REQUEST FOR APPROVAL under the Generic Clearance for NASA Education Performance Measurement and Evaluation (Testing), OMB Control Number 2700-0159, expiration 04/20/2018

_____

I. TITLE OF INFORMATION COLLECTION:

Office of Education Performance Assessment, Evaluation, and Information Management Data Collection Screens: One Stop Shopping Initiative (OSSI) Student-level Data

II. TYPE OF COLLECTION:

- ☐ Attitude & Behavior Scale
- ☐ Baseline Survey
- ☐ Cognitive Interview Protocol
- ☑ Data Collection Screens
- ☐ Focus Group Protocol
- ☐ Follow-up Survey
- ☐ Satisfaction Survey
- ☑ Usability Protocol

III. GENERAL OVERVIEW: In compliance with the Government Performance and Results Modernization (GPRAMA) Act of 2010, NASA collects data on its educational activities to ensure that progress is being made toward Strategic Objective 2.4 and its associated performance goals, and to collect evidence of the impact of NASA educational programs (NASA Education, 2016, pp. 8-9). The information from this data collection, project activity data, will be used in accordance with the criteria established by NASA for monitoring research and education projects. This information collection is also necessary to provide NASA Education projects with information on participants necessary to determine participant eligibility, selection for activity participation, identify accommodations participants may have, and provide other information necessary for effective activity implementation. Preparations are underway by the NASA Office of Education Performance Assessment, Evaluation, and Information Management (PAEIM) team to consolidate tested, redesigned, and improved project activity data collection screens underneath one unified, modernized IT infrastructure provided by the Office of Education Information Technology (OE IT) team. Rigorous testing of data collection screens underneath this clearance is the first step in PAEIM's plan for a redesigned and consolidated project activity data collection. Currently project activity data is collected via the Office of Education Performance Management (OEPM) application and the NASA Internship, Fellowship, and Scholarship (NIFS) One Stop Shopping Initiative (OSSI). This request pertains to NIFS-OSSI data collection screens only, which constitute student-level data.

IV. INTRODUCTION AND PURPOSE: Project activity data is collected through data collection screens, a set of data fields strategically aligned to the NASA Education architecture of four lines of business and accompanying programs (NASA Education, 2016, pp. 38-47) to facilitate performance measurement, analysis, and accurate reporting of NASA's contributions to STEM education. The project activity data collected characterizes our recruitment pool, educational opportunities, participant pool, participant experiences, our partners, our outputs, and outcomes, and enables effective program administration, communication, and program and project monitoring and review. The data also allows NASA Education to assess portfolio performance by tracking activity outputs, which help identify best practices and constitutes information vital to strategic planning and continuous process improvement. In so doing,

internal users of this data collection use it to make data driven decisions and to monitor and assess performance of the NASA education portfolio of projects administered by NASA field centers and facilities.

While the OE IT team is responsible for the IT infrastructure, the PAEIM team is responsible for the constitution, reliability and validity of NASA Education data collection instruments, to include attitude and behavior scales, surveys, and project activity data collection screens. Testing is an integrated, two-pronged approach supported by subject matter expertise from the PAEIM and OE IT teams. OE IT will execute quality assurance and operability testing and PAEIM will implement usability testing and determine reliability and validity characteristics. The teams' combined efforts will accomplish two objectives:

- Gain an understanding of why particular fields within the data collection screens are yielding inconsistent, unreliable data, meaning understanding the degree to which users are misinterpreting specific questions, and determining what to change to positively improve the user experience and increase the quality of data collected in those fields of concern; and

- Obtain baseline quality assurance and operability (QA & O) measures in support of current use and in support of the efforts to modernize the information technology infrastructure supporting the data collection screens.

The purpose of testing this compendium of data collection screens is to garner information regarding the challenges and success of the user experience to shape and improve the future design of the data collection screens. Specific fields within the data collection screens in question will undergo methodologically rigorous testing to identify issues with the wording of questions and instructions, for instance, and navigation within and between data collection screens. Lastly, testing under this clearance will allow NASA Education to establish a procedure for testing annual updates to the data collection screens that arise in response to changes in project activities, the NASA Education portfolio, Congressional mandates and budget requirements.

V. **RESEARCH DESIGN OVERVIEW:** OE IT QA & O measures are comprised of testing techniques to establish documented evidence that the IT system accomplishes its intended requirements, and validate that the product being developed does what the user is expecting it to do. This is facilitated by validating that requirements are adequately defined, designs and functionality conform to requirements, data is treated correctly, and that test results are accurate. Some phases of dynamic testing techniques may include:

- Unit testing – Validates that individual units of product are working as designed
- Integration testing – Units of product are combined and tested as a group
- Function testing - Involves validating product functionality against defined requirements
- System testing – Testing of both hardware and software on a completely integrated system
- User acceptance testing – End user testing of product functionality

PAEIM's efforts are both qualitative and quantitative in nature, seeking to identify description of impact on user experience as well as measure impact of usability on burden, and determine reliability and validity of the data collection screens. PAEIM may first choose to establish baseline information by applying the following multi-method testing techniques as described briefly:

- 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).
- 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 (Jaaskelainen, 2010). Think-aloud protocols will be especially useful towards the improvement of existing 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.

