEM activities integrating NASA content by March 2012.

SoI programming requirements for teachers also differs, as follows:

- Each national awardees is required to provide 150 certified middle school teachers 40 hours of professional development by March 2012.
- NASA Center partnerships are not required to provide professional development.

Part A describes the data collection activities for the FY 2012 SoI national evaluation, including the parent consent form and associated survey, student activities, event and professional development implementation data, teacher surveys, and student surveys. Parent consent forms (and associated survey) will be collected from all parents of student participants in SoI activities held by the national awardees and NASA Centers. In addition, baseline and follow-up surveys will be collected from sampled students at both national awardee and NASA Center programs. Implementation data will be collected from all national awardees but not from the Centers as significantly more NASA resources are invested in the national award programs and the requirements for the NASA Center are fewer. Similarly, baseline and

follow-up teacher surveys will only be administered at the national awardee programs as NASA Centers are not required to reach classroom teachers.

These surveys will collect the data needed to examine whether SoI student interest in science and SoI teachers' comfort in teaching NASA topics and their access and use of NASA resources changes over time. The implementation forms will inform NASA about what content the awardees are using in their SoI programs. Surveys will be administered to classroom teachers (at national awardees only) and students at national awardees and NASA Centers prior to and at the end of the SoI summer activities. Given that linking student and teacher surveys is beyond the scope of this formative evaluation, particularly because students may have multiple teachers in their summer classrooms, student and teacher surveys will be analyzed as two distinct samples. An additional teacher and student survey will be administered after the school-year activities at the national awardees; the teacher survey will be online while the third student survey will be mailed to their home address. Because requirements for the school-year activities are minimal for the NASA Centers, their participating students will not participate in the third wave of survey data collection.

B.1 Respondent Universe

The evaluation activities will involve the SoI programs run by the eight national awardees who were initially funded in FY 2011 and the NASA Centers. Awardee PIs, Center leads and evaluation coordinators will participate in the implementation data collection efforts. The eight awardees are required to reach a total of 20,000 students and the ten Centers must reach a total of 15,000 for a total of 35,000 SoI student participants. The parents of all 35,000 students will fill out a consent form and survey. A sample of approximately 6,200 students will be surveyed (3,190 students at awardee programs and 3,010 students at Center programs). All 1,200 teachers participating in SoI will be surveyed and may fill out school-year quarterly implementation log forms.

All parents will receive and be asked to complete the parent consent form as part of the program's registration materials. The national evaluation will need to obtain consent from parents prior to administering any student surveys. In addition, NASA will need the demographic data for project monitoring and compliance assessments, thus the data of the universe of SoI participants will be used regardless of whether the student is sampled for the national evaluation. Likewise, the target population for the teacher surveys is all classroom teachers who participate in SoI. Teacher surveys will be administered to the census of classroom teachers participating in SoI national awardee programs. Further, implementation data will be collected from all national awardees. As a result, there will be no sampling considerations for parent consent forms, teacher surveys, and awardees' implementation forms.

However, students will be sampled from the anticipated 120 camps from the national awardees and 90 camps of the NASA Centers, based on the number of camps held in summer 2011. A sample of students of 3,190 students will be identified from all students participating in the national awardee programs and a separate sample of 3,010 students will be selected from all the students participating at the NASA Center partnership programs. The samples will be selected in accordance with the sampling plan outlined in the following section. As expectations and supports for awardees are different from those for NASA Centers, the national evaluation will analyze their data separately.

B.2. Procedures for the Collection of Information

The data collection procedures and instruments were designed to capture information on the implementation of SoI at National Awardees and to investigate teacher and student-related outcomes. Exhibit 1 outlines the data collection schedule to be implemented in the 2012 national evaluation. Because of differing programmatic requirements at national awardee sites and NASA Centers, survey administration will differ between the two programs. National awardees will administer teacher and student baseline and two follow-up surveys, one at the end of the summer activities and one at the end of the school-year activities (June 2013). NASA Centers, however, will only administer student surveys at two points in time, prior to and at the end of the summer activities. NASA Centers will not administer teacher surveys.

Exhibit 1. FY2012 Data Collection to Be Analyzed using Statistical Methods		
<i>Instrument</i>	<i>Timing of Data Collection</i>	
	<i>National Awardees</i>	<i>NASA Centers</i>
Student Data		
Parent Consent and Survey	At time of registration (Spring 2012)	At time of registration (Spring 2012)
Student Baseline Survey	First day of SoI program (July and August 2012)	First day of SoI program (July and August 2012)
Student Post-Summer Survey	Last day of SoI program (July and August 2012)	Last day of SoI program (July and August 2012)
Student School-year Survey	End of school-year activities (June 2013)	NA
Teacher Data		
Teacher Baseline Survey	At registration (Spring 2012)	NA
Teacher Post-Summer Survey	End of all summer activities (September 2012)	NA
Teacher School-year Survey	End of school-year activities (June 2013)	NA
Program Data		
Student Summer Activities Implementation Reporting Forms	At the end of each activity/SoI program (Summer 2012)	NA
Student School-Year Activities Implementation Reporting Forms	After each activity (School-year 2012-2013)	NA
School-Year Teacher Implementation Log Forms	At the end of each quarter (School-year 2012-2013)	NA
Summer Professional Development Implementation Forms	At the end of each activity/SoI program (Summer 2012)	NA
School-Year Professional Development Implementation Forms	At the end of each activity/SoI program (School-year 2012-2013)	NA

Student Data: Procedures for Data Collection

Parent Consent Form

As part of the registration process, awardees and NASA Centers will obtain parent consent forms and the associated survey (Appendices 1 and 2) from all students. The parent consent form will be available in two formats, paper or online. Offering multiple survey modes will ease the burden of data collection on the awardees and Centers, allowing them to use the most convenient mode for their participants.

