

<b>Section</b>	<b>Page</b>
<b>B. COLLECTION OF INFORMATION EMPLOYING STATISTICAL METHODS .....</b>	<b>1</b>
1. Respondent Universe and Sampling Methods .....	1
2. Procedures for Collecting Information .....	3
3. Methods to Maximize Response .....	7
4. Testing of Procedures .....	9
5. Contacts for Statistical Aspects of Data Collection .....	10
<b>References .....</b>	<b>11</b>
<b>APPENDIX A: Descriptions of Methodological Testing Techniques .....</b>	<b>13</b>
<b>APPENDIX B: Privacy Policies and Procedures .....</b>	<b>16</b>
<b>List of Tables .....</b>	<b>18</b>

**GENERIC CLEARANCE FOR THE NASA  
OFFICE OF EDUCATION/PERFORMANCE MEASUREMENT AND EVALUATION  
(TESTING) SUPPORTING STATEMENT**

**B. COLLECTION OF INFORMATION EMPLOYING STATISTICAL  
METHODS**

**1. RESPONDENT UNIVERSE AND SAMPLING METHODS**

*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.*

***Respondent Universe***

The respondent universe for NASA Education methodological testing consists of individuals who either participate in NASA Education activities or are staff managing educational activities (both at NASA and funded through NASA grants, cooperative agreements, and contracts). It is difficult to anticipate and define all the types of potential respondents under this generic clearance beyond the most immediate needs for this generic clearance, but below are descriptions of the individuals who could represent the respondent universe in this generic submission:

- Undergraduate and graduate students participating in NASA-funded internships, scholarships, and fellowships;
- P-12 and informal educators and higher education faculty participating in NASA-funded educator professional development;
- NASA civil servants who manage projects and activities; *and*
- Primary investigators and managers of NASA-funded grants, cooperative agreements, and contracts.

Respondent categories with a corresponding estimate for each Potential Respondent Universe (N) anticipated for this generic clearance can be found below (See Table 3). Expected Response Rate is defined as the past rate of response observed in NASA Education for that particular Respondent Category. For instance, precollege, undergraduate, graduate, and post-graduate students are especially responsive to NASA Education requests for information because at the onset of establishing a relationship with NASA, they are not eligible to apply for NASA internships, fellowships, and scholarships (NIFS) with incomplete information. For the reason that these respondents are aware of the potential to obtain multiple opportunities after having been awarded a first time, these respondents willingly partner with NASA Education to maintain current contact information in order to access current information pertaining to NIFS opportunities. OEID's newest IT applications will avail participants of the opportunity to update contact information in a way that is less burdensome, through automated delivery of links to

OEID’s databases wherein opening the link will take the participant directly to a log in screen appropriate to her or his project activity or program.

Further, for these categories of respondents who are affectively characterized as highly motivated individuals, they understand the value of submitting feedback to optimize future NIFS opportunities they may be awarded. Therefore, they tend to be highly motivated to cooperate with NASA Education requests for information at a rate of 60%. The same can be said for educator participants who must complete information in our systems in order to partake of professional development opportunities.

External program managers are required to submit information to our online data collection systems and therefore it is not difficult to leverage Center points of contact to obtain data submitted in a timely fashion. Therefore, 100% compliance with a request for information, even in the form of participation in data collection instrumentation, is a reasonable expectation. Note that some testing methods (e.g., focus groups, cognitive interviews) require nine participants or less. These numbers are not reflected below. Data collection through focus groups and cognitive interviews for testing purposes will not be used to generalize results, but rather for preliminary item and instrument development, and piloting only<sup>1</sup>. Table 1 below reflects potential respondent universe, expected response rates, and statistically adjusted number of respondents for each respondent category:

**Table 1: Respondent Universe and Relevant Numbers**

	<b>Respondent Category</b>	<b>Potential Respondent Universe (N)</b>	<b>Expected Response Rate (R)</b>	<b>Statistically Adjusted Number of Respondents (n)</b>
<b>Office of Education Performance Measurement System</b>	Undergraduate and graduate student profiles	22,435	0.6	629
	Educator participant surveys	183,040	0.6	639
	External program manager- Data collection screens	844	1.0	264
<b>One Stop Shopping Initiative</b>	Pre-College surveys <sup>2</sup>	1,615	0.6	517
	Undergraduate surveys	10,486	0.6	618
	Graduate surveys	870	0.6	444
	Post-Graduate surveys	241	0.6	247
				<b>3,358</b>

<sup>1</sup> Further description of methodological testing techniques can be found in Appendix A.