  Think-aloud protocols will be in the form of a semi-structured qualitative data collection method in which there are consistencies across the ways in which each testing session is introduced and initiated and the ways in which the moderator is trained to intervene when a participant falls silent. The differences will be in the object of the test scenario, the various uses of prompts by the moderator to maintain a steady flow of engagement from the participant, and the length of time (burden) it takes each participant to proceed through a test scenario.

  A test scenario contains the task that a user needs to accomplish during a test session and when a participant completes the task, the session is over (UX Passion, 2016). A usability testing scenario will not include any information about how to accomplish the task. Hence, the usability test will show how the participant accomplishes the task and demonstrates whether the interface, the NIFS-OSSI data collection screens, facilitates completing the scenario. After the test scenario, a comparison of how it was anticipated that the user would complete the task to how they actually completed the task will provide insight into the effectiveness of the NIFS-OSSI architecture and navigation (U.S. Department of Health & Human Services, 2016).

- 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).

Given the user-centric nature of NASA Education data collections, users' requirements should drive initial usability testing for determining the completeness and accuracy with which users achieve their goals, the speed (with accuracy) with which the information solicited can be inputted, how satisfying the NIFS-OSSI user interface is to use, how well the data collection screens prevent errors and help users recover from errors, and how easy it is to learn to use the NIFS-OSSI data collection screens (Quensenbery, 2011).

This formative testing will utilize a small study environment in an iterative process, which includes:

- Identifying a specific user profile (or profiles) for the study;
- Creating scenarios that are task based and goal directed;
- Encouraging users to think out load as they work;
- And testing again to confirm that the changes worked for users (Barnum, 2010).

The time and effort reflected below in the burden estimate chart (Table 2.) indicate time spent across all NIFS-OSSI user experience levels and roles only to provide a perspective on the extent of this testing effort and the depth of information PAEIM is seeking in this testing endeavor to enhance the redesigned data collection screens in partnership with OE IT.

VI. **TIMELINE:** The research literature demonstrates that three to five participants per scenario yields approximately 85% of useful observations (Nielsen & Landauer, 1993, p. 212), and in this instance, number of scenarios is dependent upon the number of different roles and the number of different tasks each role must perform. Therefore, a timeline to completion would be devised in collaboration with the OE IT Manager and the information technology provider or vendor. In addition, usability testing of the NIFS-OSSI data collection screens will take place over the same time period during which the IT infrastructure modernization is underway.

VII. **SAMPLING STRATEGY:** The maximum cost-benefit ratio, derived by weighing costs of testing and the benefits gained, can be achieved with three to five-person participants per scenario. Participants will be randomly solicited from three user experience categories: Novice (less than 1 year of NIFS-OSSI use), Moderate-experience User (between 1 year and 3 years of NIFS-OSSI use), and Super User (More than 3 years of NIFS-OSSI use.) Categories of users are further defined by Roles as indicated in Table 1.:

Table 1. *Categories of participants in usability testing*

Data Collection Source (Roles)	Novice	Moderate-experience	Super User
Center Education (NIFS-OSSI)	2	2	1
Administrator (NIFS-OSSI)	2	2	1
Site Administrator (NIFS-OSSI)	2	2	1
Student (NIFS-OSSI)	2	2	1
Total Participants	8	8	4

VIII. BURDEN HOURS: The scenarios would be determined by data collection source and so each scenario would have no more than five participants. Note, students are the only participants considered members of the public and therefore burden only applies to their participation time in usability testing.

Table 2. *Burden hours for usability testing*

Data Collection Source	Novice			Moderate-experience	Super User		Subtotal hours per Role/Across User Levels	Number of Scenarios for Usability Testing	Total Response Burden in Hours
	Testing Participants per Role	Testing Hrs.		Testing Hrs.		Testing Hrs.			
Center Education (NIFS-OSSI)	2	1.0	2	0.75	1	0.5	4	5	20
Administrator (NIFS-OSSI)	2	1.0	2	0.75	1	0.5	4	14	56
Site Administrator (NIFS-OSSI)	2	1.0	2	0.75	1	0.5	4	2	8
Student (NIFS-OSSI)	2	1.0	2	0.75	1	0.5	4	7	28
Total									***28**

*Burden estimate for testing is solely based on that as determined by the number of Student testing participants because they are members of the public whereas all other Participant roles are filled by civil servant employees for whom burden does not apply. Further, given that this testing will be focused within scenarios on data collection fields that traditionally yield inconsistent data, 28 burden hours is a maximum estimate.