Awardees and Centers that would like to offer the parent consent form online would provide registering parents with the survey URL and a site-specific PIN, needed to gain access to the survey (see Appendix 20 for an example of the PIN and agreement to participate screens). PIN numbers would be unique to each Awardee/Center and PIN authentication would be necessary for access to the parent consent form and survey. A link to the online survey will also be available on the NASA Summer of Innovation website. The data collected via the online surveys would be maintained on the survey vendor's secure server and then safely transferred to the national evaluator. Data from the online survey will be collected, stored, and transferred in accordance with NASA's privacy and security requirements.

The consent form has information about the evaluation, the purpose of data collection, potential risk, and privacy assurances. The associated survey asks for information about the parent's and the student's demographics. Parents will return the consent form and parent survey to the awardee/Center as part of the materials required to enroll their student in the SoI activities. Students whose parents do not grant consent to participate in the national evaluation can still take part in the SoI activities.

Baseline Student Surveys

Prior to the start of the summer program, the national evaluator will obtain Institutional Review Board (IRB) approval for the modifications to the FY2012 study design and instruments and will provide training to awardees/Centers' national evaluation PIs, coordinators and education leads to ensure rigorous and systematic data collection procedures. Throughout the program, the national evaluators will support the awardees/Centers in their data collection efforts. Evaluation guidance will be provided in FY2012 to awardees and Centers in the form of a comprehensive guide to the evaluation activities available online and in hardcopy.

The awardees' PIs or evaluation coordinators will provide the national evaluator with information about the number of students and the number of camps in an awardee/Center by May 1, 2012. This information will be used by the national evaluator to sample at the camp level (more detail regarding the sampling plan is provided in the next section). After sampling is complete, the national evaluator will provide the awardees' evaluation coordinators/PIs/Center leads with the list of the camps selected for survey administration. The awardees' evaluation coordinator/PI/Center lead will then be responsible for ensuring the proper administration of the paper-and-pencil baseline surveys on site at the start of the program to those students with parental consent; they provide students without parental consent an alternative activity to participate in during survey administration. As is true throughout the duration of the study, consent to participate may be withdrawn at any time without penalty or change in participation status.

The contractor expects that most students in a camp will have parental consent. Parental consent rates for a program with a similar parental consent form process yielded relatively high consent rates (almost 70%; Martinez & Consentino de Cohen, 2010). In FY2011, the parent consent rate for national awardees was 98% and at NASA Centers was 95% for the forms that were returned.¹ Further, given that the survey is benign in nature and that there are no consequences for not granting consent, the contractor does not expect that consenting parents will be markedly different from non-consenting parents. However, to test this assumption, the contractor will compare demographic information and reasons for enrolling students in the program for non-consenting students to those with consent. Demographic and enrollment information about non-consenting students may be available from parents who fill out the parent survey

¹ It should be noted that the response rates to the 2011 SoI consent forms were low: 63 percent for awardees and 15 percent for Centers. See pages 15-16 for discussion of these response rates and the changes that will be used to improve them in 2012.

(but do not give consent) or from sites collecting this information for their own data needs. Finally, survey data results will clearly state that inferences can only be made about students with consent.

First Follow-up Student Surveys

The summer (first) follow-up student survey will be administered to students with consent in sampled camps on the last day of students' summer activities. Again, awardees' evaluation coordinators/PIs/Center leads will be responsible for administering surveys to consenting students and providing alternative activities to students without consent. The purpose of the summer follow-up survey is to measure changes in science interest using the same questions included in the baseline survey. As such, summer follow-up survey outcomes will be compared to those from the baseline survey.

Second Follow-up Student Surveys

Where programs remain intact throughout the school-year, second follow-up surveys will be administered using the same process as for the first follow up surveys. However, as this will not be true of all programs, NASA's contractors will mail school-year (second) follow-up surveys to the home addresses of students with parental consent who participated in summer activities at awardees after the completion of the school-year activities in June 2013 (students who participated in summer activities at NASA Centers will not be administered a third survey, as requirements for school-year follow-up activities are minimal). The purpose of the school-year follow-up survey is to measure if science interest is sustained over the course of the school-year using the same questions as the summer follow-up survey.

The contractor expects that some of the students will move between the time their parents completed the parent survey and when the contractor would mail the survey. Accordingly, the contractor will email parents an online form in late 2012/early 2013 allowing them to update their address information. In addition, the contractor will perform one updating of parent addresses using the Lexis Nexis database, which provides access to public records to verify information. If response rates to the mailed survey are low, the national evaluation team will follow up with students using the email addresses and phone numbers provided by parents on the consent form.

Classroom Teacher Surveys: Procedures for Data Collection

Baseline Classroom Teacher Surveys

Classroom teachers at awardees who register to participate in the SoI program will receive a registration packet that includes registration materials and the baseline survey (because there are no classroom teacher requirements at NASA Centers, any teachers participating there will not be administered surveys). The PIs will be accountable ensuring that instructing teachers complete the baseline survey included in the registration packet, collecting the registration packet, and returning the baseline surveys to the national evaluator.