<sup>2</sup>In this instance, the category “pre-college” refers to students who are over the age of consent, but have not formally enrolled in a college or university. As such, this group of students applies for opportunities associated with college preparation as a means to become more competitive for enrollment in college or as a means to explore potential STEM majors prior to enrolling in college or university.

## ***Sampling Methods***

### **Systematic Random Sampling**

For each Respondent Category, for the purposes of piloting instruments, technology support to the Office of Education will systematically randomly generate a list in length corresponding to  $n$  Statistically Adjusted Number of Respondents wherein every  $n$ th element from the population list will be selected (Hesse-Biber, 2010, p. 50.). This process attempts to create a sampling frame that closely approximates in characteristics pertinent to the Respondent Universe for each data collection instrument.

### **Nonprobability Purposive Sampling**

For the purposes of focus groups and cognitive interviews, nonprobability purposive sampling will be used wherein the research purpose determines the type of elements or respondents selected in the sample. This sampling strategy gathers a collection of specific informants deemed likely to exemplify patterns of behavior or characteristics reflective of the Respondent Universe from which they are drawn necessary for the purposes specific to a particular data collection instrument under development (Hesse-Biber, 2010, p. 126). Even in the event that a focus group or cognitive interview fails to yield persuasive results, OEID will not interview a participant more than once. Instead, OEID will recruit an entirely new focus group or set of participants for cognitive interviews. Obtaining statistical rigor later on in the process begins by avoiding introduction of confounding variables in the preliminary stages of instrument design. Interviewing a participant twice in a cognitive interview or including her or him in a new focus group may be a source of confounding variables and should be entirely avoided.

## **2. PROCEDURES FOR COLLECTING INFORMATION**

*Describe the procedures for the collection of information including:*

*\* Statistical methodology for stratification and sample selection:*

Not applicable. For the purposes of this data collection instrument development, NASA Education has no need for instrumentation specific to subgroups within any of the Respondent Universe categories of interest.

*\* Estimation procedure:*

For the reason that NASA Education has experienced poor survey response rates in some Respondent Categories pertinent to this clearance package, the number of respondents to reach in order to obtain a statistically significant response is based on the following criteria:

Where:

$n$  = statistically adjusted number of respondents or number of respondents in required in final sample size

**N** = potential respondent universe (number of people in the population)

**P** = estimated variance in respondent universe (population), as a decimal: (0.5 for 50-50, 0.3 for 70-30)

**A** = Precision desired, expressed as a decimal (i.e., 0.03, 0.05, 0.1 for 3%, 5%, 10%)

**Z** = Based on confidence level: 1.96 for 95% confidence, 1.6449 for 90% and 2.5758 for 99%

**R** = Expected (estimated) response rate, as a decimal

Thus, utilizing those criteria, Yamane (1973) and Blalock (1972) offer this equation for determining the statistically adjusted number of respondents for the final sample size:

$$n = \left[ \frac{\frac{A^2}{Z^2} + \frac{P[1-P]}{N}}{R} \right]$$

Steps in Selecting a Sample Size:

1. **Estimating N, potential respondent universe (population size):** For the instance of NASA Education project activity participants, we will use prior trends of participation to estimate N.
2. **Determining A, the desired precision of results:** The level of precision is the closeness with which the sample predicts where the true values in the population lie. The difference between the sample and the real population is called the sampling error or margin of error. The level of precision accepted depends on balancing accuracy and resources. High levels of precision require larger sample sizes and thus higher costs to achieve those samples, but a high margin of error can produce meaningless results. For social science application in general, an acceptable margin of error or precision level is between 3% and 10%. For the purpose of this first phase of field testing, 5% is an acceptable margin of error. In the future, given greater availability of funds for data collection instrument development, it would be ideal to integrate a more stringent 3% margin of error into determining sample size for the next phase of statistical testing as OEID continues to monitor and maintain the psychometric properties of NASA Education instruments.
3. **Determining Z, confidence level:** Confidence level reflects the risk associated with accepting that the sample is within the normal distribution of the population. A higher confidence levels require a larger sample size, but avoids statistically insignificant results

associated with a low confidence level. For this social science application, a 95% confidence level has been adopted for these purposes.