IX. DATA CONFIDENTIALITY MEASURES: Any 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.

X. PERSONALLY IDENTIFIABLE INFORMATION:
    1. Is personally identifiable information (PII) collected? ☑Yes ☐No
    2. If yes, will any information that is collected by included in records that are subject to the Privacy Act of 1974? ☑Yes ☐No
    3. If yes, has an up-to-date System of Records Notice (SORN) been published?
        ☑Yes ☐ No
    Published in October 2007, the Applicable System of Records Notice is NASA 10EDUA, NASA Education Program Evaluation System -
    http://www.nasa.gov/privacy/nasa_sorn_10EDUA.html.

    APPLICABLE RECORDS: Submitted data will be retained in accordance with NASA Records Retention Schedule 1, Item 68D. Records will be destroyed or deleted when ten years old, or no longer needed, whichever is longer.

XI. PARTICIPANT SELECTION APPROACH:
Does NASA Education have a respondent sampling plan? ☑Yes ☐ No

If yes, please define the universe of potential respondents. If a sampling plan exists, please describe? **The universe of potential usability testing participants includes NASA NIFS-OSSI Center Education users, Administrators, Site Administrators, Broker Facilitators, and NIFS-OSSI student applicants.**

If no, how will NASA Education identify the potential group of respondents and how will they be selected? **Not applicable.**

XII. INSTRUMENT ADMINISTRATION STRATEGY
Describe the type of Consent: ☑ Active   ☐ Passive

4. How will the information be collected:
☑ Web-based or other forms of Social Media
☐ Telephone
☑ In-person
☐ Mail
☐ Other
5. Will interviewers or facilitators be used?   ☑ Yes ☐No

XIII. DOCUMENTS/INSTRUMENTS ACCOMPANYING THIS REQUEST:
☑ Consent form (Appendix B: Confidentiality, Consent, & Recording Release-Adult)
☐ Instructions
☑ Instrument (Appendix C: NIFS-OSSI Data Collection Screens)
☑ Protocol script (Appendix A: Sample Usability Testing Script)
☐ Other (Specify _____)

XIV. GIFTS OR PAYMENT: ☐ Yes ☑ No

XV. ANNUAL FEDERAL COST: The estimated annual cost to the Federal government is $6,171. The cost is based on an annualized effort of 187 person-hours at the evaluator's rate of $33/hour for administering the usability testing protocols, collecting and analyzing responses, and drafting a report for the Information Technology vendor's use towards development of the modernized system of data collection screens for ultimate approval under the methodological testing generic clearance with OMB Control Number 2700-0159, exp. 04/30/2018.

XVI. CERTIFICATION STATEMENT:
I certify the following to be true:
1. The collection is voluntary.
2. The collection is low burden for respondents and low cost for the Federal Government.
3. The collection is non-controversial and does raise issues of concern to other federal agencies.
4. The results will be made available to other federal agencies upon request, while maintaining confidentiality of the respondents.
5. The collection is targeted to the solicitation of information from respondents who have experience with the program or may have experience with the program in the future.

Sponsor: Dr. Roosevelt Johnson
Title: Deputy Associate Administrator
 Office of Education
Email address or Phone number: roosevelt.y.johnson@nasa.gov
Date: 12/5/16

References

Barnum, C. M. (2010). *Usability testing essentials: Ready, set...test!* Burlington, MA: Morgan Kaufmann.

Colton, D., & Covert, R. W. (2007). *Designing and constructing instruments for social reserch and evaluation.* San Francisco: John Wiley and Sons, Inc.

Jaaskelainen, R. (2010). Think-aloud protocol. In Y. Gambier, & L. Van Doorslaer (Eds.), *Handbook of translation studies* (pp. 371-373). Philadelphia, PA: John Benjamins.

NASA Education. (2016, January). Retrieved from NASA Education Implementation Plan: http://www.nasa.gov/sites/default/files/atoms/files/nasa_education_implementation_plan_ve 4_2015-2017.pdf

Nielsen, J., & Landauer, T. K. (1993). A Mathematical Model of the Finding of Usability Problems. *INTERCHI 93 Conference on Human Factors in Computing Systems* (pp. 206-213). Amsterdam: ACM Press. Retrieved July 26, 2016, from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.544.5899&rep=rep1&type=pdf

Quensenbery, W. (2011, July 25). *WQUsability.* Retrieved from http://wqusability.com/publications.html: http://wqusability.com/handouts/righttechnique-uf2008.pdf

U.S. Department of Health & Human Services. (2016, July 28). *Scenarios*. Retrieved from Usability.gov: https://www.usability.gov/how-to-and-tools/methods/scenarios.html

UX Passion. (2016, July 28). *Usability: What is a test scenario and how to design it*. Retrieved from UX Passion: http://www.uxpassion.com/blog/usability-what-is-test-scenario/