At some sites access to the internet is readily available and teacher feedback from the FY2011 evaluation suggested that teachers at these sites would appreciate the option of taking the survey online. Awardees and Centers that would like to offer the teacher survey online would provide registering teachers with the survey URL and a site-specific PIN, needed to gain access to the survey (see Appendix 20 for an example of the PIN and agreement to participate screens). PIN numbers would be unique to each Awardee/Center and PIN authentication would be necessary for access to the teacher survey. A link to the online survey

will also be available on the NASA Summer of Innovation website. The data collected via the online surveys would be maintained on the survey vendor's secure server and then safely transferred to the national evaluator. Data from the online survey will be collected, stored, and transferred in accordance with NASA's privacy and security requirements.

First Follow-up Classroom Teacher Surveys

The baseline survey collected at the beginning of SoI programming will contain the classroom teacher's contact information, including their email address. The email address included in the registration form will be used to send classroom teachers a message asking them to complete an online survey immediately at the end of SoI summer activities. Similar to the student surveys, the teacher surveys are designed to detect changes in outcomes between the follow-up summer survey and baseline survey using the same questions across survey waves. In the event that classroom teachers do not respond to the email (or it bounces back), the contractor will use additional information from the registration forms (e.g., phone number) to follow up with teachers by sending up to three email reminders and making up to three follow-up calls to encourage them to fill out the online survey at home or wherever they have internet access. The contractor will also offer them the option of taking a paper and pencil survey that the contractor will mail to them along with a pre-paid, pre-addressed, Business Reply Envelope. Two national awardees indicated in FY2011 that their teachers may not have internet access. The contractor will print paper surveys and mail them to the evaluation coordinator for administration at any awardee site where there is limited internet access.

Second Follow-up Teacher Surveys

An email message asking teachers to complete a school-year follow-up survey will be sent to teachers in June 2013 so that the contractor may assess whether there are any differences between school-year follow-up and summer follow-up surveys. In the event that classroom teachers do not respond to the email (or it is bounced back), the contractor will use additional information from the registration forms (e.g., phone number) to follow up with teachers by sending up to three email reminders and making up to three follow-up calls to encourage them to fill out the online survey at home or wherever they have internet access. Similar to the administration of the first follow-up survey, the contractor will also offer teachers with limited access to the internet the option of taking a paper and pencil survey that the contractor would be mailed to them along with a pre-paid, pre-addressed, Business Reply Envelope.

Program Data: Procedures for Data Collection

Implementation Forms

The implementation forms were designed to collect the information necessary to understand each awardees' implementation and discern lessons learned. These instruments allow NASA to capture awardees' plans for its program models and implementation strategies for later comparison with actual implementation. Links to electronic implementation forms will be sent to the evaluation coordinators at each awardee site to collect information about the awardees' professional development and student activities that are actually implemented. The awardee PIs and Center leads will be accountable for ensuring that leads of each camp complete a student activities implementation form at the conclusion of the summer camp sessions, at the conclusion of the student school-year events coordinated by the awardee, and at the end of each summer and school-year professional development session.

These forms ask awardees to report the actual dates of implementation, the content used, the number of contact hours, the number of hours during which NASA content was used, the number of participants enrolled and attending, reasons for why participants did not complete the activity, and who led the activities. The data collected through these forms allow for descriptions of the different approaches taken by awardees to meet the NASA requirements.

During the FY 2011 school-year activities, national awardees were frequently not involved in the delivery of school-year SoI activities. This structure necessitates collecting school-year implementation data directly from teachers/after-school instructors rather than from the awardees' coordinators. To accomplish this, the contractor designed a school-year teacher implementation log form to be electronically completed quarterly. Teachers associated with awardees not providing structured school-year activities receive reminders to complete the form via e-mail. This form is a shortened version of the summer implementation report forms; teachers are asked whether and what NASA resources they used and the number of hours of NASA content they provide. Collecting this data allows NASA to learn how school-year activities are implemented when an awardee does not coordinate them.

B.2.1 Statistical Methodology for Sample Selection

Sampling Plan

Sampling Frame for Teacher Surveys

The teacher respondent universe of 1,200, based on NASA requirements included in the cooperative agreements made with the national awardees. Each of the eight national awardees will be included in the teacher survey sample and each awardee is expected to provide professional development to least 150 classroom teachers. The universe of classroom teachers participating in SoI at awardees will be asked to complete the teacher surveys; thus, there are no sampling considerations.

Sampling Frame for Student Surveys

The student respondent universe of 35,000 is based on NASA requirements included in the agreements with the national awardees and NASA Centers. National awardees are expected to reach 20,000 students and NASA Centers are expected to reach 15,000 students. Given the limitations of surveying such a large number of students, a sample will be drawn to obtain a representative sample of students, as is described in the following section.

Power Calculations for Student Surveys

The basic approach to power analysis was to estimate the design effect (*deff*) for the Horvitz-Thompson (HT) estimator of the mean (for our sampling design) for each of the four difference score outcome variables, and then use this to compute the number of students who would be needed in order to achieve adequate statistical power.