4. **Estimating P, the degree of variability:** Variability is the degree to which the attributes or concepts being measured in the questions are distributed throughout the population sampled. The higher the degree of variability the larger the sample size must be to represent the concept or attribute within the sample. For the instances of this social science application located within the context of STEM education activities, we will assume moderate heterogeneity and estimate variability at 50%.
5. **Estimating R, expected response rate:** Base sample size is the smallest number of responses required for statistically meaningful results. Calculation of sample size must overcome non-response and should also consider a guesstimate at what a response rate might be or it can consider response rates experienced with the population of interest. NASA Education response rates to survey and data collection instrumentation have been very low in some instances. Response rates vary between 20% and 60% with the exception of program managers who are required to enter data and thus have a response rate of 100%. Regardless of this difference in response rates within our community, characteristics of respondents may differ significantly from non-responders. For this reason, follow-up samples of the corresponding non-respondent population may be undertaken to determine differences, if any exist.

For the purposes of large-scale statistical testing, consideration of the aforementioned variables within the context of this methodological testing package to ensure that the collection of responses statistically resembles each Respondent Universe (data collection source) results in the following Table 2:

**Table 2: Respondent Universe and Sampling Calculations**

	<b>Data Collection Sources</b>	<b>(N)</b>	<b>(A)</b>	<b>(Z<sup>2</sup>)</b>	<b>(P)</b>	<b>Base sample size</b>	<b>(R)</b>	<b>(n)</b>
<b>Office of Education Performance Measurement System</b>	Undergraduate and graduate student profile	22,435	0.0025	3.8416	0.5	378	0.6	629
	Educator participant surveys	183,040	0.0025	3.8416	0.5	384	0.6	639
	External program manager- Data collection screens	844	0.0025	3.8416	0.5	267	1	264
<b>One Stop Shopping Initiative</b>	Pre-College surveys <sup>3</sup>	1,615	0.0025	3.8416	0.5	310	0.6	517
	Undergraduate surveys	10,486	0.0025	3.8416	0.5	371	0.6	618
	Graduate surveys	870	0.0025	3.8416	0.5	267	0.6	444
	Post-Graduate surveys	241	0.0025	3.8416	0.5	148	0.6	247
								<b>3,358</b>

- Information collected under the purview of this clearance will be maintained in accordance with the Privacy Act of 1974, the e-Government act of 2002, the Federal Records Act, and as applicable, the Freedom of Information Act in order to protect respondents’ privacy and the confidentiality of the data collected<sup>4</sup>. Further information on data security is provided in Appendix B.

*\* Degree of accuracy needed for the purpose described in the justification,*

NASA Education project activities target STEM-related activities. Hence, instrumentation and the sample with which data collection instrumentation is tested must correspond with a high degree of accuracy. Moreover, because data from these instruments is used to inform policy, a high degree of accuracy must be integrated throughout the entire data collection instrument process.

*\* Unusual problems requiring specialized sampling procedures,*

Not applicable. NASA OEID does not foresee any unusual problems with executing pilot or large-scale statistical testing via the procedures described.

*\* Any use of periodic (less frequent than annual) data collection cycles to reduce burden.*

<sup>3</sup>Again, in this instance, the category “pre-college” refers to students who are over the age of consent, but have not formally enrolled in a college or university. As such, this group of students applies for opportunities associated with college preparation as a means to become more competitive for enrollment in college or as a means to explore potential STEM majors prior to enrolling in college or university.

<sup>4</sup> [http://www.nasa.gov/privacy/nasa\\_sorn\\_10EDUA.html](http://www.nasa.gov/privacy/nasa_sorn_10EDUA.html)

- Since this information collection request applies to methodological testing activities, data collection activities will occur as needed to gather statistically significant data to appropriately determine the validity and reliability characteristics of instruments, where applicable, and the psychometric properties of instrumentation, where applicable.
- Rigorously tested data collection instrumentation is a requirement for accurate performance reporting. If these testing activities are not conducted, NASA will not be able to conduct basic program office functions such as strategic planning and management.
- Without the timely and complete set of planning, execution, and outcome (survey) data collected by valid and reliable instruments in both OSSI and OEPM, NASA Education will be unable to assess program effectiveness, meet federal and agency reporting requirements, or make data informed management decisions.
- Less timely and complete information will adversely affect the quality and reliability of the above-mentioned endeavors. The degradation of any single component of our data collection would jeopardize the integrity and value of the entire OEID suite of applications and the integrity of our databases.

### **3. METHODS TO MAXIMIZE RESPONSE**

*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.*

Maximizing response rates and managing issues of non-response are equally relevant concerns in recruiting participants for pilot testing and routine data collection instrument administration. In that regard, OEID intends to utilize such methods to reach each targeted population to yield statistically significant data from a random sample of at least 200 respondents to determine initial reliability coefficients and validity (Komrey and Bacon, 1992; Reckase, 2000). Furthermore, the same procedures will be employed during regular data collection through OMB-approved instruments. Meaning, similar patterns of effectiveness of participant recruitment strategies and response rates are inextricably linked and any procedures for maximizing response rates, as complex as they may be, are interdependent (Barclay, Todd, Finlay, Grande, & Wyatt, 2002). Therefore, despite the wide range of data sources being recruited for study participation—undergraduate student, graduate student, or educator, for instance—the same strategies for maximizing response apply.

#### **Study Participant Recruiting**

The OEID team will use a combination of recruitment by NASA Education Center Education Directors and automatic email reminders adopted from Swail and Russo (2010) to maximize participant response rates for data collection instrument testing. Participant contact lists will be solicited from the appropriate Center Point of Contact (POC) for the respondent population sampled. Center POCs will use one month to identify respondents who agree to participate and



submit their contact information to NASA Education OEID. Bi-weekly reminders will be sent and follow-up phone calls will be made to POCs as needed.

### Participant Assignment to Study

Using random assignment, respondents will be assigned to an instrument for which their responses are appropriate with the goal of having equal numbers of participants completing instruments across testing sites and to avoid Center effects, meaning, responses to survey instruments related to a participant's Center culture.

### NASA OEID IT Infrastructure for Testing

New information technology applications, the Survey Launcher and Composite Survey Builder, are in development with the NEACC and only a few months from completion. The Survey Launcher is a communication vehicle between OEID and the NASA education community of project activity participants. The primary uses of the Survey Launcher application are to administer OMB-approved surveys and data collection instruments to project activity participants and to notify participants of the opportunity to volunteer updated contact information to OEID. Since the Survey Launcher is designed to reach several hundred project activity participants through a single emailed survey web link, this capability will be leveraged to maximize response rates during piloting and large-scale statistical testing of data collection instruments as a vehicle for instrument delivery as well as the vehicle for important reminders to study participants.

The Composite Survey Builder is the companion, content management application to the Survey Launcher in that it will provide OEID complete control over content electronically delivered to project activity participants. During the testing phases of instrument development for instance, the Composite Survey Builder will be used to electronically deliver draft instruments for testing and consent forms whereas post-testing, the application will be used to administer finalized, OMB-approved data collection instruments and notification reminders for updating contact information in OEID's data bases.

Together, the Survey Launcher and Composite Survey Builder enable OEID to maximize response rates during data collection instrument development and afterwards, when said instruments will be used to collect information from participants in NASA Education STEM project activities.

### Procedure

1. Using the Survey Launcher, survey invitations will be emailed to individual respondents. They will be informed of the purpose of the test, the estimated time to complete the data collection instrument, a short description of the questions they will be answering, and a consent form to be electronically signed and returned.
2. Respondents will be given 5 business days to respond to the invitation before an automated "gentle" reminder email is sent to those who have not responded to the initial

invitation and/or have not completed the data collection instrument. \*This process will be repeated two more times during this phase of the test, but will include an edited subject line expressing this sentiment, “If you have already responded to this request, please kindly disregard this reminder.”

3. Prior to the final “gentle” reminder email, participants will be informed that their POC will be notified about their participation status. Each POC will be contacted and provided with a list of participants who have not completed the assigned data collection instrument. The purpose of this contact is to encourage the highest level of participation from respondents. Throughout the entire testing procedure, there will be an open line of communication between participants and OEID POCs. This communication will be essential to assuring all individuals involved in this process are completely informed about the progress of field testing efforts and to resolve any issues or concerns.

OEID will employ split-half methodology to obtain reliability and validity measures on instruments and performance data on individual items while limiting burden on respondents to one testing situation. This method presents less burden on respondents than some other methods, such as test-retest methods, since the requirement for testing is for a single test administration.