For each of the outcome difference scores, the contractor defined a minimal mean difference score, m , that would be substantively meaningful. For the measures on a 1-5 scale, $m = .1$; for the measures on a 10-50 scale, $m = 1$. For each of the changes between survey administrations, the contractor will test the null hypothesis that the mean difference score equals zero with 80% power at the alternative that the mean difference score equals m . Because of attrition and the duration between baseline and the second

follow-up, the contractor expected that adequate power to test hypotheses about this change would require the largest sample sizes. To compute the sample sizes needed to achieve adequate power here, the contractor estimated the variance of these changes. Because the contractor only had estimates of the variance of changes occurring between baseline and the first follow-up, the contractor applied an “inflation factor” of 100% to these estimates to come up with plausible values for the baseline to second follow-up variances. In line with this doubling of variance, the contractor included a commensurate decrease in within-summer-camp correlation (the contractor was able to estimate this for the change between baseline and the first follow-up).

The contractor then used simulations to estimate the variance of the HT estimator under the proposed sampling design, varying the number of camps selected per awardee/Center proportionate to program size. The contractor expected the variance of the HT estimator would be larger than the variance of the sample mean of a random sample of the same size because (1) there will be some degree of within-SoI-camp correlation in outcomes; (2) sampling weights are not uniform, in part due to sampling based on camp to size; and (3) the presence of missing data. Regarding the latter, the contractor expects 85% completion rate at baseline, and then 30% attrition at each of the follow-up occasions. Thus, the contractor expects about 42% of the sample to have complete data over the three survey administrations, requiring nonresponse adjustments to the sampling weights of the complete cases. The simulations modeled the sampling of camps with the probability of selection based on camp size within awardees and Centers and attrition between baseline and the second follow-up, averaging the results from a large number of simulation iterations to get the approximate variance of the HT estimator under the sampling plan. Based on summer 2011 awardee data, the contractor knew that the average classroom size was 21 and that summer camps averaged about 170 students. Since each awardee and Center are expected to have 2,500 and 1,500 students in summer 2012, respectively, this means they will average about 15 camps per awardee and 9 camps per Center. The contractor did not have dependable data on the distribution of camp sizes within awardees or Centers, so the distribution was assumed unimodal, with a few large and small camps. Regarding the attrition mechanism, the contractor posited three weighting classes of equal size, with 32%, 42%, and 52% complete data rates, respectively. The contractor also included a small degree of confounding, in which students in the third weighting class averaged slightly larger outcomes than students in the other weighting classes. Based on these simulations the estimated range of design effect for awardees was 4.49-5.47 and was 3.15-4.61 for Centers.

The results of the power calculations, including adjustments for response rates, attrition, sampling based on camp size and the design effect indicate that the contractor will have to sample a total of 3,190 students at awardees and a total of 3,010 students at NASA Centers. Thus, the contractor will sample a total of 6,200 students.

Sampling Method

A stratified single-stage cluster design will be used to select a representative sample of SoI students. Awardees (or, separately, Centers) are the strata and camps are the clusters within strata. First, the contractor will select a systematic sample of camps at each awardee/Center. Systematic sampling after sorting by size increases the likelihood of having a wide distribution of camp sizes in the selected sample. Within camps, all students with parental consent will be administered a survey. The resultant sample of students will be clustered (or nested) within sampled camps. This strategy will be utilized for each awardee/Center to enable separate analyses for these subpopulations.

Camps will be the primary sampling unit because the FY2011 national evaluation documented the obstacles to sampling at the classroom level. Awardees may only be a few weeks or days between when they determine the number of SoI classes and the start of the SoI program, impeding the ability of the national evaluator to implement a sampling plan at the classroom level.

Each of the ten NASA Centers and eight national awardees will be included in the student survey samples, however the sample for the Centers will be analyzed separately from the national awardees given the different programming requirements. The number of students to be sampled within each awardee/Center will be determined using allocation proportional to the number of students within each awardee/Center (note that although each awardee/Center is expected to reach a minimum number of students, the contractor expects that each awardee/Center will vary in the number of students they reach). As such, the number of students needed from each awardee/Center (SS_i) can be calculated as (acknowledging that these targets may have to be adjusted once the contractor obtains the number of students enrolled at each awardee/Center):

$$SS_i = SS * \frac{n_i}{N} ; \text{ where}$$

SS is the required sample size across all awardees (SS=3,190) and all Centers (SS=3,010);

n_i is the total number of students engaged in SoI activities at the i^{th} awardee/Center; and
N is the total number of students engaged in SoI activities across all awardees/Centers.

For example, the required sample size across all national awardees is 3,190. If national awardees have enrolled a total of 20,000 students and National Awardee A has enrolled 3,100 students, using the formula above, National Awardee A would have to sample about 500 students. If National Awardee A has 4 camps each with 250 students, the contractor would select 2 camps for a total of 500 students.

Sampling Weights

To produce population-based estimates, each responding student will be assigned a sampling weight. The sampling weight is a combination of a base weight and an adjustment for student non-response to the survey. The base weight is the inverse of the probability of selection of the responding student. The probability of selecting a student is the probability of selecting the camp in which the student is located in an awardee/Center site, making the overall base weight the camp weight. The weights of responding students in a camp are adjusted to account for students who belong to that camp but do not respond. The non-response-adjusted weights are used for producing estimates and for all statistical analyses.

As the contractors are surveying the universe of teachers and expect a high response rate (as noted in Section B.3), no weights are needed to produce population-based teacher estimates.

B.2.2 Analytic Approach

Implementation Data

Analysis of the implementation forms will be descriptive, using counts, ranges, frequencies, means, and standard deviations. The implementation data will allow us to explore how summer activities were implemented and how strategies were similar or different between awardees. Further, implementation data will be used to explore associations with survey outcomes and to generate hypotheses.