If this procedure fails to yield satisfactory measures on enough items per instrument, a second round of cognitive interviews or focus groups on the failed items and then a second full administration of the instrument to a new sample of the target population may be necessary to obtain satisfactory reliability coefficients. If a second, new administration of items is necessary, the process will start again, but with an entirely new pool of potential test participants to avoid burdening any one respondent more than one time.

#### **4. TESTING OF PROCEDURES**

*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.*

This submission is in itself a request for authorization to conduct tests of data collection instruments that are in development and/or require OMB approval. The purpose of cognitive and other forms of intensive interviewing, and of the testing methods in general covered by this request, is not to obtain data, but rather to obtain information about the processes people use to answer questions as well as to identify any potential problems in the question items or instruments prior to piloting with a statistically relevant sample of respondents. In some cases, focus group and/or cognitive interview protocols will be submitted for OMB approval. In other cases where the evidence base provided by the educational measurement research literature has provided a basis for a reasonable instrument draft consistent with a program activity, the instrument draft will be submitted to OMB for approval for pilot testing. The testing procedures and methodologies to be used by NASA Office of Education and its contractors are, overall,

consistent with the educational measurement research literature evidence base and other Federal agencies engaged in STEM program performance data collection.

## **5. CONTACTS FOR STATISTICAL ASPECTS OF DATA COLLECTION**

*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.*

Valador, Inc. has availed Dr. Lisa E. Wills, subject matter expert in data collection instrument design, and contracted and dedicated full-time to the NASA Office of Education Infrastructure Division. Her areas of expertise are as follows: Quantitative, Qualitative, & Mixed Research Methods; Cognitive, Psychometric & Survey Instrument Development; Inferential & Descriptive Statistics; Big Data Analytics; Multi-level modeling; and Discourse, Narrative, & Case Study Analyses.

## References

- Barclay, S., Todd, C., Finlay, I., Grande, G., & Wyatt, P. (2002). Not another questionnaire! Maximizing the response rate, predicting non-response and assessing non-response bias in postal questionnaire studies of GPs. *Family Practice, 19*(1), 105-111.
- Blalock, H. M. (1972). *Social statistics*. New York, NY: McGraw-Hill.
- Colton, D., & Covert, R. W. (2007). *Designing and constructing instruments for social research and evaluation*. San Francisco: John Wiley and Sons, Inc.
- Costello, A. B., & Osborne, J. W. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment, Research & Evaluation, 10*(7), 1-9.
- Davidshofer, K. R., & Murphy, C. O. (2005). *Psychological testing: Principles and applications*. (6th ed.). Upper Saddle River, NJ: Pearson/Prentice Hall.
- DeMars, C. (2010). *Item response theory*. New York: Oxford University Press.
- Fabrigar, L. R., & Wegener, D. T. (2011). *Exploratory factor analysis*. New York, NY: Oxford University Press.
- Haladyna, T. M. (2004). *Developing and validating multiple-choice test items* (3rd ed.). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- Hesse-Biber, S. N. (2010). *Mixed methods research: Merging theory with practice*. New York: Guilford Press.
- Jaaskelainen, R. (2010). Think-aloud protocol. In Y. Gambier, & L. Van Doorslaer (Eds.), *Handbook of translation studies* (pp. 371-373). Philadelphia, PA: John Benjamins.
- Komrey, J. D., & Bacon, T. P. (1992). Item analysis of achievement tests based on small numbers of examinees. *Paper presented at the annual meeting of the American Educational Research Association*. San Francisco.
- Kota, K. (n.d.). Testing your web application: A quick 10-step guide. Retrieved from [http://www.adminstrack.com/articles/testing\\_web\\_apps.pdf](http://www.adminstrack.com/articles/testing_web_apps.pdf).
- Reckase, M. D. (2000). The minimum sample size needed to calibrate items using the three-parameter logistic model. *Paper presented at the annual meeting of the American Educational Research Association*. New Orleans.
- Swail, W. S., & Russo, R. (2010). *Instrument field test: Quantitative summary*. Library of Congress- Teaching with Primary Sources: Educational Policy Institute.

Wilson, M. (2005). *Constructing measures: An item response modeling approach*. New York: Psychology Press.

Yamane, T. (1973). *Statistics: An introductory analysis*. New York: Harper & Row.

## APPENDIX A: Descriptions of Methodological Testing Techniques

- Usability testing: Pertinent are the aspects of the web user interface (UI) that impact the User's experience and the accuracy and reliability of the information Users submit. The ease with which Users navigate the data collection screens and the ease at which the User accesses the actions and functionality available during the data input process are equally important. User experience is also impacted by the look and feel of the web UI and the consistency of aesthetics from page to page, including font type, size, color scheme utilized and the ways in which screen real estate is used (Kota, n.d.). The foundation for Usability testing will be a think-aloud protocol analysis as described by Jääskeläinen (2010) that exposes distractions to accurate input of data whereas a short Likert Scale survey with qualitative questions will determine the extent of distraction and nature of the distractions that impede accurate data input.
- Think-aloud protocols (commonly referred to as cognitive interviewing): This data elicitation method is also called 'concurrent verbalization', meaning subjects are asked to perform a task and to verbalize whatever comes to mind during task performance. The written transcripts of the verbalizations are referred to as think-aloud protocols (TAPs) (Jääskeläinen, 2010, p 371) and constitute the data on the cognitive processes involved in a task (Ericsson & Simon, 1984/1993). When elicited with proper care and instruction, think-aloud does not alter the course or structure of thought processes, except with a slight slowing down of the process. Although high cognitive load can hinder verbalization by occupying all available cognitive resources, that property is of no concern regarding the tasks under analysis that are restricted to information actively processed in working memory (Jääskeläinen, 2010, p. 371). For the purposes of NASA Education, think-aloud protocols will be especially useful towards the improvement of existing and developing of new data collection screens, which are different in purpose from online applications. Whereas an online application is an electronic collection of fields that one either scrolls through or submits, completed page by completed page, data collection screens represent hierarchical layers of interconnected information for which user training is required. Since user training is required for proper navigation, think-aloud protocols capture the user experience to incorporate it into a more user-friendly design and implementation of this kind of technology. Lastly, data from think-aloud protocols is used to ensure that user experiences are reliable and consistent towards collecting robust data.
- Focus group interviews: With groups of nine or less per instrument, this qualitative approach to data collection is a matter of brainstorming to creatively solve remaining problems identified after early usability testing of data collection screen and program application form instruments (Colton & Covert, 2007, p. 37). Data from this type of research will include audiotapes obtained with participant consent, meeting minutes taken

by a subject matter expert in administration assistance, and reflective comments submitted by participants after conclusion of the focus group. Focus group interviews may be used to refine items that failed initial reliability testing for the purposes of retesting. Lastly, focus group interviews may be used with participants as a basis for a grounded theory approach to instrument development or for refining an already existing instrument to be appropriate to a specific audience.

- **Comprehensibility testing:** Comprehensibility testing of program activity survey instrumentation will determine if items and instructions make sense, are ambiguous, and are understandable by those who will complete them. For example, comprehensibility testing will determine if items are complex, wordy, or incorporate discipline- or culturally-inappropriate language (Colton & Covert, 2007, p. 129).
- **Pilot testing:** After program activity survey instruments have performed satisfactorily in readability and comprehensibility testing, the next phase is pilot testing with a sample of the target population that will yield statistically significant data, a random sample of at least 200 respondents (Komrey and Bacon, 1992; Reckase, 2000). The goal of pilot testing is to yield preliminary validity and reliability data to determine if items and the instrument are functioning properly (Haladyna, 2004; Wilson, 2005). Data gleaned from pilot testing will be used to fine-tune items and the instrument in preparation for more complex statistical analysis upon large-scale statistical testing.
- **Large-scale statistical testing:** Instrument testing conducted with a statistically representative sample of responses from a population of interest. In the case of developing scales, large-scale statistical testing provides sufficient data points for exploratory factor analysis (EFA), a multivariate statistical method used to uncover the underlying structure of a relatively large set of variables and is commonly used when developing a scale, a collection of questions used to measure a particular research topic (Fabrigar & Wegener, 2011). EFA is a “large-sample” procedure where generalizable and/or replicable results is a desired outcome (Costello & Osborne, 2005, p.5). This technique is particularly relevant to examining relationships between participant traits and the desired outcomes of NASA Education project activities.
- **Item response approach to constructing measures:** Foundations for testing that address the importance of item development for validity purposes, address item content to align with cognitive processes of instrument respondents, and that acknowledge guidelines for proper instrument development will be utilized in a systematic and rigorous process (DeMars, 2010). Validity will be determined as arising from item development, from statistical study of item responses, and from exploring item response patterns via methods prescribed by Haladyna (2004) and Wilson (2005.)
- **Split-half method:** This method for determining test reliability is an efficient solution to parallel-forms or test/retest methods. Split-half method does not require developing alternate forms of a survey and it places a reduced burden on respondents in comparison to other methods, requiring participation in a single test scenario rather than requiring retesting at a later date. This method involves administering a test to a group of