Parent Survey Data

Given the descriptive nature of the information to be collected from parent, the use of simple descriptive statistics, such as counts, ranges, and frequency, in conjunction with content analytic methods, is most appropriate for these data sources in this evaluation.

Student Descriptive Analyses: Single Time Point

When the appropriate weights are used, our sampling design allows for the calculation of representative, cross-sectional averages of survey data at the student level across all awardees/Centers and at the awardee/Center level.

Student Outcomes Across All Awardees/Centers

To make statements about, “the percent of students that..,” the contractor will design our analysis such that the interpretation of “percent of students” corresponds to the percent of students out of all SoI students, not just the students that happen to be in the sample. In order to calculate statistics that are representative of all SoI students, the sampling design must be taken into account. The calculation algorithm is below. Note that if the survey item is dichotomous (0/1), then the process described below to estimate a mean actually results in the estimation of a proportion. Multiplying the proportion by 100 will give a percentage.

Let:

- y_{ij} be the response on a survey item for student j in camp i ,
- w_{ij} = the sampling weights for student j in camp i across all awardees/Centers, adjusted for non-response
- $\bar{P}$ = the estimator of the population percentage,
- $\bar{Y}$ = the estimator of the population mean,
- $\hat{Y}$ = the estimator of the population total,
- $\hat{M}$ = the estimator of the number of elements (students) in the population,
- i = 1, ..., I enumerate the camps,
- j = 1, ..., n_j , enumerate the sampled students in camp i

Then:

$$\hat{Y} = \sum_{i=1}^I \sum_{j=1}^{n_j} w_{ij} y_{ij},$$

$$\hat{M} = \sum_{i=1}^I \sum_{j=1}^{n_j} w_{ij},$$

$$\bar{Y} = \hat{Y} / \hat{M},$$

$$\bar{P} = 100 * \bar{Y}$$

Student Data at the Awardee/Center Level

In the event that NASA would like to make a statement about, “the percent of students at Awardee A that...,” the contractor will design our analysis such that the interpretation of “percent of students at Awardee A” corresponds to the percent of students out of all SoI students at the awardee/Center, not just the students that happen to be in the sample. In order to calculate statistics that are representative of all students at a particular awardee/Center, the contractor will apply the same calculation algorithm described above, but adjust the weight to reflect all students at a particular awardee/Center rather than all students across awardees/Centers.

Teacher Descriptive Analyses: Single Time Point

Teacher Data Across and Within Awardees

Because the universe of teachers will be sampled, to make a statement about, “the percent of teachers that,” or “the percent of teachers within an awardee that...,” the descriptive statistics for a single point in time do not need to be adjusted for a sampling design. Means and standard deviations will be used to describe central tendency and variation for survey items using continuous scales. Frequency distributions and percentages will be used to summarize answers given on ordinal scales. Descriptive analyses about all awardees will be conducted on the full teacher sample, while descriptive analyses about teachers within particular awardees will be restricted only to respondents from that awardee.

Statistical Software for Calculating Parameter Estimates and Standard Errors

For the student analyses, the estimator of the population mean can be easily calculated in statistical software packages that are designed for analysis of complex survey data including the estimation of mean and variance (e.g. SAS, SUDDAN). The contractor can use the variance estimates to produce standard errors and 95% confidence intervals around the estimates of the population means for the student level data. The teacher descriptive statistics can also be easily calculated in statistical software packages (e.g., SAS, SUDDAN).

Student Descriptive Analyses: Change Over Time Analyses

By “change over time analyses,” NASA means simple descriptions of change in a variable over time. This is distinct from a model where the contractor tries to assess the relationship between some predictor variable(s) and the change in the outcome variable over time.

Difference in Proportions over Time for Student Sample

The contractor plans on testing whether the difference in proportions between two time points is zero. The null and alternative hypotheses are specified as:

$$H_0: \pi_2 - \pi_1 = 0 \quad \text{vs} \quad H_a: \pi_2 - \pi_1 \neq 0$$

where π_1 and π_2 are population proportions, and p_1 is an estimate of the proportion obtained from the first sample p_2 is the proportion in the second.

In addition to reporting the estimated difference between two population proportions from two points in time, the precision of the estimate needs to be reported. This is done by computing the standard error of the estimated difference. The simple case has two independent samples at the two time points, thus, the variance of the difference between the two sample proportions is simply the sum of the variance of the first proportion and the variance of the second proportion.

The SoI study, however, is not an example of the simple case. Rather, the same population is surveyed at both time points. Therefore, the variance of the difference is no longer a sum of the variances; rather, it also includes a covariance. The variance of the difference is smaller than if there were two independent samples. The degree to which the variance in the estimate for the overlapping samples is reduced depends on the amount of overlap and the correlation coefficient between the estimates at two time periods. Kish (1965) provides a formula that computes the variance of the difference taking into account the amount of overlap and the correlation between two time periods as described below.

Let p_1 denote the estimated proportion from the baseline sample, with sample of size n_1 . Let p_2 denote the proportion from the follow-up sample, with sample of size n_2 . Let m denote the amount of overlap between the two samples. Of interest is identifying the standard error of the difference between the two sample proportions. The estimated variance of the difference between the two sample proportions can be written as:

$$v(p_1 - p_2) = v(p_1) + v(p_2) - 2m \text{cov}(p_1, p_2) \quad 2$$

where, $v(p_1)$ is the estimated variance of the baseline proportion based on a sample of n_1 units and $v(p_2)$ is the estimated variance of the post proportion based on n_2 units.

Under simple random sampling, the estimated variance of the difference in two sample proportions becomes

$$v(p_1 - p_2) = \frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2} - 2m \frac{p_{12} - p_1 p_2}{n_1 n_2}$$

where p_{12} is the proportion having the attribute in both the samples based on a sample of m units.

The student-level data from the two samples must be merged to calculate the quantity p_{12} .

For estimating the variance under the complex design used in SoI and proposed for the current study, the contractor can first estimate the variance under simple random sampling using the formula given above but with weighted proportions. Then, the contractor multiplies the variance by the design effect.

² Note that m may be close to 1 as samples may be fully overlapping.

To implement this method the contractor will obtain the variances under the complex design for the two samples using SAS as the values for $v(p_1)$ and $v(p_2)$, and estimate the covariance as

$$\text{cov}(p_1, p_2) = \rho_{p_1 p_2} \sqrt{v(p_1) v(p_2)},$$

where the correlation term $\rho_{p_1 p_2}$ is calculated as the correlation between baseline and follow-up survey measurements for the students that were measured at both time points.

The square root of the variance gives the standard error of the difference in the two proportions recognizing that the samples have overlap and thus that the samples are not independent. The standard error will be used in a statistical test of the null hypothesis of equivalent proportions in the two groups. The specification of the hypothesis test is,

$$H_o: p_2 - p_1 = 0 \text{ vs } H_a: p_2 - p_1 \neq 0$$

and procedure for determining the test statistic is:

$$z = \frac{(p_2 - p_1) - 0}{SE(p_2 - p_1)}$$

If the observed value of z as calculated above is greater than the critical value for $\alpha=0.05$ from the z-distribution (1.96), the null hypothesis will be rejected at the $p<0.05$ level.

Difference in Means over Time for Student Sample

If the goal of the calculations is to perform a test of whether the difference between two means from two points of data collection is equal to zero then the challenging part of conducting the test is again calculating the variance of the difference of the means.

In order to calculate the variance (and standard error) of the difference (Kish, 1965), let $\bar{x}_1$ denote the estimated mean from the first sample of size n_1 . Let $\bar{x}_2$ denote the estimated mean from the second sample of size n_2 . We are interested in testing the difference between the two sample means. The estimated variance of the difference between the two sample means can be written as:

$$v(\bar{x}_1 - \bar{x}_2) = v(\bar{x}_1) + v(\bar{x}_2) - 2\text{cov}(\bar{x}_1, \bar{x}_2)$$

Under our sampling design, the variance of the difference can be written as

$$v(\bar{x}_1 - \bar{x}_2) = v(\bar{x}_1) + v(\bar{x}_2) - \frac{2\rho_{x_1 x_2} m \sqrt{v(\bar{x}_1) v(\bar{x}_2)}}{\sqrt{n_1 n_2}},$$

where $v(\bar{x}_1)$ is the estimated variance of the first mean based on a sample of n_1 units, $v(\bar{x}_2)$ is the estimated variance of the second proportion based on n_2 units and m is the amount of overlap

between the two samples. The correlation ($\rho_{x_1x_2}$) is estimated based on the overlap. SAS can calculate the variance under our sampling design for the first mean and for the second mean.

The square root of the variance gives the standard error of the difference in the two means, which can be used in a statistical test recognizing that the samples overlap and are not independent. The specification of the hypothesis test is,

$$H_o: \bar{x}_2 - \bar{x}_1 = 0 \quad \text{vs} \quad H_a: \bar{x}_2 - \bar{x}_1 \neq 0$$

and procedure for determining the test statistic is:

$$t = \frac{(\bar{x}_2 - \bar{x}_1) - 0}{SE(\bar{x}_2 - \bar{x}_1)}$$

If the observed value of t as calculated above is greater than the critical value from the t-distribution with $n - 2$ degrees of freedom and $\alpha = 0.05$, the null hypothesis will be rejected at the $p < 0.05$ level.

Teacher Descriptive Analyses: Change Over Time Analyses

Similar to the student analyses, the contractor will conduct simple descriptions of change in a variable over time for classroom teacher outcomes. Again, this is distinct from a model where the goal is to try to assess the relationship between some predictor variable(s) and the change in the outcome variable over time.

Difference in Teacher Outcomes over Time

The contractor plans on testing whether the difference in proportions and/or means for teachers between two time points is zero. To do so, the contractor will use a McNemar test or paired t-test, depending on the distribution of the outcome variables.

B.3 Methods to Maximize Response Rates

Response rates for parent consent, student surveys, and teacher surveys were all low for the national evaluation of summer 2011.³ Multiple factors were involved in preventing the return of the materials, including the brief amount of time available for planning between the award announcement and program implementation, delayed access to SoI funding, and lack of clarity and prescriptiveness regarding evaluation responsibilities and requirements. NASA and its contractor are taking the following steps to increase the response rates for summer 2012:

Steps taken during FY2011 data collection:

- Providing opportunities during the summer camp sessions to complete the student survey at the beginning and end of the summer activities;
- Conducting a webinar for national evaluation coordinators, PIs and NASA Center education leads when the evaluation materials are distributed to review data collection processes,

³ For example, the overall parent consent response rate was 60 percent for national awardees and 15 percent for the NASA Centers.