individuals, dividing the test in half along odd and even item numbers, and then correlating scores on one half of the test with scores on the other half of the test (Davidshofer & Murphy, 2005).



## APPENDIX B: Privacy Policies and Procedures

- Information collected under the purview of this clearance will be maintained in accordance with the Privacy Act of 1974, the e-Government act of 2002, the Federal Records Act, NPR 7100.1, and as applicable, the Freedom of Information Act in order to protect respondents' privacy and the confidentiality of the data collected<sup>5</sup>.
- Data is maintained on secure NASA servers and protected in accordance with NASA regulations at 14 CFR 1212.605.
- Approved security plans are in place for the Office of Education Performance Measurement (OEPM) system in accordance with the Federal Information Security Management Act of 2002 and Office of Management and Budget, Circular A-130, *Management of Federal Information Resources*.
- Only authorized personnel requiring information in the official discharge of their duties are authorized access to records from workstations within the NASA Intranet or via a secure Virtual Private Network (VPN) connection that requires two-factor hardware token authentication.
- OEPM resides in a certified NASA data center and has met strict requirements relating to application security, network security, and backup/recovery of the NASA Office of the Chief Information Officer's security plan.
- Data will be secured and removed from this server and location upon guidelines set out by the NRRS/1392, 68-69. Specific guidelines relevant to the OPEM system include the following:
  - Project management records documenting basic information about projects and/or opportunities, including basic project descriptions, funding amounts and sources, project managers, and NASA Centers, will be destroyed when 10 years old or when no longer needed, whichever is longer.
  - Records of participants (in any format), maintained either as individual files identified by individual name or number, or in aggregated files of multiple participants identified by name or number, including but not limited to application forms, personal information supplied by the individuals, will be destroyed 5 years after the last activity with the file.
  - Survey responses and other feedback (in any format) from project participants and the general public concerning NASA educational programs, including interest area preferences, participant feedback, and reports of experiences in projects, will be destroyed when 10 years old or when no longer needed, whichever is longer.

---

<sup>5</sup> [http://www.nasa.gov/privacy/nasa\\_sorn\\_10EDUA.html](http://www.nasa.gov/privacy/nasa_sorn_10EDUA.html)

The following confidentiality statement, edited per data collection source, will be posted on all data collection screens and instruments, and will be provided to participants in methodological testing activities per NPR 7100.1:

In accordance with the Privacy Act of 1974, as amended (5 U.S.C. 552a), you are hereby notified that this study is sponsored by the National Aeronautics and Space Administration (NASA) Office of Education Infrastructure Division (OEID), under authority of the Government Performance and Results Modernization Act (GPRMA) of 2010 that requires quarterly performance assessment of Government programs for purposes of assessing agency performance and improvement. Your participation is important to the success of this study. The information we collect will help us improve the nature of NASA education project activities and the accuracy with which NASA Office of Education can report to the stakeholders about the project activities offered. The NASA OEID will use the information provided for statistical purposes related to data collection instrument development only and will hold the information in confidence to the full extent permitted by law. Information will be secured and removed from this server and location upon guidelines set out by the NASA Records Retention Schedule 1392, 68-69. Although the following efforts will be taken to ensure confidentiality, there remains a remote risk of personal data becoming identifiable. A non-identifying code number will be assigned to participants' data records, which will be stored in accordance with federal regulatory procedures and accessible only to the investigator. Any use of individual data to illustrate specific assessment results will be labeled in a manner to preserve the participants' anonymity. Any photographs or video of participants involved in the study will not be released without prior written consent. In no way does refusing participation in this instrument development study preclude you from eligibility for NASA education project activities now or in the future.

## List of Tables

Table 1: Respondent Universe and Relevant Numbers .....	2
Table 2: Respondent Universe and Sampling Calculations .....	6