- reiterate grant requirements regarding the evaluation, and emphasize the importance of collecting the baseline survey before the start of SoI programming;
- Requiring awardees to include the teacher baseline survey as part of their registration materials;
- Providing a toll-free number that participants can call to ask questions and verify the legitimacy of the evaluation; and,
- Sending up to three email reminders and making up to three follow-up calls to encourage teachers to fill out the surveys at home or wherever they have internet access.

Additional steps to be taken during the FY2012 data collection:

- Providing funding to awardees and Centers no later than February, 2012;
- Updating the SoI website to include comprehensive evaluation information and materials for increased accessibility;
- Reviewing of the evaluation activities and progress toward data collection goals at monthly meetings between NASA and awardee PIs;
- Providing both paper and online versions of adult surveys and forms (i.e., parent consent, teacher baseline and follow-up surveys);
- Distributing the parent consent forms to awardees and NASA Centers in February so that they can include them in the registration materials;
- Sending email reminders about the parent survey to those parents that did not complete the parent survey. Email information could be collected from contact information gathered by awardees and Centers.

With these revised operations, and given that students will have opportunities to fill out the baseline surveys during summer programming, the contractor expects to achieve a response rate of 85% or higher for the baseline surveys and parent consent forms. The contractor assumes that some attrition will occur, particularly if a third survey is mailed to students after the school-year activities. Given that attrition rates for the FY2011 student surveys were about 30% between baseline and summer follow-up, the contractor expects a response rate of about 60% for the summer follow-up survey and about 42% for the school-year follow-up survey.

For teachers, the contractor expects to achieve response rates of 85% or higher for the baseline survey. The contractor assumes that some attrition will occur between the baseline and summer follow-up survey and between the summer follow-up survey and the school-year follow-up survey. Based on the pilot teacher survey attrition rates of 10% between survey waves, the contractor would expect a response rate of about 77% at summer follow-up and about 69% at school-year follow-up.

Non-Response Bias

Nonresponse may be a problem in our analyses if it introduces bias into our population estimates. Bias occurs if the students or teachers that refuse to participate or leave the study would give systematically different responses to the survey (had they responded to it) than the students or teachers who complete the surveys. Poor response rates do not guarantee a biased estimate, as the decision to not participate or leave the study could be completely unrelated to survey answers.

In general, the effects of potential non-response bias cause little concern if the non-response rate is less than 20 percent; accordingly, the contractor will conduct a nonresponse bias analysis that is described below, if our response rate is less than 80 percent.

Student Non-Response

The contractor will construct a propensity model to estimate the probability of a student responding to the survey (propensity score) both for responding and non-responding students. These propensity scores are estimated by a logistic regression model that will use demographic variables (e.g., gender, grade level, race, ethnicity) collected on the original parent consent form/survey that will be available both for non-responding and responding students. The contractor will then group students using the estimated propensity scores and examine the demographic characteristics of responding and non-responding students within each group. This grouping will provide a method of forming weighting classes to adjust the weights of responding students and reduce nonresponse bias.

Teacher Non-Response

In 2011, an appreciable number of teachers were not asked to complete the baseline survey prior to the start of their professional development (PD) activities because the survey had not yet been approved, and this contributed to low baseline survey response rates. This year, the contractor expects approval to be received well in advance of any summer PD activities, thus allowing all SoI teachers to be invited to complete the baseline survey. Because baseline surveys will be distributed with registration forms and these forms must be completed, the contractor expects teachers will tend to see the survey as part of the registration process and this will result in high response rates for the surveys.

However, teacher response may be a concern during the FY2012 evaluation. To account for potential nonresponse if the teacher response rate is below 80%, the each time a teacher fills out a survey (whether it be the baseline, the summer follow-up or the school-year follow-up) the teacher will be asked questions about their demographics. In doing this, the only teachers for whom there will not be demographic data are teachers who responded to none of the surveys. The contractor could then perform t-tests and chi square tests to compare baseline responders and non-responders on demographic variables as well as follow-up responders to non-responders.

To develop adjustments to statistical procedures to account for teacher non-response, the contractor will alter the models to include weights that compensate for the missing data from non-responders. These weights will be derived from estimates of propensity scores, defined as the probability of being a complete case (i.e. a responder to multiple survey waves) given a responder's demographic characteristics. The estimates will be derived from a logistic regression predicting whether or not a teacher is a complete responder based on her demographic variable values.⁴ For teachers who responded to one but not all of the survey waves (i.e. partial responders), estimated probabilities can be obtained from the logistic regression, and multiplying these estimated probabilities by one minus the proportion of complete non-responders gives estimates of the propensity scores. Weights derived from these propensity score estimates can be used to prevent biased data analyses if (i) data from complete non-responders is missing "completely at random," (ii) non-response on a single survey only is missing "at random" with respect to the demographic variables, and (iii) the logistic regression model is correct. While, in practice, it is unlikely that these assumptions strictly hold, if the complete non-response rate is low, then they are sufficiently plausible that weights based on them will have some value in limiting bias due to non-response.

⁴ Note that teachers that do not respond to any waves of the survey (i.e. complete non-responders) are necessarily excluded from this analysis, as their demographic data are never obtained.

Under these assumptions, weights equal to the reciprocals of the estimated propensity scores can be used in complete case data analyses to produce approximately unbiased results; e.g., performing weighted t-tests on continuous outcomes. However, the presence of observations with large weights (i.e., reciprocals of very small propensity scores) may result in estimates with high variability. It is therefore often useful to “trade off” some bias for a lessening of variance by developing weighting classes based on the estimated propensity scores of complete cases and of teachers who only responded to one survey. All of these teachers are sorted by their estimated propensity scores, and the sorted list is partitioned into quintiles. Each quintile constitutes a weighting class, and all teachers in a weighting class are assigned the same weight, namely, the reciprocal of the proportion of complete cases in the weighting class. Again, this approach relies on the proportion of complete non-responders being small.

B.4 Test of Procedures

Estimates for the parent consent form are based on time requirements from similar surveys conducted on comparable evaluations.

Survey development and procedures were tested and refined as follows. The 2010 pilot surveys were fielded in summer 2010, revised in fall 2010, and updated in winter 2010 to measure outcomes of interest in FY2011. For the student surveys, existing instruments with established psychometric characteristics were selected after an extensive literature review (e.g., Modified Attitudes Towards Science Inventory (mATSI), Weinburgh and Steele, 2000; Test of Science Related Attitudes (TOSRA), Fraser, 1981; and the Math and Science Interest Survey, Hulett, Williams, Twitty, Turner, Salamo, and Hobson, 2004). The student surveys were piloted with seven middle school students in spring 2011 to refine the language and estimate time for completion. Given that the minor changes made to the FY2012 student survey were all in response to feedback from the FY2011 administration, no testing was completed on the FY2012 survey.

Key questions on the teacher survey were designed specifically for the SoI evaluation and constructed to capture outcomes from the program’s logic model. Teacher background questions were taken from existing, nationally fielded instruments (see Appendix 10). The surveys, including the newly constructed questions, were piloted with six teachers in 2011 to test and refine the language and estimate time for completion. Given the similarities between the FY2011 and FY2012 surveys and that fact that the minor changes were made to the FY2012 survey (in response to teacher feedback received during the FY2011 administration, the FY2012 teacher surveys was not re-tested. Experts in the field reviewed draft and final instruments for content validity and clarity, including Marian Pasquale, Senior Research Scientist at the Education Development Center (EDC) and a former middle school teacher, with expertise in middle school curriculum development, technology implementation, and student learning.

All FY2012 revisions were made based on awardee and NASA Center feedback and careful discussion between NASA and the contractor. Finally, NASA Office of Education staff reviewed the instruments for final approval.

B.5 Individuals Consulted on Statistical Aspects of Design

The plans for statistical analyses for this study were primarily developed by Abt Associates, Inc. and the Education Development Center (EDC). The team is led by Hilary Rhodes, Project Director; Ricky Takai,

Principal Investigator; Alina Martinez, Principal Associate; Amanda Parsad, Project Quality Advisor; Kristen Neishi, Deputy Project Director; Ed Bein, Psychometrician; Melissa Velez, Task Leader; and Tamara Linkow, Task Leader, all of Abt Associates, Inc. The surveys were refined by Jacqueline DeLisi, Abigail Levy, and Yueming Jia at EDC. Contact information for these individuals is provided on the next page. Additionally, Laura LoGerfo, the Project Officer for High School Longitudinal Study of 2009 at the U.S. Department of Education National Center for Education Statistics, provided feedback on the parent, student, and teacher survey instruments.

Abt Associates, Inc.			
Hilary Rhodes	Project Director	617-520-3516	hilary_rhodes@abtassoc.com
Alina Martinez	Principal Associate	617-349-2312	alina_martinez@abtassoc.com
Ricky Takai	Principal Investigator	301-634-1765	ricky_takai@abtassoc.com
Amanda Parsad	Project Quality Advisor	301-634-1791	amanda_parsad@abtassoc.com
Kristen Neishi	Deputy Project Director	301-634-1759	kristen_neishi@abtassoc.com
Melissa Velez	Task Leader	617- 520-2875	melissa_velez@abtassoc.com
Ed Bein	Psychometrician	617-520-3029	ed_bein@abtassoc.com
Tamara Linkow	Task Leader	617-520-2978	tamara_linkow@abtassoc.com
Education Development Center			
Jacqueline DeLisi	Survey Task Leader	617-969-5979	jdelisi@edc.org
Abigail Levy	Survey Developer	617-969-5979	alevy@edc.org
Yueming Jia	Survey Developer	617-969-5979	yjia@edc.org

References

- Fraser, B. J. (1981). *TOSRA test of science related attitudes handbook*. Hawthorn, Victoria, Australia: Australia Council for Educational Research.
- Hulett, L.D., Williams, T.L. Twitty, L.L., et al. (2004). *Inquiry-based Classrooms and Middle School Student Perceptions about Science and Math*. Paper presented at the 2004 Annual Meeting of the American Educational Research Association, San Diego, CA.
- Kish, L. (1965). *Survey Sampling*. New York: John Wiley & Sons, Inc.
- Martinez, A. & Consentino de Cohen, C. (March 31, 2010). *The National Evaluation of NASA's Science, Engineering, Mathematics and Aerospace Academy (SEMAA) Program*. Cambridge, MA: Abt Associates, Inc.
- Weinburgh, M..H., & Steele, D.
(2000). The Modified Attitudes Toward Science Inventory: Developing an instrument to be used with fifth grade urban students. *Journal of Women and Minorities in Science and Engineering* 6, 87-94.